Automatic Positioning Data Correction for Sensor-annotated Mobile Videos

Guanfeng Wang, Beomjoo Seo, Roger Zimmermann
School of Computing, National University of Singapore
Singapore 117417
{wanggf, seobj, rogerz}@comp.nus.edu.sg

ABSTRACT

Video associated positioning data has become a useful contextual feature to facilitate analysis and management of media assets in GIS and social media applications. Moreover, with today’s sensor-equipped mobile devices, the location of a camera can be continuously acquired in conjunction with the captured video stream without much difficulty. However, most sensor information collected from mobile devices is not highly accurate due to two main reasons: (a) the varying surrounding environmental conditions during data acquisition, and (b) the use of low-cost, consumer-grade sensors in current mobile devices. In this paper, we enhance the noisy positioning data generated by smartphones during video recording by analyzing typical error patterns for real collected data and introducing two robust algorithms, based on Kalman filtering and weighted linear least square regression, respectively. Our experimental results demonstrate significant benefits of our methods, which help upstream sensor-aided applications to access media content precisely.

Categories and Subject Descriptors
I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Sensor Fusion

General Terms
Algorithms, Experimentation, Performance

Keywords
Location sensors, positioning data, GPS, mobile video

1. INTRODUCTION

Recently an increasing number of GIS and social media applications utilize auxiliary GPS location information for assistance and as complementary feature to improve multimedia content analysis performance. A study by Divvala et al. reported on the contribution of contextual information in challenging object detection tasks [2]. Slaney also concluded that certain information is just not present in the signal and researchers should not overlook the rich meta-data that surrounds a multimedia object, which can help to build better feature analyzers and classifiers [7]. Auxiliary positioning data is also employed by various multimedia processing methods such as video encoding complexity reduction [8], video indexing and tagging [6], video summarization [10, 4], travel recommendations [3] and others. Moreover, the smartphones are becoming an important provider of geo-tagged content and an essential contributor to location-based services (LBS) since almost all present smartphones are equipped with GPS receivers.

However, the limitations of smartphone embedded GPS units are also well known. The best accuracy acquired by GPS can approach below ten meters under excellent conditions. However, conditions are not always favorable due to some factors, such as the surrounding environmental conditions (satellite visibility and multipath reception), the number of satellites in view and the satellite geometric dilution of precision (HDOP, GDOP, PDOP), etc. Moreover, most of the sensors used in mobile devices like smartphones are quite low cost, which may also result in decreased accuracy. Some post-processing algorithms and software solutions have been proposed to enhance data accuracy by a number of researchers. However, these methods require additional sources of data to determine a more accurate position in addition to the GPS measurements, e.g., vehicular ad-hoc network or WLAN information. During the GPS data collection on a smartphone, such information is not always available. Therefore, a post-processing correction method purely based on GPS measurement data itself is desirable. The goal of our approach is to automatically and transparently process the positioning data of sensor-annotated videos and then provide more accurate data to existing media-utilizing applications.

2. REAL SENSOR OBSERVATIONS

To have a better understanding of noisy positioning data and their typical error patterns, we collected and carefully examined more than 80 mobile videos associated with positioning information, publicly available from the GeoVid web site at http://geovid.org. We report our observations and summarize two typical inaccuracy patterns that we observed from those real positioning data.

Very high inaccuracy at start—A standalone GPS system needs orbital information from satellites to calculate the current position and it provides the first location estimation after approximately 30 to 40 seconds. To avoid
empty measurements within the first 30 seconds in location data collection, many mobile devices combine A-GPS and other location services [9] to improve startup performance. However, those assistive systems require additional network resources to provide locations and also use satellites under poor signal conditions. Since network resources are not always available when users begin to record videos, we observed that some location data generated by GPS sensors at the beginning of a location sequence are extremely noisy.

Sudden moderate inaccuracy—GPS operation uses radio signals from satellites. Under very poor signal conditions, for example in a city, these signals may suffer from multipath propagation where signals bounce off buildings, or are weakened by passing through atmospheric conditions, walls or tree covers. Thus, when users encounter these conditions during video recording, some GPS navigation devices without network connections may not be able to report a position due to the fragmentary signal, rendering them unable to function until a clear signal can be received again. As a result, we observed some sudden moderate incorrect location measurements generated throughout some of the location sequence files.

