GPS Multipath Mitigation: A Nonlinear Regression Approach

Quoc-Huy Phan[a], Su-Lim Tan[b], Ian McLoughlin[c]

[a] School of Computer Engineering, Nanyang Technological University, 50, Nanyang Avenue, Singapore 639798

[b] Academic Programmes Division, Singapore Institute of Technology, 25 North Bridge Road, EFG Bank Building, Singapore 179104

[c] Earth Observatory of Singapore, Nanyang Technological University, 50, Nanyang Avenue, Singapore 639798

Email: phan0035@ntu.edu.sg

Abstract

Under the assumption that the surrounding environment remains unchanged, multipath contamination of GPS measurements can be formulated as a function of the sidereal repeatable geometry of the satellite with respect to the fixed receiver. Hence, multipath error estimation amounts to a regression problem. We present a method for estimating code multipath error of GPS ground fixed stations. By formulating the multipath estimation as a regression problem, we construct a nonlinear continuous model for estimating multipath error based on well-known sparse kernel regression, e.g. Support Vector Regression (SVR). We will empirically show that the proposed method achieves state-of-the-art performance on code multipath mitigation with 79% reduction on average in terms of standard deviation of multipath error. Furthermore, by simulation, we will also show that the method is robust to other coexisting signals of phenomena, such as seismic signals.

Keywords: GPS, multipath mitigation, nonlinear regression, support vector machine.

Introduction

Multipath is defined as one or more indirect replicas of the line-of-sight signal arriving at the receiver antenna after reflection on objects in the surroundings. It constitutes a major error source that contaminates receiver measurements, causing performance degradation of GPS positioning solutions. The errors induced by multipath are typically up to 15 meters for C/A code (Hoffmann-Wellenhof and Lichtenegger 2001) and up to a few centimeters for carrier phase measurements (Leick 2004). That is significant for precise applications requiring centimeter-

level accuracy. Although many positioning algorithms rely on the carrier phase measurements, accurate code measurements (pseudoranges) are important to a variety of applications.

It is still challenging to either rule out or accurately model multipath errors. Various approaches have been proposed to mitigate multipath in the GPS measurement domain. These approaches can be classified as either frequency-domain or time-domain processing. The methods in the frequency-domain are based on spectral analysis of the multipath error. After converting multipath sequences into the frequency domain, the multipath error can be ruled out by nullifying the spectrum corresponding to its bounded frequency region (Zhang and Bartone 2004a). More recently, many attempts have been made to apply wavelet decomposition to de-noise high-frequency multipath or to extract low-frequency multipath signatures from GPS measurements which are then applied to correct multipath errors in subsequent measurements (Zhang and Bartone 2004b; Elhabiby at el. 2008, Souza at el. 2008a; Souza at el. 2008b). However, a shortcoming of these methods is the tendency to rule out phenomena signals of interest, such as seismic signals, since their spectrum likely overlaps with the spectrum of multipath signals. Other approaches attempt to link signal-to-noise (SNR) measurements, which are independent of phenomena signals, to multipath to obtain a map of the multipath environment (Bilich at el. 2008; Rost and Wanninger 2009). Unfortunately, SNR measurements are not always available to use; furthermore, different GPS receiver types may report SNR ratios in inconsistent units and resolution, making it inapplicable in many situations.

Pertaining to time-domain processing, the pioneering work is credited to the carrier smoothing filter (CSF) (Hwang at el. 1999; Misra and Enge 2006). This technique plays the role of a low-pass filter, using precise but ambiguous carrier phase measurements to smooth noisy code measurements. Nevertheless, it is effective at reducing receiver noise and very high-frequency multipath error but less effective for lower frequency multipath error. The filter-based techniques like band-pass finite impulse response (FIR) filter (Ge at el. 2000; Liu at el. 2009) have also been used but they are of limited value as they require impractical multipath-free reference stations. Furthermore, it also tends to filter out other signals of interest. Another major trend is the modeling approach, which takes advantage of repeatability of multipath sequences on a daily basis at a fixed

station, to stack multiple multipath sequences of consecutive days in order to map the multipath environment (Axelrad at el. 2005; Larson at el. 2007; Zhong at el. 2010). The stacked multipath sequences are then time-shifted with a repeatable period to calibrate multipath errors for subsequent measurements. However, it is not obvious which time shift to use when different satellites are visible at different times of the day. In addition, high-rate data is required to boost its performance (Axelrad at el. 2005).

Our previous work (Phan and Tan 2011) suggests a non-parametric model by posing multipath error as a function of repeatable geometry of the satellite with respect to a fixed receiver, and the function is then determined by learning from historical multipath data. We extend the work with an in-depth analysis of geometrical models of code multipath error. We derive the geometrical model of GPS code multipath error, formulate code multipath error estimation as a nonlinear regression problem, and efficiently solve multipath error estimation by Support Vector Regression (Vapnik 1995; Smola and Schölkopf 2004) to achieve mitigation.

Section II reviews the mathematical models of GPS measurements, followed by the derivation of the geometrical model of code multipath error. Section III presents the theory of Support Vector Regression and the formulation of code multipath estimation as a regression problem. Experimental results and discussions will be presented in Section IV and a conclusion will follow in Section V.

