GPS Self-adaption Data Correction Based on the Data Change Rate

Dong-fei ZHU¹, Neng-kai WU²*, Xiao-mei Li¹ and Dong-yao JIA²

¹Wuhan Metro Group Co., Ltd, China
²School of Electronic Information Engineering, Beijing Jiaotong University, China
*Corresponding author

Keywords: Data change rate, GPS correction, Data pretreatment, Eliminate the error points, Interpolate.

Abstract. As GPS data is affected by various factors, subsequent pretreatment is extremely important. The current GPS data preprocessing algorithms have prevalence of inefficient, pertinence and other defects, causing that the local details of the GPS positioning accuracy is not enough. In order to improve the efficiency and instantaneity of the data pretreatment, a real time self-adaption dynamic correction method based on the data change rate is proposed in this paper. It determines the upper and lower correction limits by the change rate of rolling sample data, and then find out the error points according to this correction range to eliminate or interpolate them. The field experiment results show that after the dynamic correction, the GPS data accuracy increases by 57.1%, so the accuracy of map matching greatly improves.

Introduction

GPS data preprocessing refers to processing collected GPS information before map matching, which mainly achieves the following goals: format standardization, abnormal data clearing, error correction, etc, to ensure the precision and the integrity of data to improve positioning accuracy. Many GPS data preprocessing methods have been proposed at present mainly including single epoch calculating method, dynamic kalman filtering method, spectrum analysis method and so on. In paper [1], the initial assumption of single epoch fast ambiguity calculating was put forward and then paper[2] developed it. However the method owns high success rate just for the baseline of less than 1 km. According to paper[3] the dynamic kalman filtering method could be used in multivariable system and non-stationary random process, while it would be make filtering result inaccurate in practice because the system noise and observation noise could not be white noise strictly. Paper [4] proposed the spectrum analysis algorithm for GPS data available in both time domain and frequency domain, but its demand of constant duration made data processing difficulty and damaged its real-time ability. GPS data preprocessing requires us to deal with new collected GPS points real-time, delete some GPS points with great error and then make data interpolation in some of the deleted points to improve the accuracy of the positioning algorithm finally.

In order to improve the efficiency and instantaneity of the data pretreatment, a real time self-adaption dynamic correction method based on the data change rate is proposed in this paper, which determines the upper and lower correction limits by the change rate of rolling sample data, and then find out the error points according to this correction range to eliminate or interpolate them. The experiment results show that the accuracy of map matching improves obviously after the dynamic correction.

Adaptive Correction based on Data Variation Rate

In data variation rate analysis, we commonly use the statistical characteristics of residual sequence and statistical decision methods to judge the appearance of error value. Fixed threshold is applied in most algorithms, which can be calculated by a large amount of prior knowledge and sampled points. Evidently the method is too complex to deal with error real-time and effectively. Therefore an
adaptive correction algorithm is proposed based on data variation rate in the paper, which applies the variation rate of collected data for error data judgment and includes right data's variation rate into a corrected interval changing with GPS data. The data will be regarded as error if its variation rate beyonds the corrected interval.

Historical data must be used in order to calculate the variation rate of positioning data and the matching efficiency limits its number. In the paper every \( n (n \leq 15) \) data will be divided into a group. When calculating the variation rate of sampled data, we must choose suitable historical data as original sampled data and here an adaptive sliding sampling method with high real-time ability is used as follows:

\[
\text{Group 1: } D(1), D(2), \ldots, D(n-1), D(n) \\
\text{Group 2: } D(2), D(3), \ldots, D(n), D(n+1) \\
\text{...} \\
\text{Group m: } D(m), D(m+1), \ldots, D(m+n-1), D(m+n)
\]

(1)

Calculate the data variation rate of one group in the moment \( i \):

\[
R(i) = \frac{D(i+1) - D(i)}{2}, 1 \leq i \leq n - 1
\]

(2)

Where \( R(i) \) and \( D(i) \) refers to the variation rate of GPS data and collected GPS data value in the moment \( i \). Calculating the weighed average rate of the group \( \text{weiR}(n) \) as follows:

\[
\text{weiR}(n) = w_1R_1 + w_2R_2 + \cdots + w_iR_i + \cdots + w_nR_n, 1 \leq i \leq n
\]

(3)

Where \( w_i \) refers to the weight of \( R_i \) and \( w \) is subjected to the following formula:

\[
\begin{cases}
w_1 + w_2 + \cdots + w_i + \cdots + w_n = 1 \\
w_1 < w_2 < \cdots < w_i < \cdots < w_n
\end{cases}
\]

(4)

Because data nearer the matching point will bring about greater influence to its change trend, so the weighted coefficient \( w \) must meet the formula \( w_1 < w_2 < \cdots < w_i < \cdots < w_n \).

Different correction range can be obtained according to different positioning accuracy and here we assume that the data rate floating range is 20%, namely:

\[
0.8 \text{weiR}(k) \leq \text{AccR}(k) \leq 1.2 \text{weiR}(k)
\]

(5)

\( \text{AccR}(k) \) is the allowed change range within which the point is normal.

