International Core Journal of Engineering Volume 7 Issue 3, 2021
ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

Sensor Detection Target Trajectory Fitting Based on Local Principal Curves

Weijie Zhang

School of Shanghai Maritime University, Shanghai 201306, China.

zhangwj_524@163.com

Abstract

The target information detected by sensors such as radar will change according to different sensor positions, which will cause some errors, but there may be a certain relationship between the points detected by sensors in different positions. The master curve is a feature extraction method based on nonlinear transformation. It is a smooth curve that passes through the middle of the data distribution and satisfies self-consistent, which can better reflect the structural characteristics of the data distribution. Curve fitting refers to selecting an appropriate curve type to fit the observed data according to the characteristics of the given data and reflect the changing trend of the observed data. This article attempts to use the existing data to obtain the master curve, then use the characteristics of the sensor to add noise to the master curve, curve fitting the points after adding noise, and finally analyze the error of the fitted curve and the master curve. The experimental results show that the master curve can well reflect the structure of the data, and the detection errors between various sensors have a certain correlation.

Keywords

Principal Curve; Curve Fitting; Sensor Detection.

1. Introduction

All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper. Use *italic* for emphasizing a word or phrase. Do not use boldface typing or capital letters except for section headings (cf. remarks on section headings, below).

Do not number your paper: All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question.

In the military, sensor simulation is an important part of the simulation of the battlefield situational awareness system. Various sensors such as radar, photoelectric, and infrared are used to detect various targets. There are many factors that may affect sensor detection errors. Among them, radar can be used to detect the data information of some aerial targets, such as the flight trajectory information of a certain fighter aircraft in a certain period of time. The flight trajectory information is obtained from the information of the detected discrete points, which are detected by sensors in different positions. The points are not the same, but there may be some associations between the two points. Through error analysis, we can see whether there is a connection between the two. Therefore, this paper

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

combines the master curve theory with the curve fitting method, and proposes a method to study the trajectory fitting of the sensor detection target and the detection error between sensors.

As a non-linear extension of linear principal component analysis (PCA), the concept of master curve was first given by T. Hastie and W. Stuetzle in 1989 after 5 years of research. After that, based on its wide application prospects, scientists all over the world made a series of improvements and expansions on the theory and algorithm of the master curve. In 1992, the BR master curve given by Banfield and Raftery, although it made some improvements to the HS master curve, the error was relatively large. Subsequently, the T master curve and the D master curve were generated. In 1999, B. Kegl et al. proposed the polygonal algorithm, which made the master curve theory step into a new milestone. In recent years, many documents have also appeared from the theory and algorithm of master curve to the application of master curve. As a geometrically intuitive and algorithmically feasible method of nonlinear data reduction, the master curve has attracted the attention of the computer science community.

Curve and surface fitting technology plays an important role in many fields, and is widely used in computer-aided design, computer graphics, data visualization, virtual reality, digital geometry processing, data mining and other fields. In applications such as geometric modeling, specific expressions of curves and surfaces can construct specific models of curves and surfaces. However, in industrial applications, there are usually specific curve and surface models first, and when these models are stored or displayed digitally, an analytical formula is constructed to express such curve and surface models. In general, the relatively simple B-spline polynomial is chosen as the analytical formula, because B-spline has the advantages of geometric invariance, segmental smoothness and local support [6]. When using B-spline for curve and surface fitting, it is necessary to determine the node vector, control vertices and parameters. The usual method is to first determine the node vector and parameters, and then solve the control vertices. Once the node vector and parameters are determined, the problem becomes a linear problem of solving the control vertex, that is, solving a linear equation system [7]. Usually, the number of fitting data points is more than the control vertices, so only the data points can be approximated, so the least square method is adopted. Least square fitting is a classic method for solving B-spline fitting, which can directly solve a linear system [8]. However, when the scale of the fitted data points is large, the dimension of the coefficient matrix is high, and a large sparse linear equation system needs to be solved. The method of directly inverting the matrix consumes a lot of computing resources and is low in efficiency. The method is to use iterative methods to solve linear equations, such as Jacobian iteration and Gauss-Seidel iteration. In addition, there are B-spline curves and surfaces based on the asymptotic approximation iteration method proposed by Lin HW et al. [9-10] Fitting algorithm (Progressive and Iterative Approximation, PIA) and Deng CY et al. [11] proposed Least Square Progressive and Iterative Approximation (LSPIA).

This article mainly analyzes the relationship between the master curve and the refitted curve. First, make a small improvement to the existing master curve algorithm to obtain the master curve of the original data and use it as the true trajectory detected by the sensor. Then according to the characteristics of the sensor, the main curve is processed with noise, the data after the noise is processed and fitted to obtain a new curve, and the relationship and error between this curve and the main curve of the original data are studied.