GPS Measurements and Geometrical Model of Code Multipath Error

The mathematical models of GPS measurements will be reviewed. In the sequel, the relationship between satellite-specific code multipath error and geometry of the satellite with respect to a receiver will be explored to derive geometrical model of code multipath error. As a result, the code multipath error can be viewed as a function of satellite geometry which is characterized by azimuth and elevation angles.

GPS Measurements

During visible period of a satellite, the GPS code measurements ρ and carrier phase measurements ϕ are given respectively as (Misra and Enge 2006; Parkinson 1996):

$$\rho_1 = r + c(\delta_u - \delta_s) + I_1 + T + M_1^\rho + \varepsilon_1^\rho \qquad (1)$$

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

where r denotes the true range between a satellite and a receiver, c speed of light, δ_u and δ_s the receiver and satellite clock biases respectively, T the tropospheric delay, I the ionospheric delay, λ the wavelength, and N is the ambiguous integer. The symbols M^ρ and ε^ρ represent code multipath error and random receiver noise while M^ϕ and ε^ϕ have similar meanings for carrier phase measurements. The opposite signs of I in (1) and (2) are due to the fact that the ionosphere affects code and carrier phase measurements equally but in opposite directions (Leick 2004).

Geometrical Model of Code Multipath Error

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. Receiver antenna receives one or more replicas of the direct signal reflected from objects near the line-of-sight path, particularly those in the vicinity of the receiver. As a result, the receiver will track the composite signal that is a combination of the direct signal 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, and $\alpha = A_m / A_d \leq 1$ 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 center. Let β denote the reflection angle relative to the reflecting surface, the induced code multipath error is given as (Axelrad at el. 2005):

$$M^{\rho} = \frac{\alpha\delta\cos\psi}{1+\alpha\cos\psi} \qquad (3)$$

In order to obtain the geometric model, we need to relate multipath error with geometric parameters, i.e. azimuth and elevation angles. Assuming that the satellite, antenna, and normal vector to the reflecting surface are all in a same plane, multipath reflections fall into two categories: *forward-scatter* and *backscatter* (Bilich and Larson 2007) as illustrated in Fig. 1.
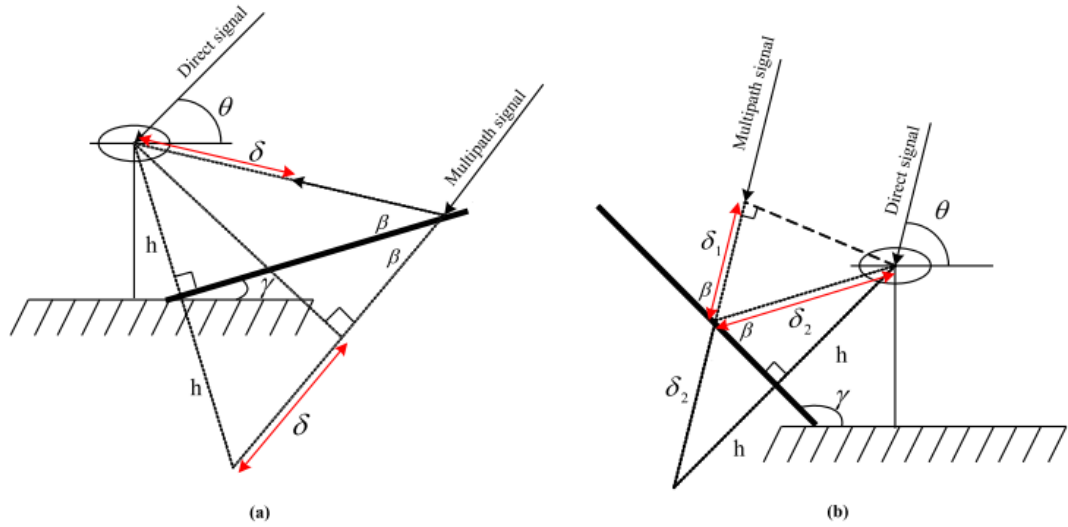
Fig. 1 Multipath reflections: (a) Forward-scatter, (b) Backscatter

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

$$\delta = 2h\sin\beta \qquad (4)$$

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

Furthermore, using the convention of angles measured counter-clockwise over the interval $[0°\ 180°]$, $\beta + \gamma = \theta$ as shown in Fig. 1a whereas $\beta + \theta = \gamma$ in Fig. 1b. To generalize:

$$\beta = |\theta - \gamma| \qquad (6)$$

Substituting (4), (5), and (6) into (3), the code multipath equations corresponding to one reflecting signal can be re-written:

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

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

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

Under the assumption that the multipath environment is held fixed, the quantities α_i, h_i, γ_i are constant. That is, code multipath M^P is a complicated function 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.

Code Multipath Estimation with Nonlinear Support Vector Regression

The theory of Support Vector Regression (SVR) for function approximation will be briefly overviewed. Thereafter, given the derived geometric model, code multipath estimation is formulated as a nonlinear regression problem which is then solved by the SVR framework.