\( \text{ThrU} (k) \) and \( \text{ThrD} (k) (k > n) \), the maximum and the minimum limit of the correction interval, is defined as follows:

\[
\begin{cases}
\text{ThrU} (k) = a \times \text{weiR}(n), a > 1 \\
\text{ThrD} (k) = b \times \text{weiR}(n), b < 1
\end{cases}
\]

(6)

In formula (6), \( a \) and \( b \) are the correction coefficient. They may be taken according to the actual accuracy requirement, for example \( a = 1.5, b = 0.7 \). Current data rate value will be considered as error value if it is beyond the correction interval set by the upper limit \( \text{ThrU} (k) \) and the lower limit \( \text{ThrD} (k) \). And then it will be done with in the following four conditions:
\[
\begin{align*}
R(k) &< \text{ThrD}(k) \\
R(k) &> \text{ThrU}(k) \\
R(k) &\in \text{AccR}(k) \\
\text{ThrD}(k) &< R(k) < \text{ThrU}(k) \quad \text{if} \quad R(k) \notin \text{AccR}(k)
\end{align*}
\]

(7)

\(R(k)\) refers to current data rate and the concrete method will be introduced then.

Resulting from the adaptive sliding sampling method, \(R(k), \text{ThrU}(k), \text{ThrD}(k)\) and \(\text{AccR}(k)\) need to be calculated real-time to ensure the reliability of the following judgment. At the same time, the fourth data cannot be replaced by estimated value while it is matched with the current sampling value to avoid that more continuous estimated values will lead to incorrect matching.

**Eliminating Error Point and Interpolation Processing**

Correction processing refers to dealing with the sampling point with great error which will improve the matching accuracy and efficiency of following points. We need to convert GPS data into coordinate point in the cartesian coordinate system.

In order to realize real-time positioning, it can be divided into the following two types to process error points:

- **Repeat points**
  
  As is shown in figure 1, solid lines are a group of adjacent parallel tracks in the orbit station and hollow dots represent the collected GPS information to be matched. A lot of repeat information exists at the end of an orbit resulting from repeat recording, repeat measurements and staying still. The data rate of the points with repeat information will be certainly less than the lower correction limit \((R(k) < \text{ThrD}(k))\) after the adaptive correction and then the point will be eliminated and replaced with its previous data \(D(k-1)\). In the place with repeat points, the positioning point is the GPS data arriving firstly.

![Before treatment](image1)

Before treatment

![After treatment](image2)

After treatment

**Figure 1. Eliminating repeat points.**

1. **Deflected points**

   As is shown in figure 2, solid lines are a group of adjacent parallel tracks in the orbit station and hollow dots represent the collected GPS information to be matched. It can be seen that points a, b and c deviate far away from the track probably resulting from error recording or transmission problem. Calculate the data rates of these points and compare them with the upper and the lower limit of the correction interval:
   
   - i. If \(R(k) < \text{ThrD}(k)\), the point is a repeat point mentioned above.
   - ii. If \(\text{ThrD}(k) < R(k) < \text{ThrU}(k)\) and \(R(k) \notin \text{AccR}(k)\), the point will be taken as an error point. Because its weighed average rate is within the correction interval, the data of the point is not deviate from the precise value completely but is accurate partly, and then we may use the previous two sampling data to replace the point, such as \(2 \times D(k-1) - D(k-2)\).
   - iii. If \(R(k) > \text{ThrU}(k)\), the point deviates from the track of other points obviously and then we can calculate the angle of the two attachments between the point and its previous point and its next point. The angle value decides how to process the error points.
Calculate the angle $\theta_k$ according to the formula 8 and the result will be divided into different types as what the formula 9 says.

$$\theta_k = \arctan \left( \frac{\Delta y}{\Delta x} \right) \times \frac{180}{\pi} = \arctan \left( \frac{y_k - y_{k-1}}{x_k - x_{k-1}} \right) \times \frac{180}{\pi}$$ (8)

$$\theta_k = \begin{cases} 
\theta_k & \Delta y \geq 0, \Delta x > 0 \\
\theta_k + 180 & \Delta y \geq 0, \Delta x < 0 \\
\theta_k - 180 & \Delta y < 0, \Delta x < 0 \\
90 & \Delta y < 0, \Delta x = 0 \\
-90 & \Delta y < 0, \Delta x = 0 
\end{cases}$$ (9)

Assume that $\Delta \theta_k (\Delta \theta_k = \theta_k - \theta_{k-1})$ denotes the angle variety value and it needs to be corrected as follows to avoid that it is beyond positive or negative 180 degrees:

$$\Delta \theta_k = \begin{cases} 
\Delta \theta_k - 360 & \Delta \theta_k > 180 \\
\Delta \theta_k + 360 & \Delta \theta_k < -180 
\end{cases}$$ (10)

The station’s orbit is relatively flat in a small area, so the changes in the angle between the adjacent sampling points should be small, which make us able to set a fixed threshold $\sigma$ to process it based on the actual situation.

