Many positioning applications utilize global navigation satellite systems (GNSS) derived position estimates for stationary positions. Inexpensive navigation-grade receivers provide estimates within a few meters in relatively open skies, while more specialized devices, typically distinguished by specialized antenna design and additional post-processing can achieve sub-meter accuracy. These latter devices can be two orders of magnitude more expensive than navigation-grade receivers but are still subject to measurement error due to severe multipath in built-up areas.

In our experiments we post-process positions computed by an inexpensive receiver by applying wavelet filtering following by clustering and characterization. This produces a reliable and significant reduction in variance of the estimate, a normalization of the data scatter-distribution and a characterization of the estimate that is amenable to a wider range of statistical comparisons and tests than would be possible for unfiltered, highly non-Gaussian distributions, especially as occur in urban canyon circumstances.

**Keywords:** GNSS; Urban canyon; Multipath mitigation; Wavelets; k-means; RAIM

1. Introduction

Ongoing developments in GNSS space segment (Galileo and GPS modernization) are poised to provide significantly more and better ranging signals for positioning applications. Recent innovation in high-sensitivity receiver technology (HSGNSS) enables the acquisition of attenuated and obstructed signals. These additional signals dramatically lower the probability of a gap [5,6,7] (loss of lock on enough signals to compute a position) in challenging signal environments such as in “urban canyon”, heavy foliage, indoors, etc. While inertial navigation may fill in those gaps in dynamic applications (navigation, logistics tracking), it cannot help stationary or near-stationary applications such as survey, E911, asset or personnel location, or metered parking.

HSGNSS signal measurements are biased and especially noisy due to excessive multipath and low-power signals [2]. Taken together, GPS modernization, Galileo and HSGNSS, means the potential opportunity of many more applications, but generally in harsh signal environments. Specific noise sources are entirely dependent on conditions local to the antenna of the receiver in question and are not addressable by augmentation such as differential GPS (DGPS) or wide area augmentation systems (WAAS), or broad-area correction, such as atmospheric, etc. Even traditional receiver autonomous integrity monitoring (RAIM) has diminished utility since it was developed for signal environments with an assumption of zero or one fault in a field of 5 to 11 signals. We can now project near-future, integrated GPS/Galileo applications with 4 to 22 signals in harsh environments where many or all signals are disturbed.

To tackle these harsher signal environments, new antenna designs [2] and new fault detection and elimination techniques (FDE) that extend RAIM approaches [6,7,8] are being developed. Specialized antennas add system costs and the FDE techniques are computationally complex so that they may be impractical for larger signal sets.

This paper describes an alternate approach: a process that includes wavelet filtering, weighted clustering and characterization of position estimates from a stationary receiver. This approach results in reduced variance of the estimate and a normalization of the data-scatter which, in turn, provides an inexpensive method for applications that require accuracy of 1-2m for short-dwell readings (under ten minutes) in many multipath circumstances. As space segment improvements (Galileo, GPS modernization) and receiver design improvements (high sensitivity) continue to come on-stream, multipath mitigation such as we propose here tends to reduce the relative difference in accuracy between open skies and urban canyon.

The next section of this paper describes our experimental methods, including data collection and processing algorithms. The third section describes and demonstrates our results for each of wavelet filtering, widowing and clustering.

2. Experimental Methods

2.1. Data Collection

To support a variety of experiments, we gathered street-level, urban canyon, carrier phase and position data at multiple locations in downtown Toronto (Canada). Four sites were selected to represent distinct levels of urban canyon effects ranging from moderate to extreme multipath interference. At
each location we collected ten 15-minute samples over five sidereal days, for a total of forty 900-second data sets.

Figure 1 shows the data collection setup we used.

For this particular experiment, we simply used the 3D position estimates generated by the receiver without consideration for outlier removal. Fig 2 shows a typical sample showing high positioning variability. Fig 3 demonstrates that even the geometric mean of a 15-minute sample can be highly variable in severe multipath. We wish to mitigate both forms of variability.

Figure 1: Data collection equipment consisted of: u-blox TM-LP 15 (not HS) evaluation kit with u-center ANTARIS software and a laptop. An active antenna was mounted on a portable antenna mount 1m above the ground. We did not use an external ground plane.

Figure 2: 15 minutes of data collected with the equipment in Fig 1. This is typical of about half of our street-level readings in downtown Toronto.

Figure 3: We sampled the same locations in 15-minute samples over 5 sidereal days. The geometric mean of each of these samples can drift considerably. At this location, a spread of about 40 meters in both Easting and Northing is apparent over the 10 samples taken.

2.2. Processing Algorithms

Our process comprises of two fundamental steps: filtering using wavelet analysis, and an inverse-variance weighted estimate of the mean position using either a moving window or a k-means algorithm to cluster the data.

Positions from receiver | Wavelet filtering | moving window variance weighting | or | K-means variance weighting | Gaussified data scatter with lower variance

Since multipath error is a time varying process, wavelet analysis can be used effectively to mitigate its effects. We tested various wavelets including Daubechies, Coiflets, Symlets, Morlet, and Meyer. Although the results from Symlets and Daubechies were very similar, the analysis was carried out using the ‘Daubechies order 7 (db7)’ filter and wavelet coefficients were modified based on thresholding [4]. Outlier removal was applied to the wavelet output by excluding all points exceeding $3\sigma$ from the mean of the filtered data (where $\sigma$ is the standard deviation).