Support Vector Regression

Given a set of N training samples $\{(x_1, y_1), \ldots, (x_N, y_N)\} \subset X \times \Re$, for a regression problem, our goal is to learn a function $f : X \mapsto \Re$ to map from an observation vector x to an estimate of target $\hat{y}$. Due to noise in the training data, it is unlikely that $f(x_i)$ will 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.

Among various kinds of regression models, SVR (Vapnik 1995; Smola and Schölkopf 2004) are well-known to be more robust and accurate than traditional models, such as neural networks (Bishop 1996). We employ ε-SVR (Smola and Schölkopf 2004) which is most commonly used. In ε-SVR, the estimate $f(x_i)$ has at most ε deviation from the actually target y_i 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.

We denote the regression function $f(x) = \langle w, \varphi(x) \rangle + b$ where $\varphi : X \mapsto F$ is a possibly mapping function that maps a data point x in the input space X to a feature space F, w is the weight vector in the feature space, $b \in \Re$ is a bias term, and $\langle \cdot, \cdot \rangle$ denotes the dot product. The *ε-insensitive loss function* 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:

$$L(f(x), y) = \begin{cases} 0 & \text{if } |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon & \text{otherwise} \end{cases} \qquad (9)$$

$f(x)$ can be solved through the following optimization problem (Smola and Schölkopf 2004):

$$\text{minimize} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)$$

$$\text{subject to} \quad \begin{cases} y_i - \langle w, \varphi(x_i) \rangle - b \leq \varepsilon + \xi_i \\ \langle w, \varphi(x_i) \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* > 0, \ i = 1, \ldots, N \end{cases} \qquad (10)$$

where $\varepsilon > 0$ is the parameter in the ε-insensitive loss function that controls the accuracy of the regressor. The constant $C > 0$ adjusts a trade-off between model complexity and training error. $\xi = \{\xi_1, \ldots, \xi_N\} \in \Re^N$ and $\xi^* = \{\xi_1^*, \ldots, \xi_N^*\} \in \Re^N$ are slack variables allowing errors around the regression function for exceeding the target value by more than ε and for being below the target value by more than ε, respectively. Fig. 2 depicts equation (10) graphically.
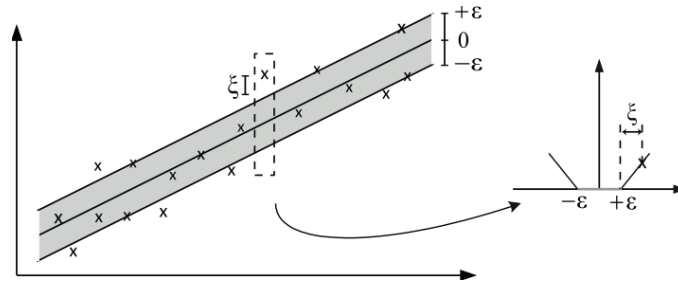
Fig. 2 The ε-insensitive loss function setting with soft margin corresponds for a linear support vector machine. Image from (Schölkopf and Smola 2002).

By introducing Lagrange multipliers (Schölkopf and Smola 2002) to the optimization problem (10), the following dual form can be obtained:

$$\text{maximize} \quad -\varepsilon \sum_{i=1}^{N} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) y_i$$

$$-\frac{1}{2} \sum_{i,j=1}^{N} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \kappa(x_i, x_j) \quad (11)$$

$$\text{subject to} \quad \begin{cases} \alpha_i, \alpha_j \in [0, C] \\ \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) = 0 \end{cases}$$

where α_i, α_i^* are Lagrange multipliers and $C, \varepsilon > 0$ are chosen *a priori*. $\kappa(x_i, x_j)$ is a kernel function which represents the inner product $\langle \varphi(x_i), \varphi(x_j) \rangle$ in feature space. This leads to the regression solution, which is so-called *support vector* (SV) expansion, taking the form:

$$f(x) = \sum_{i=1}^{N_{SV}} (\alpha_i - \alpha_i^*) \langle \varphi(x), \varphi(x_i) \rangle + b$$

$$= \sum_{i=1}^{N_{SV}} (\alpha_i - \alpha_i^*) \kappa(x, x_i) + b \quad (12)$$

where N_{SV} is the number of support vectors. Note that all training samples fall inside the ε-tube, i.e. $|f(x_i) - y_i| < \varepsilon$, will have the Lagrange multipliers α_i, α_i^* vanished. Only for samples with $|f(x_i) - y_i| \geq \varepsilon$, which is so-called support vectors, the Lagrange multipliers may be non-zero (Schölkopf and Smola 2002). The w is implicitly contained in (12) without the need to compute explicitly when evaluating the function. It can be completely described as a linear combination of the training patterns x_i:

$$w = \sum_{i=1}^{N_{SV}} (\alpha_i - \alpha_i^*) \varphi(x_i) \quad (13)$$

and b can be computed as follows:

$$b = y_i - \langle w, \varphi(x_i) \rangle - \varepsilon \ \text{ for } \ \alpha_i \in (0, C)$$

$$b = y_i - \langle w, \varphi(x_i) \rangle - \varepsilon \ \text{ for } \ \alpha_i^* \in (0, C) \quad (14)$$

It can be seen that $f(x_i)$ only depends support vectors through the representation of the kernel function κ. Any function can be chosen as the kernel function providing that it satisfies Mercer's theorem (Schölkopf and Smola 2002).