2. Existing Principal Curve Algorithm

2.1 HS Principal Curve

The master curve is a nonlinear extension of the first principal component, which is the best description of the one-dimensional linearity of the data set. The master curve emphasizes finding a smooth one-dimensional curve that passes through the middle of the data distribution and satisfies self-consistent. The theoretical basis is to find a non-Euclidean low-dimensional manifold embedded

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

in a high-dimensional space. The master curve is to find a geometrically intuitive, theoretically complete, and algorithmically feasible method to describe the internal structure of the data set.

In 1989, T. Hastie and W. Stuetzle published the article "Master Curve" [1], which was a pioneering work on the theory of master curves, after which they called the master curve HS master curve.

The HS main curve is defined as follows:

Definition 1: Under the random distribution density function h, if $E(x|t_f(x)=t)=f(t)$, then the curve f is said to be self-consistent. The schematic diagram of self-consistent is shown in the figure below.

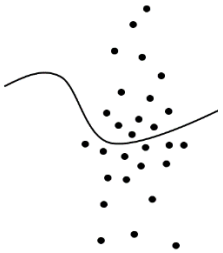

Fig. 1 Self-Consistent Curve

Definition 2: For a given data distribution or data set, a self-consistent smooth curve is called the master curve.

With the definition and self-consistent characteristics of the master curve given by Hastie, we know the hypothesis of the smoothness of the HS master curve: that is, if two points x and y are adjacent, then the projections of these two points on the curve are also adjacent, Therefore, it is ensured that the topological structure of the data distribution remains unchanged on the master curve. According to the static characteristics of the distance function, the master curve can be approximated by any smooth curve. In this way, the master curve distance function can be expressed as:

$$D^2(f)=E\,\|x-f(t_f(x))\|^2=E_{t_f(x)}E[\|x-f(t_f(x))\|^2|\,t_f(x)] \quad (1)$$

The above equation is equivalent to finding f and t so that $D^2(f,t)=E\,\|x-f(t)\|^2$ is the smallest.

When f is fixed, t is selected according to the nearest neighbor principle to reduce the above formula (projection step): in the same way, t is fixed, and f is obtained according to the self-consistent characteristic to further reduce it (desired step).

The master curve algorithm given by Hastie under the condition of known data distribution is as follows:

Step1 Initialization: Let $f^{(0)}(t)=\bar{x}+ut$, where u is the first linear principal component of distribution h. Assume $t^{(0)}(x)=t_f(0)(x)$, j=0.

Step2 Projection step: For $f^{(j)}(t)$, calculate the projection index $t_f{}^{(j)}(x)$ of all data points $x\in R^d$:

$$t_f{}^{(j)}(x)=t^{(j)}(x)=\max\{t:\|x-f^{(j)}(t)\|=\min_{\tau}\|x-f^{(j)}(\tau)\|\} \quad (2)$$

Step3 Expection step: Through self-consistent characteristics, obtain the smooth curve f, define and estimate: $f^{(j+1)}(t)=E(x\,|\,t_f{}^{(j)}(x)=t)$.

Step4 If $(1-\Delta(f^{(j+1)})/\Delta(f^{(j)}))$ is less than a certain threshold, stop; otherwise, let j=j+1, go to Step2.

2.2 Local Principal Curves

Do not number your paper: All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

Assume a d-dimensional data cloud $X_i \in R^d, i = 1,...,n$, where $X_i = (X_{i1},..., X_{id})$. We try to find a smooth curve which passes through the "middle" of the data cloud. The curve will be calculated by means of a series of local centers of mass of the data, according to the following strategy:

1. Choose a suitable starting point x(0). Set $x = x(0)$.

2. Calculate the local center of mass μ_x around x.

3. Perform a principal component analysis locally at x.

4. Find the new value x by following the first local principal component γ_x starting at μ_x.

5. Repeat steps 2 to 4 until μ_x remains (approximately) constant. The series of the μ_x determines the local principal curve. In the sequel we will explain these steps in detail:

In principle,everypoint $x(0) \in R^d$ which is in or close to the data cloud can be chosen as starting point.There are two ideas which suggest themselves:

• Based on a density estimate the point with the highest density $x(0) = \max$, $x \in R^{f(x)}$ is chosen.

• A point $x(0) = X_i$ is chosen at random from the set of observations.

The advantage of the density method is that one can be quite sure not to start in a blind alley, whereas a randomly chosen point could be an outlier far from the data cloud which stops the algorithm already in the first loop. However, in this case it is easy to draw another starting point, and the computational costs of these cond approach are much lower. Moreover, for the handling of crossings a randomly chosen starting point is even superior to a high density point.

Let H be a bandwidth matrix and $K_H(.)$ a d-dimensional kernel function. Given that all components of X are measured on the same scale, we set $H = \{h^2 I : h > 0\}$, with I denoting the d-dimensional identity matrix. For a detailed description of multivariate kernels and bandwidth matrices see Wand and Jones (1993). The local center of mass around x is given by

$$\mu(x) = \frac{\sum_{i=1}^{n} K_H(X_i - x) X_i}{\sum_{i=1}^{n} K_H(X_i - x)} \tag{3}$$

Comaniciu and Meer (2002) studied the properties of the mean shift,which is given by $\mu(x) - x$, and investigated the relation of $\mu(x)$ to the Nadaraya-Watson estimator. For ease of notation, we will use the abbreviation $\mu_x = \mu(x)$, and denote the j-th component of $\mu(x)$ by μ_j^x.