Rather than simply computing the geometric mean of the wavelet filtered data as a new position estimate, a subsequent, independent process was applied to the position data output from the wavelet filter. Noting that the variance of positioning data, especially in urban canyon, is non-stationary (varies with time), we reasoned that weighting each datum inversely with its local variance would tend to suppress the contribution to the mean estimated position from high-velocity data segments. Such data segments can be caused by a satellite rising or falling at the horizon or changing from line-of-sight to non-line of sight multipath (or vice-versa) and other biasing effects.
We tried two ways of estimating local variance: temporal and spatial. Temporal variance weighting is easily achieved by computing the variance of short temporal data segments (windows) and then by inversely weighting the local means of those temporal windows relative to their local variance.

Spatial variance weighting can be achieved via spatial data clustering. Over a 15 minute sample in a harsh signal environment one can observe the spatial non-stationarity of the position estimate as two or more clusters of points in the scatter (fig 4). If we use a statistical clustering algorithm, such as k-means, we would tend to group spatially similar estimates regardless of whether they are temporally adjacent. The mean of each such cluster can then be weighted by the inverse of its variance. k-means is more computationally intensive than a moving temporal window, but it can be expected to perform somewhat better. This is because a cluster is unlikely to span a positioning discontinuity, while a temporal window is more likely to do so.

3. Experimental Results

3.1. Wavelet filtering

The effect of our wavelet filtering was to always reduce variance (fig 4) and to often Gaussify a sample (normalize its data scatter) by reducing both skew and kurtosis (table 1).

![Figure 4: Plots (a) and (b) show raw data (black) and wavelet filtered results (red). A dramatic reduction in variance, as measured in table 1, is apparent.](Image)

![Figure 5: Variance was reliably and significantly reduced by wavelet filtering in every sample tested.](Image)

![Figure 6: Skew and kurtosis are reduced by wavelet filtering, but only in cases of extreme skewness or kurtotic behavior in the raw data. Kurtosis, here, has been normalized to 0 ("i.e., kurtosis proper-3"). This means that wavelet filtering "gaussifies" these scatter plots in the event of highly non-Gaussian distributions, as frequently occurs in urban canyon.](Image)

Table 1: Corresponding to Figure 4, this table shows large reductions in variance both in GPS space (northing and easting), as well as in eigenspace. Skew and kurtosis were calculated in GPS space only and in (a) shows no change and in (b) shows improvement. In future this will be measured in eigenspace, as well.
The output of this wavelet filtering is consistent: lower variance (fig 5) and Gaussifed data (fig 6), centered very nearly at the same geometric mean. However we know that the mean itself “wanders” over time (fig 3), so that the first moment still retains a bias effect that we now wish to reduce.

3.2. Windowing

Recognizing that the variance process for these data sets is non-stationary, we wish to weigh more heavily data segments that are momentarily stationary (low instantaneous velocity) and weigh less heavily data segments that exhibit high instantaneous velocity. While such a process can not necessarily discriminate between multipath and non-multipath contaminated data, it does take advantage of the fact that ranging signals undergoing rapid change in multipath circumstances exhibit more high-velocity bursts, hence we can reduce the impact of these momentary data subsets for a stationary receiver.

For our temporal windowing process, we experimented with several window lengths and window overlaps. We report here using windows of 20 seconds that overlap by 10 seconds. We then inversely weighted each local mean by the local variance and computed a new weighted mean for the full sample as:

\[ \sum_{i=1}^{N} \frac{X_i}{\sigma_i^2} \]

This process has the effect of causing a set of means from a single location to show reduced scatter. In other words, this process tends to remove noise from the mean of a 15-minute data set collected in urban canyon (fig 7).

3.3. k-means clustering

When examining raw GPS data plots, especially the noisier ones, one often sees areas of two or more clusters of data connected by sparse, high-velocity segments. We reasoned that if we could isolate those clusters, calculate local means and once again weight them by their inverse variance we would see an even greater improvement in the ability to reduce the spread in the geometric means, sample-over-sample.

To do this, we applied a k-means clustering algorithm \((k=15)\), calculated the mean and variance for each of these 15 clusters and computed a weighted mean for the entire dataset, as before.

The overall result of this latter approach (fig 8) provides a further improvement over the windowing approach (fig 7) reducing the variance in the “wandering means” (fig 3). By examining the concentric black-red 3σ ellipses one can see a reduction of 20-35%.

4. Conclusions

Wavelet filtering can be used to reduce variance, skew and kurtosis in GPS position data collected by a stationary receiver. Temporal windowing and spatial clustering of that output can be used to further reduce data biases in urban canyon that tend to make even aggregated mean estimates “wander” about their true position.

These experiments, while successful, leave considerable room for refinement. Future work includes: setting dynamic thresholds for wavelet filtering, non-linear treatment of the inverse-weighting for the moving windows, determining \(k\) dynamically for the k-means algorithm, or using fuzzy c-means. Indeed, the fixed, 15-minute sampling period of this experiment can also be dynamic allowing greater accuracy when time/cost permits and more rapid results in locations of modest multipath.

Figure 7: Shows the relative shift in final position estimates for all 40 15-minute datasets in our experiment. There is one black and one red ellipse representing the 3σ bounds for each of four locations and 10 means each calculated from 900 per-second samples for each location. The black points and ellipses are for the wavelet output and the red are the same for the output after the windowing process.

Figure 8: Shows the same information as in fig 7, except that the red data and ellipses represent the output after the k-means process. It is evident in individual results as it is in these summary plots that k-means out-performs the moving window process in our experiment.
Acknowledgements

This work of Ryerson University Multipath Mitigation (RUMM) labs was supported by Ryerson University (Toronto, Canada) and a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References