3. POSITIONING DATA CORRECTION

The positioning data that we target with our correction approach represent the conceptual description of the video content based on the geospatial properties of the scene it captures, which is modeled by its field-of-view (FOV, also called the viewable scene) [1]. In our context, GPS data of a mobile video is a time-series dataset consisting of locations and their accuracies. Let \( L = \{l_1, l_2, \ldots, l_n\} \) and \( A = \{a_1, a_2, \ldots, a_n\} \) be the sequences of GPS readings and their corresponding accuracies for the time sequence \( T = \{t_1, t_2, \ldots, t_n\} \). Each position measurement \( l_i \) consists of a GPS coordinate, i.e., longitude \( x_i \) and latitude \( y_i \). We further denote the ground truth of the position sequence data as \( G = \{g_1, g_2, \ldots, g_n\} \). The location measurement error for \( l_i \) is the distance between its true and its measured locations \( \delta_i = \|g_i - l_i\| \). The location error of \( L \) is the average of every sample’s location error, i.e., \( E_L = \frac{1}{n} \sum_{i=1}^{n} \delta_i \).

3.1 Problem Formulation

Given a sequence of positions \( L \) and their related sequences \( T \) and \( A \), find a sequence of estimated position coordinates, \( F : l_i \rightarrow \tilde{\tau}_i \), such that the processed position sequence \( L' = \{\tau_1, \tau_2, \ldots, \tau_n\} \) is more accurate when comparing \( E_L \) with \( E'_{L'_k} \), where \( E'_{L'_k} = \frac{1}{n} \sum_{i=1}^{n} \delta'_i \), and \( \delta'_i \) represents each processed position’s distance to the ground truth. Each \( \tau_i \) also includes an updated longitude \( x'_i \) and latitude \( y'_i \).

An original GPS reading \( l_i \) has an associated accuracy measurement value \( a_i \), which indicates the degree of closeness between \( l_i \) and its true, but unknown position \( g_i \). If the value of \( a_i \) is relatively large, it means that the actual position \( g_i \) is possibly far away from \( l_i \). We utilize the model of location measurement noise with \( l_i \) and \( a_i \) [5], where the probability of the real position data is assumed to be normally distributed with a mean of \( l_i \) and its standard deviation \( \sigma_i \). We set \( \sigma_i^2 = g(a_i) \), where the function \( g \) is monotonically increasing.

Let a small sub-sequence of a given GPS dataset (termed GPS segment or just segment) be \( S_k = \{l_i, l_{i+1}, \ldots, l_j\} \). It has a relatively short duration and its respective moving speed \( v_k \) is assumed constant. The original GPS readings can also be expressed as a series of disjoint segments \( L = \{S_1, S_2, \ldots, S_m\} \) with their corresponding velocity \( V = \{v_1, v_2, \ldots, v_m\} \). For a given segment \( S_k \), we can estimate the accurate position \( \tau_k \) based on the fusion of two sources of data: (1) the raw, noisy measurement \( l_k \) directly from the GPS receiver and (2) the displacement calculation based on the last estimated position \( \tau_{k-1} \), the velocity \( v_k \) in this segment, and the time duration between \( t_k \) and \( t_{k-1} \).

3.2 Kalman Filtering-based Correction

We model the process in accordance with the framework of Kalman filtering. The filter operates recursively on two streams of noisy data to produce an optimal estimate of the underlying positions. We describe the position and velocity of the GPS receiver as a linear state space:

\[
\tau_k = \begin{bmatrix} x_k \\ y_k \\ v_{kx} \\ v_{ky} \end{bmatrix}^T
\]

where \( v_{kx} \) and \( v_{ky} \) are the longitude and latitude components of the velocity \( v_k \). In each segment \( S_k \), \( v_k \) is estimated from less uncertain coordinates and their timestamp information. We define the state transition model \( F_k \) as

\[
F_k = \begin{bmatrix} 1 & 0 & \Delta t_k & 0 \\ 0 & 1 & 0 & \Delta t_k \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},
\]

where \( \Delta t_k \) is the time duration between \( t_k \) and \( t_{k-1} \). We also express the observation model \( H_k \) as

\[
H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.
\]