Code Multipath Error Estimation

From (8), estimating code multipath error is equivalent to the function approximation. Therefore, satellite-specific code multipath error estimation can be posed as a regression problem with geometric features: azimuth and elevation angles. Specifically, estimating code multipath error is a 2-dimensional regression setting:

$$M^\rho = f(\theta, \mu) \qquad (15)$$

We will cast this problem into ε-SVR framework with the input vector $x \equiv (\theta, \mu) \in \Re^2$ and the target output $y \equiv M^\rho \in \Re$. The radial basis function (RBF) kernel (Schölkopf and Smola 2002) is reasonably chosen due to its ability to handle nonlinearity:

$$\kappa(x_i, x_j) = \exp(-\eta \|x_i - x_j\|^2) \qquad (16)$$

where η is the kernel bandwidth. Learning from training data, the multipath estimator will estimate code multipath error for each visible satellite at each epoch. The estimated code multipath error will be used to correct the code measurements to achieve multipath mitigation.

The rationale behind this method is the observation that a fixed receiver experiences highly sidereal day-to-day correlation of satellite-receiver geometry and multipath error. These observations have been used for multipath research, specifically the multipath stacking-based methods (Axelrad at el. 2005, Bock at el. 2000; Zhong at el. 2010). In this work, our proposed method leverages this observation with a machine learning approach to approximate the function of the repeatable multipath error. We train the multipath estimators using historical data of a few days before using them to estimate multipath error and correct measurements of the successive day.

Experiments

We will describe experiments conducted to train multipath estimators using historical data and subsequently use them for multipath estimation. We demonstrate that our approach outperforms state-of-the-art results in multipath

mitigation in term of standard deviation. The advantages of the exploited methods will be also discussed.

Experimental Data Set

The data set used for evaluation was recorded at 0.1 Hz at the station equipped with a dual-channel Trimble NetRS receiver on the rooftop of N2 building in Nanyang Technological University (NTU) campus during five consecutive days: from the day of year (DOY) 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. 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 Fig. 3, 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.

Fig. 3 The observation site on the rooftop of N2 building (NTU campus)

Extraction of Training Data

Training data is the key ingredient for any learning algorithm. Fortunately, the target code multipath error can be extracted from code measurements by batch-processing a sequence of contiguous data (i.e. 1-day data). By differencing (1) and (2), we obtain the code-minus-carrier (CmC) as follows:

$$\chi = 2I - N\lambda + M^{\rho} + \varepsilon^{\rho} \qquad (17)$$

Note that carrier phase multipath error and receiver noise are far smaller than those of code measurements, i.e. $M^\rho >> M^\phi \approx 0$ and $\varepsilon^\rho >> \varepsilon^\phi \approx 0$; therefore, they can be ignored for simplicity. Since N is a constant bias as long as no cycle slip occurs and $(M^\rho + \varepsilon^\rho)$ is zero-mean (Misra and Enge 2006), N can be computed by averaging over a whole orbit arc and subtracted from χ at each epoch. Pertaining to ionospheric delay I, it can be removed by fitting a low frequency polynomial to the CmC sequence in case of single-frequency receiver (Zhang and Bartone 2004a) or forming ionosphere-free combinations in case of dual-frequency receiver (Misra and Enge 2006). We used ionosphere-free combinations throughout this experiment. Eventually, only code multipath error and receiver noise remain in CmC χ.

In order to illustrate the day-to-day repeatability of code multipath error, Fig. 4a and Fig. 4b plots the PRN 12 code multipath sequences extracted from the recorded data set during four days from DOY 306 to DOY 309. The correlation is more clearly revealed after we smoothed the sequences with a 50-second CSF to largely remove high-frequency noise. The CSF is implemented in recursive form as follow (Zhang and Bartone 2004b; Misra and Enge 2006):

$$\bar{\chi}(t+1) = \frac{\tau-1}{\tau}\bar{\chi}(t) + \frac{1}{\tau}\chi(t+1) \qquad (18)$$