If $\Delta \theta_k > \sigma$, it means that the point deviate from the track seriously and must be replaced by by Interpolation processing just like point b and c in figure 2. Or the point will be replaced by the point it projects in the line between its previous point and its next point just like point a in figure 2.

iv. Recount $R(k)$, $weiR(k)$, $ThrU(k)$, $ThrD(k)$ and $AccR(k)$ based on the above formulas and then process the next data using the results to ensure the reliability of the data judgment.

Interpolation processing is necessary after eliminating the deviation points mentioned in the previous chapter and when GPS data is not enough due to low GPS receiver frequency or leakage written data by recorders.

It is divided into two types for interpolation processing in the paper, namely interpolation processing in a straight track or a curve orbit. According to paper[5] linear interpolation is available in the straight track while Parabolic interpolation is more useful in the curve orbit.

It’s certainly that all data is standardized basically and the map matching accuracy will be improved definitely after error points eliminating and interpolation processing. It should be pointed out that we must calculate the percentage of the error data in all the data real-time. If the ratio is great we should find out the faults and solve it in time to improve positioning accuracy finally.

**Experiment**

We conduct four measurements on a chosen orbit which is divided into sixteen sections (twenty meters long each) using GPS sensors in the simulation experiment. In every simulation graphs, X axis says latitude value and Y axis denotes longitude value. It shows four simulation curves with different
colors, where we can see three error points (two points in the red tracks and one point in the green tracks). The three points are definitely not the logical data collected by the maintainers in normal circumstances and should be eliminated. As is shown in figure 4, the curves become much more smooth after GPS data pretreatment, including error points eliminating and interpolation processing, which contributes to higher map matching accuracy.

In the experiment, we assume that the average of the four measurements is the precise value of the GPS data. Compare the GPS data both before pretreatment and after pretreatment with the precise values and calculate the absolute error severally, including the latitude error and the longitude error. Figure 5 and figure 6 show the absolute error curves of each measurement and the yellow lines of the figures denote X axis (y=0).

As is shown in the two figures, the maximum error value is 0.014 before pretreatment while it changes into 0.006 after pretreatment. Therefore the GPS data pretreatment method is useful to increase the matching accuracy and the result can be obtained as follows:

$$E' = \frac{20}{0.02} \times 0.014 = 14 \text{ m}$$ (11)

$$E = \frac{20}{0.02} \times 0.006 = 6 \text{ m}$$ (12)

$$\rho = \frac{14-6}{14} \times 100\% = 57.1\%$$ (13)

Where $E'$ denotes the matching accuracy before pretreatment, $E$ denotes the matching accuracy after pretreatment and $\rho$ denotes the rate it increases. According to the results it is obvious that the matching accuracy achieves 6 meters and increases 57.1 percent after the GPS data pretreatment.
Field experiments show that the GPS data error turns down and the measurements become more precise after the GPS data is corrected dynamically, which makes the curve more smooth and increases the matching accuracy distinctly.

**Conclusion**

Based on the analysis of the characteristic of the GPS data, an adaptive correction method of the GPS data change rate is put forward to increase the efficiency of the data pretreatment. It determines the upper and lower correction limits by the change rate of rolling sample data, and then find the error points according to this correction range, and eliminate or interpolate them. The field experiment results show that after the dynamic correction, the GPS data accuracy improved by 57.1%, so the accuracy of map matching is greatly improved.

**References**

[1] Sharma S, Banerjee P. Effect of errors in position coordinates of the receiving antenna on single satellite GPS timing[J]. Indian Journal of Pure & Applied Physics. 2010, 48(6):429-434.

[2] Santamaria-Gomez A, Bouin MN, Collilieux X, et al. Correlated errors in GPS position time series: Implications for velocity estimates[J]. Journal of Geophysical Research-Solid earth.2011, 116.

[3] Jia Chen, Jihua Hu, Feizhou Zhang. Research on map matching algorithm for vehicle monitoring navigation[J]. Journal of Peking University. 2009, 45(2):299-305.

[4] Nontasud S, Leelaruji N. Applied the artificial neural network for reduce the position error on GPS receiver due to ionospheric irregularities[C].2008 International Symposium on Communications and Information Technologies. 2008:649-653.

[5] Jianchao Hu, Zhong Wang, Wei Zhang. The improved map matching algorithm research[J]. Communication technology. 2009, 11(42):166-167.

[6] Jiang Zeng, Zibin Jiang. The data error analysis and accurate positioning of the GPS navigation system[J]. Modern surveying and mapping, 2004, 27(1):25-27.

[7] Jie Deng. The realization of the vehicle navigation and GPS positioning system[J].Journal of Gansu Sciences. 2003, 15(3):123-125.

[8] Lecce V. Di, Amato A., Piuri V. Neural technologies for increasing the GPS position accuracy[C]. Computational Intelligence for Measurement Systems and Applications. 2008:4-8.

[9] Ochieng W.Y., Quddus M., Noland R.B. Map-Matching in Complex Urban Road Networks[J]. Revista Brasileira de Cartografia, 2009-lsie.unb.br.

[10] Guanli H, Bin Y, Keyu J. New algorithms on the solution to drifting problem of GPS positioning[J]. Journal of Computers, 2011, 6(1):83-89.