\( H_k \) maps the true state space into the measured space. For the measurement noise model, we use \( a_k \) to present the covariance matrix \( R_k \) of the observed noise as follows:

\[
R_k = \begin{bmatrix} g(a_k) & 0 \\ 0 & g(a_k) \end{bmatrix}.
\]

Similarly, the covariance of the process noise \( Q_k \) can also be determined by a diagonal matrix, but using the average of \( g(a_k) \), whose corresponding position coordinates \( l_1 \) and timestamp \( t_1 \) were used to estimate \( v_k \) in this segment.

We apply this process model to the recursive estimator in two alternating phases. The first phase is the prediction,
which advances the state until the next scheduled measurement is arriving. Second, we incorporate the measurement value to update the state.

As observed from our real data traces, in many cases the accuracy measurements at the start of a GPS data sequence are worse than those towards the end. To efficiently correct such highly uncertain GPS readings, we start processing positioning data in reverse (i.e., from \( l_j \) to \( l_i \) in segment \( S_k \)) with our recursive algorithm. Finally, after processing each GPS segment \( S_k \) from \( L \), we obtain a series of updated position sequence segments \( S_k' = \{ \tau_i, \tau_{i+1}, \cdots, \tau_j \} \).

The corrected result is composed of this series of segments, \( L' = \{ S_1', S_2', \cdots, S_m' \} \).

### 3.3 Weighted Linear Least Squares Regression-based Correction

Our second correction model is based on a piecewise linear regression analysis. Since we postprocess the GPS sequence data, we can fully utilize both previous and future GPS readings between \( l_i \) to \( l_j \) to estimate the current position \( \tau_i \), where \( i < k < j \). The piecewise linear regression model computes several estimated position data sequences \( S_k' \), which can later be linked into an integrated sequence as corrected output \( L' \).

Within a segment \( S_k \), different GPS readings are associated with varying accuracy measurements. Figure 1 illustrates the concept of linear regression for one segment \( S_k \). For example, position \( l_i \) is accurate to \( a_i = 5 \) meters while for another position \( l_k \), \( a_k = 600 \) meters. This indicates that these regressors have been observed with certain errors which have varying variances. Thus, when calculating the regression estimator for \( S_k \), the contributions of different points with varying accuracy measurements to the regression line should be unequal. With the assumption that these errors are uncorrelated with each other and with the independent variables \( l_i \), we utilize the weighted least squares method to generate estimators \( \hat{\beta}_k \) for each \( S_k \), instead of utilizing the standard regression model.

Since \( x_i \) and \( y_i \) are two independent variables, we estimate the model function parameters \( \hat{\beta}_k \) for longitude and latitude values with respect to time separately. The goal is to find \( \hat{\beta}_k \) for the model which “best” fits the weighted data. By using the weighted least squares method, we need to minimize \( R \), where

\[
R = \sum_{k=i}^{j} W_{kk}r_k^2, \quad r_k = x_k - f(t_k, \hat{\beta}_k). \tag{1}
\]

Here \( r_k \) is the residual defined as the difference between the original measured longitudinal value and the value predicted by the model. The weight \( W_{kk} \) is defined as

\[
W_{kk} = \frac{1}{\sigma_k^2} \tag{2}
\]

where \( \sigma_k \) is the deviation of the measurement noise. It is proven that \( \hat{\beta}_k \) is a best linear unbiased estimator if each weight is equal to the reciprocal of the variance of the measurement. As described in Section 3.1, we modeled the measurement noise as a normal distribution with mean \( x_k \) and standard deviation \( \sigma_k = g(a_k) \) in the longitude dimension.