where τ, χ, $\bar{\chi}$ and t represent smoothing window, CmC, smoothed CmC and epoch index respectively. The original multipath sequences are plotted in blue and CSF-smoothed multipath sequences are in red. Day-to-day correlation of the two multipath sequences is numerically evaluated by their pair-wise normalized cross-correlation values which are tabulated in Table 1. The blue and red values correspond to the original and smoothed multipath sequences respectively. The cross-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.
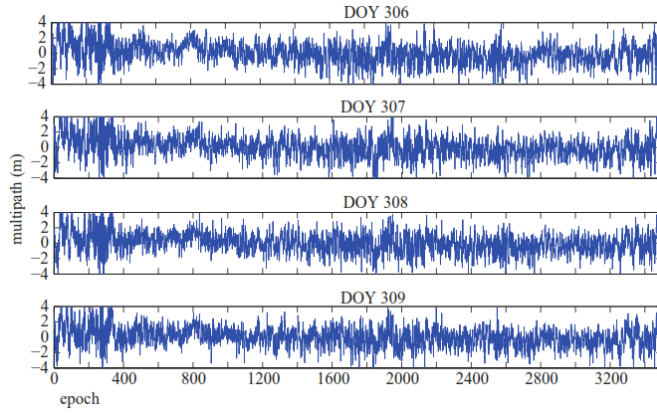
(a) Original
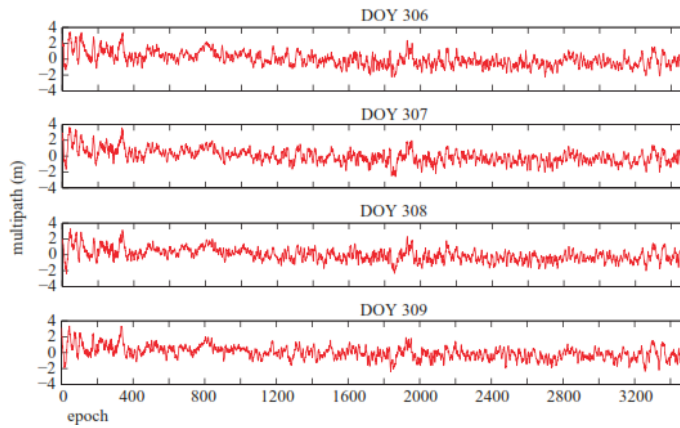
(b) 50-second CSF-smoothed

Fig. 4 PRN 12 multipath sequences in four sidereal days from DOY306 to DOY 309 of 2010

Table 1 Normalized cross-correlation of PRN 12 multipath sequences

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

Training Code Multipath Estimators

For each satellite, the training data is prepared using 4-day data from DOY 306 to DOY 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 with respect to the receiver, which are inputs for training, are computed from the broadcast navigation data. For the desired multipath outputs, after being detached from observation data, the CmC sequences containing multipath errors are filtered with CSF as equation 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 (Dickman at el. 2003; Zhang and Bartone 2004b). Thus, the receiver noise was significantly reduced without removing the multipath which was to be quantified. This smoothing operation helps to clearly expose multipath patterns; as a result, to enhance the estimators' generalization.

Scaling is applied to the training data before feeding them for training. 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 (available at 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*. Since it is not known beforehand which parameter values are best for a given problem; some kind of model selection (parameter search) must be done. For η and C, grid search and cross-validation were applied. First, the coarse grid-search was performed on exponentially growing sequences of η and C followed by cross-validation for each pair of (η, C) with a fixed ε. The result with the best 6-fold cross-validation accuracy was picked. With this strategy, better regions on the grid can be identified and finer grid-search on that region is then conducted. 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 (12) are found for multipath estimation function f of each satellite. Each trained estimator f should estimate multipath error when presented with a new observation of azimuth /elevation angles thereafter.

Experimental Results

In order to evaluate the performance of code multipath estimators, the proper multipath correction is directly applied to code measurements of DOY 310. Note

that the inputs must be scaled as they were during the training phase and the estimated multipath errors need to be de-scaled.

For the sake of demonstration, Fig. 5 presents pseudorange multipath errors of DOY 310 of 2010 and responses of the multipath estimator corresponding to PRN 12. As observed, the errors are significantly reduced after being corrected with the responses of the SVR estimators.
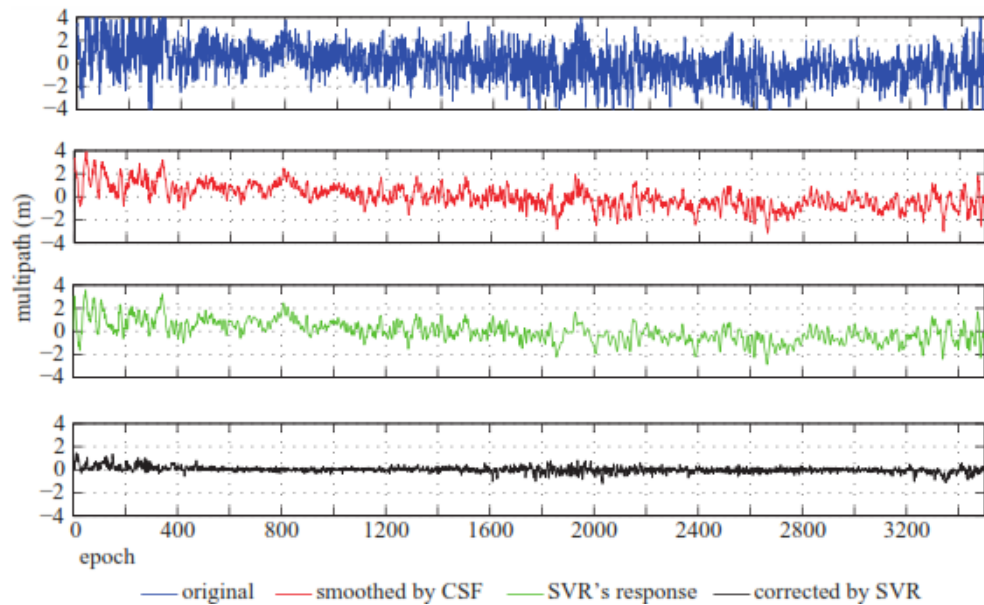

Fig. 5 Original pseudorange multipath, CSF-smoothed pseudorange multipath, SVR-estimator response, and SVR-corrected pseudorange multipath of PRN 12 in DOY 310 of 2010

Performance of all SVR multipath estimators corresponding to the visible satellites is presented in Fig. 6 where multipath mitigation is measured in term of standard deviation of multipath errors. With correction by SVR estimation, the percentages of multipath error reduction range from 67.8% with PRN 13 to 91% with PRN 30. On average, calibrating the data with CSF followed by the SVR estimator gains an improvement from 36.7% to 79.0%. Performance of the multipath estimators likely depends upon how fast the reflecting surfaces along the propagation direction change. Since the environmental changes are expected to be different for different propagation directions of the GPS satellites, the variation in performance of the SVR estimators as seen in Fig. 6 is understandable.

The goodness of the corrections is additionally illustrated in the position domain. Fig. 7 shows the variation of the solved positions from the nominal position of the receiver. Weighted Least Mean Square single-point positioning (Misra and Enge 2006) with broadcast navigation data was applied to the data of

DOY 310. The measurements were only corrected with multipath estimation. 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 2.
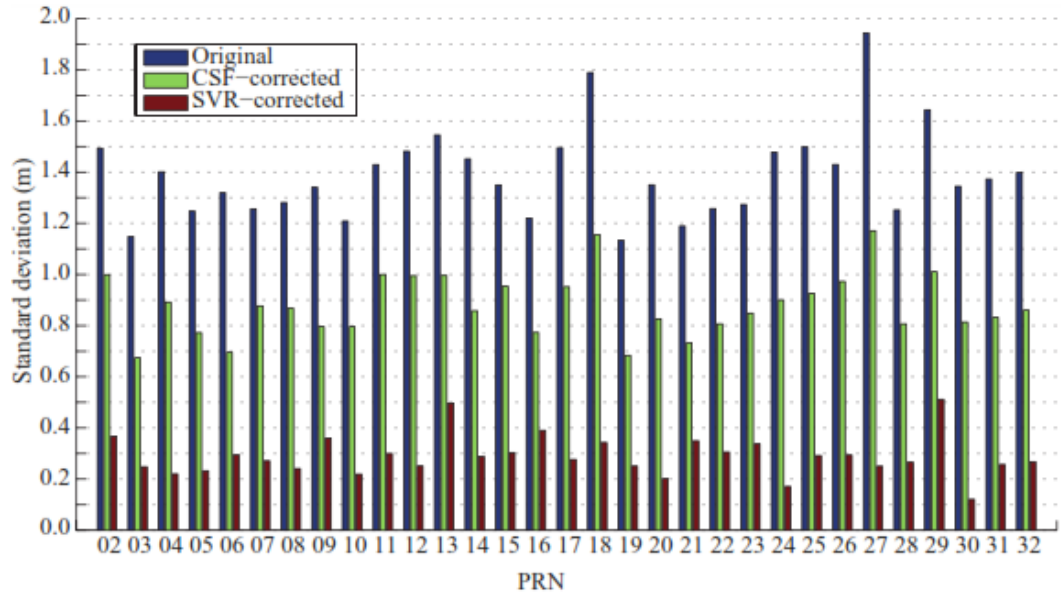
Fig. 6 Code multipath standard deviation of original, CSF-corrected, and SVR-corrected data sets

· original data   · data smoothed with CSF   · data corrected with SVR
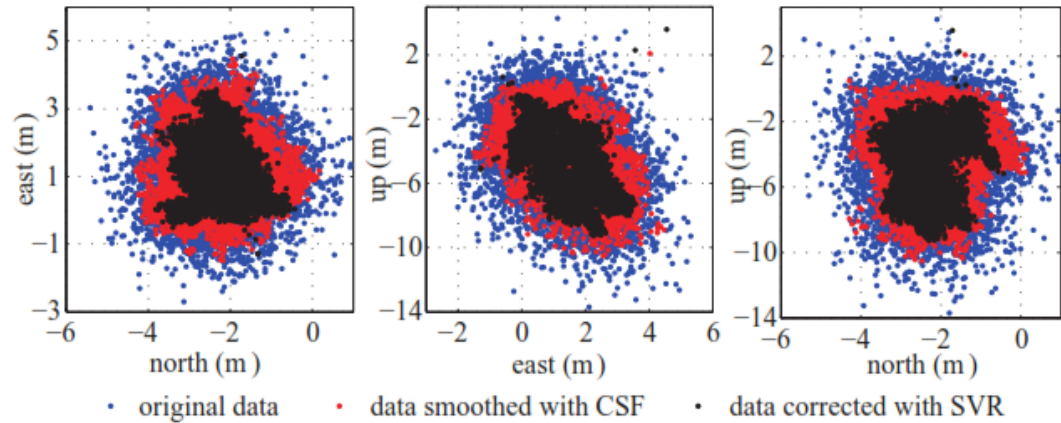
Fig. 7 Positioning solutions on original, CSF-smoothed and SVR-corrected data sets of DOY 310 of 2010

Table 2 Standard deviation (m) of coordinate time series. The values in brackets are the percentages of standard deviation reduction.