In this model, measurements \( x_k \) with a high value of \( a_k \) – which indicates high uncertainty – will not have much impact on the regression estimation. Usually, these uncertain measurements reflect outlier GPS locations, which are far away from where the real positions should be. Considering that the regression line is estimated mostly from the confidence data values which are almost consecutive in the temporal domain, we are able to correct these spotty GPS locations to positions that are much closer to their real coordinates. Thus, after we calculate the regression line \( \beta_k \), we update the longitude values based on the following rules

\[
\begin{align*}
x_k' &= x_k, \quad \text{if } a_k < T H_a \text{ and } a_k < r_k \\
x_k' &= f(t_k, \hat{\beta}_k), \quad \text{otherwise}
\end{align*} \tag{3}
\]

Here \( T H_a \) is the accuracy measurement threshold. We only update the latitude value if its corresponding accuracy is measured as being considerably uncertain, or its distance to the projection point on the regression line is less than its accuracy measurement value. We do not use projection points on the regression line for all \( x_k \), because some original longitude values have relatively high confidence with their measurements, and keeping those extremely confident GPS raw data while updating the remaining uncertain data will generate a more accurate position sequence according to the observation that the GPS carrier movement may not follow a standard linear model.

Finally, by computing every longitude value \( x_k' \) in each piece \( S_k' \), we concatenate all the approximate linear pieces together and obtain an updated position sequence \( L' \).

### 4. EXPERIMENTAL EVALUATION

We implemented both of our location data correction algorithms and evaluated them on a set of sensor-annotated videos publicly available at the GeoVid website. We report the accuracy enhancement that we achieved for location measurements from those video clips.

We selected 63 sensor-annotated videos and their sensor datasets among 87 videos retrieved from the GeoVid website through the public API. The 24 not chosen videos contained only a few GPS sample points and had a relatively small recording duration (typically less than one minute). The mobile devices employed during the data collection included various Apple iPhone smartphones and a number of Android devices (e.g., Motorola Droid, Samsung Galaxy S, ASUS Transformer, HTC Desire). In our experiments, we set the function \( g(a) = a^2 \) in terms of the physical meanings of both standard deviation of the normal distribution and accuracy measurements of the GPS generated data. We used a threshold of \( T H_a = 40 \) meters and the typical segment length was \( |S_k| = 20 \) samples.

We computed the average distance between every processed sample and its corresponding ground truth position for each GPS sequence data file, and compared these values to the average distance between every measurement sample and the ground truth position. On average, the Kalman filtering based algorithm and the weighted linear least squares regression based algorithm improve the GPS data correctness by 16.3% and 21.76%, respectively. Figure 2 illustrates the cumulative distribution function (CDF) for both algorithms. Our methods increase the proportion of GPS data with low average error distance and shorten the largest sequence average error distance by around 30 meters. Moreover, we apply our algorithms to 17 highly inaccurate datasets (i.e., the best accuracy value \( \max(a) > 50 \) meters). We found that our algorithms reduce the average distance between the measured positions and the ground truth data to
Figure 2: The cumulative distribution functions of average error distances. The \( y \)-value of each point represents the fraction of the total number of GPS sequence data files whose average distance to the ground truth positions is less than the given distance value.

(a) Kalman filtering based algorithm.
(b) Weighted linear least squares regression based algorithm.

Figure 3: The average error distance results between the corrected data and the ground truth positions of highly inaccurate GPS sequence data files.

(a) Kalman filtering based algorithm.
(b) Weighted linear least squares regression based algorithm.

a great extent. The Kalman filtering based algorithm and the weighted linear least squares regression based algorithm reduce the average error distances by 39.82\% and 48.18\%, respectively. Figure 3 illustrates the average error distance reduction of every GPS data sequence file.

5. CONCLUSIONS

We analyzed several typical error patterns for real-world positioning data and proposed two approaches to improve the location measurement accuracy while relying purely on the GPS generated data. The experimental results demonstrate that our methods are highly effective in enhancing accuracy. In our future work we plan to include other sources of geographical knowledge into the sensor data correction model, e.g., GIS databases. With additional, reliable supplemental information to aid in the fusion, the correction model may further improve its performance. We believe that such processed, highly accurate sensor data are useful for other sensor-aided social media applications.

Acknowledgment

This research has been supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

6. REFERENCES