	Original	CSF-corrected	SVR-corrected
North	0.9136	0.7223 (20.9)	0.5902 (35.4)
East	1.2180	0.9997 (17.9)	0.9033 (25.9)
Up	2.6069	2.1715 (16.7)	1.9496 (25.2)

Discussions

For comparison, Table 3 tabulates the numerical performance of different methods in term of percentage of reduced multipath error. The performance of CSF is reported with a 100-second smoothing window (Misra and Enge 2006). The performance range of the frequency analysis method using Fast Fourier Transform (FFT) (Zhang and Bartone 2004a) is reported with block sizes of 256 and 512 respectively whilst the block size of the wavelet analysis method (Zhang and Bartone 2004b) is 100 seconds. In particular, the performance of the FIR filter method (Ge at el. 2000) is unreliable as only one satellite (PRN 9) was used for analysis. It is shown that SVR estimators significantly outperform other reported methods with a large margin. 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 get more and more severe over time, performance of the estimators would temporally degrade. Therefore, the multipath estimators need to be equipped with adaptability, which has not been addressed so far.

Table 3 Performance comparison in percentages of multipath error reduction of different code multipath mitigation methods.

CSF (Misra and Enge 2006)	58
FIR Filter (Ge at el. 2000)	75
FFT (Zhang and Bartone 2004a)	50 – 70
Wavelet (Zhang and Bartone 2004b)	55 - 65
SVR estimator	**68 - 91**

Our method modeling multipath errors as functions of continuous variables in azimuth and elevation angles does not need to determine the time-shifting period as in multipath stacking-based methods (Axelrad at el. 2005; Choi at el. 2004; Larson at el. 2007). 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 sparsity of the support vector solution (Smola and Schölkopf 2004), the multipath estimators just use a subset of training data for computation, being simpler while requiring less storage. All of these imply better scalability.

One distinct advantage of the proposed approach is that it is able to preserve other signals of interest such as deformation caused by earthquakes. This is achieved by training the models with data on normal days without displacement before using them to correct data on the subsequent day when a phenomenon occurs. For the purpose of demonstration, we simulated an event to add to the PRN 12 pseudorange measurements of DOY 310. The event signal is given as:

$$e(t) = 2\cos(\frac{\pi}{10}t + \pi) + \cos(\frac{\pi}{15}t) \qquad (19)$$

Since signals interest are usually of low frequency (Bock at el. 2000; Ge at el. 2000), the frequencies of the simulated event were chosen to exhibit diminishing effects of CSF, which is a low-pass filter (Hwang at el. 1999; Zhang and Bartone 2004b). The PRN 12 pseudorange multipath sequence was smoothed by 50-second CSF and then corrected by the trained PRN 12 SVR estimator. As shown in Fig. 8, 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 sidereal daily basis (Choi at el. 2004; Liu at el. 2009). Therefore, the proposed method could be extended to apply to carrier phase multipath mitigation.

Conclusion

A nonlinear regression approach has been proposed to address the GPS pseudorange multipath mitigation problem for fixed stations. Based on analysis of geometry of multipath signal reflections, the geometrical model of code multipath error is developed. More specifically, code multipath errors corresponding to a satellite are mathematically formulated as a function of the satellite geometry with respect to a receiver which is parameterized by azimuth and elevation angles. As a result, the problem of code multipath error estimation amounts to regression problem where the multipath functions are approximated using 4-day training data using ε-SVM algorithm. Finally, the trained multipath estimators are employed to correct measurements of successive day. The proposed method demonstrates good performance and is scalable to data rate. Furthermore, the multipath estimators do not affect the simulated signal of phenomena.
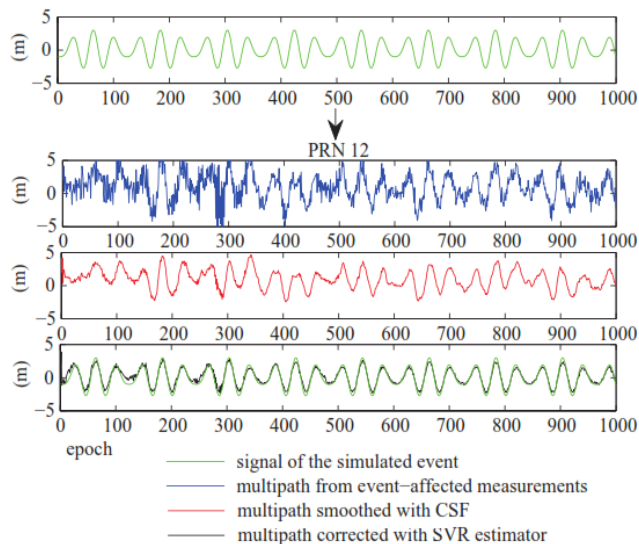
Fig. 8 Simulation of event signal added to pseudorange measurements of PRN 12. It can be seen that the event signal is indeed left intact from correction of the PRN 12 SVR estimator

Acknowledgment

References

Axelrad P, Larson K, Jones B (2005) Use of the correct satellite repeat period to characterize and reduce site-specific multipath errors. In: Proc ION GNSS 2005, Long Beach, CA, pp 2638-2648.

Bilich A, Larson KM (2007) Mapping the GPS multipath environment using the signal-to-noise ratio (SNR). Radio Sci. doi: 10.1029/2007RS003652.

Bilich A, Larson KM, Axelrad P (2008) Modeling GPS phase multipath with SNR: Case study from Salar de Uyuni, Bolivia. J Geophys Res. doi: 10.1029/2007JB005194.

Bishop CM (1996), Neural Networks for Pattern Recognition. Oxford University Press.

Bock Y, Nikolaidis R, de Jonge PJ, Bevis M (2000) Instantaneous geodetic positioning at medium distances with the global positioning system. J Geophys Res 105: 28223-28254.

Choi K, Bilich A, Larson KM, Axelrad P (2004) Modified sidereal filtering: Implications for high-rate GPS positioning. Geophys Res Lett. doi: 10.1029/2004GL021621.

Dickman J, Bartone C, Zhang Y, Thornburg B (2003) Characterization and performance of a prototype wideband airport pseudolite multipath limiting antenna for the local area augmentation system. In: Proc ION NTM2003, Anaheim, CA, pp 783-793.

Elhabiby M, El-Ghazouly A, El-Sheimy N (2008) A new wavelet-based multipath mitigation technique. In: Proc ION GNSS 2008, Savannah, GA, pp 625-631.

Ge L, Han S, Rizos C (2000) Multipath mitigation of continuous GPS measurements using an adaptive filter. GPS Solut 4:19-30.

Hoffmann-Wellenhof B, Lichtenegger H, Collins J (2001) GPS: Theory and Practice. Springer-Verlag, Wien.

Hwang PY, McGraw GA, Bader JR (1999) Enhanced differential GPS carrier-smoothed code processing using dual frequency measurements. Navig 46:127-138.

Leick A (2004) GPS Satellite Surveying, 3rd edn. Wiley, New York.

Liu H, Li X, Ge L, Rizos C, Wang F (2009) Variable length LMS adaptive filter for carrier phase multipath mitigation. GPS Solut 15:29-38.

Misra P, Enge P (2006) Global Positioning System: Signals, Measurements, and Performance, 2nd edn. IGanga-Jamuna Press, Massachusetts.

Parkinson B (1996) Introduction and heritage of NAVSTAR, the Global Positioning System. In: Parkinson B, Spilker J, Axelrad P, Enge P (ed) Global Positioning System: Theory and Applications, vol. 1. American Institute of Aeronautics and Astronautics, Washington DC, pp 3-28.

Phan QH, Tan SL (2011) Mitigation of GPS periodic multipath using nonlinear regression. In: Proc EUSIPCO 2011, Barcelona, Spain, pp 1795-1799.

Rost C, Wanninger L (2009) Carrier phase multipath mitigation based on GNSS signal quality measurements. J Appl Geod. doi: 10.1515/JAG.2009.009.

Schölkopf B, Smola AJ (2002), Learning with Kernels. MIT Press.

Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. J Stat Comput 14:199-222.

Souza EM, Monico JFG, Polezel WGC, Pagamisse A (2008a), An effective Wavelet method to detect and mitigate low-frequency multipath effects, VI Hotine-Marussi Symp Theor Comput Geod, Int Assoc Geod Symp. doi: 10.1007/978-3-540-74584-6_28.

Souza EM, Monico JFG, Polezel WGC, Pagamisse A (2008b), Spectral analysis and low-frequency multipath mitigation for kinematic applications. In: Proc IEEE/ION PLANS 2008, Monterey, CA, pp 413-417.

Vapnik V (1995), The Nature of Statistical Learning Theory. Springer-Verlag, New York.

Zhang Y, Bartone C (2004a) Multipath mitigation in frequency domain. In: Proc IEEE/ION PLANS 2004, Athens, OH, pp 486-495.

Zhang Y, Bartone C (2004b), Real-time multipath mitigation with WaveSmooth[TM] technique using wavelets. In: Proc ION GNSS 2004, Long Beach, CA, pp 1181-1194.

Zhong P, Ding X, Yuan L, Xu Y, Kwok K, Chen Y (2010) Sidereal filtering based on single differences for mitigating GPS multipath effects on short baselines. J of Geod. doi: 10.1007/s00190-009-0352-z.

Biography   about 50 words per author

Quoc-Huy Phan obtained his BSc in Computer Science from Ho Chi Minh City University of Science, Vietnam in 2007. He is currently a graduate student at School of Computer Engineering, Nanyang Technological University where he is working on embedded systems, machine learning and its applications.

Su-Lim Tan received his BSc in Applied Science from Nanyang Technological University (NTU) in 1999 and PhD degree from the University of Warwick, U.K. in 2006. Currently, he is assistant professor at School of Computer Engineering at NTU. His research interest mainly focused around embedded system designs, interactive digital media, and intelligent/smart systems.

Ian McLoughlin received his BEng and PhD from the University of Birmingham in 1991 and 1997 respectively. After spending over a decade working in the electronics R&D industry, he is currently an associate professor in the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include embedded systems, speech and audio, satellite-based electronics and high-reliability computation. In 2009, he was co-founder of the Earth Observatory of Singapore, responsible for computation, communications and data processing for earth observation projects in the South East Asian region.