2.3 Analyze HS and Local Principal Curves

The HS master curve is a smooth curve that satisfies the self-consistent characteristics through the "center line" of the data distribution. It describes the intermediate, self-consistent and non-parametric characteristics of the data distribution, but its original definition and some future applications still have shortcomings. For example, the existence of the HS master curve has not been proven yet. It only proves the existence under some special distributions (such as elliptic distribution, etc.), and there are also problems such as model deviation, estimation deviation and convergence.

The local master curve proves that the concept of combining the local principal component and the mean shift is a simple and useful tool for calculating the master curve. Compared with most other master curve algorithms, this tool shows superiority in the simulation data set. performance. The algorithm can work on simulated real data sets even for highly complex data structures. This includes data situations that cannot be handled satisfactorily by existing methods, such as data with multiple or broken branches. Especially for noisy spatial data (such as ubiquitin data), this method has a high potential to detect its latent structure. The theoretical background of this method still needs further research. Although it works well, we don't have much theoretical basis why we should use a data cloud with local principal components. This choice is wise, but it is by no means the only one, and

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

there seem to be many alternatives, such as extrapolation in which part of the curve has been estimated. Due to the good characteristics of mean shift, even if the observation point is not orthogonal to the main curve, it can even use a line in any direction. Crucially, improvementism—the main curve of the "middle" direction of the data cloud will be adjusted after the transformation of means. However, by applying local principal components, the algorithm is the fastest and most stable, and its results can be intuitively expected. We consider approximating the first local principal component to the tangent (biased) approximation of the peak line of the estimated density. μ_x uses the weight w_i in (4) to minimize the weighted distance between X_i and the line. Therefore, the first local principal component is the line that locally gives the best fit. In addition, it will be interesting to study whether the proposed algorithm can be extended to obtain the local main surface or even higher-dimensional local mainstream shapes. This is easy to do because there is no simple technique to transfer the signal translation coordinate system to a higher dimensional curve without developing new concepts.

3. Existing Fitting Algorithm

3.1 Least Squares Method

Interpolation and fitting are both the process of finding approximate curves with similar changes and characteristics based on several known data points of an unknown function (or a function that is known but difficult to solve). But what the interpolation method requires is that the approximate curve needs to pass through the data points completely, and the fitting is to get the closest result, emphasizing the concept of minimum variance.

Curve fitting refers to selecting an appropriate curve type to fit the observed data according to the characteristics of the given data and reflect the changing trend of the observed data. For a given set of positioning data $\{(x_i,y_i),i=1,2,...,N\}$, the set number can be described by a curve model, expressed as:

$$y(x)=f(a,x)=a_0+a_1x+a_2x^2+...+a_nx^n(n<N-1) \tag{4}$$

Among them, n is the order of the curve model. In order to measure the advantages of fitting, the article puts forward many standards for goodness of fit. The most commonly used method is to select parameter a to minimize the weighted sum of squares of residuals $\delta_i=f(x_i)-y_i$ between the fitted model and actual observations at each point, namely:

$$|\delta|^2=\sum_i^N[y_i-(a_0+a_1x+a_2x^2+...+a_nx_i^n)]^2 \tag{5}$$

Using the criterion of least squares to find the extreme value according to the multivariate function, the function has a partial derivative at a certain point, and the partial derivative is zero, we can get:

$$\frac{\partial\delta^2}{\partial a_j}=0 \tag{6}$$

The unique solution of the coefficient matrix $[a_0,a_1,...,a_n]^T$ can be obtained from the difference above, and the coefficient matrix can be substituted into equation 2 to obtain the optimal solution based on least squares curve fitting.

3.2 Polynomial Fitting

In practical problems, it is often necessary to predict the expression of the function $y=f(x)$ from a set of observation data $(x_i,y_i)\ i=1,2,...,n$. From a geometric point of view, this problem is that a given set of data points (x_i,y_i) to describe the approximate image of the curve $y=f(x)$. Interpolation method is a numerical method to deal with such problems. However, because the interpolation curve requires a pass through each given data point, this limitation will retain the error

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

of the given data. If the error of the individual data is large, the interpolation effect is obviously not ideal.

The problem we are facing now has such a characteristic: the data given is not necessarily reliable, the error of individual data may even be large, but the data given is a lot. The subject of the curve fitting method is to find the law from a large amount of seemingly disorganized data, that is to say, try to construct a curve, the so-called fitting curve, reflecting the total of the given data points Trend to eliminate local fluctuations in the given data.

Assume that the distribution of the given data points (x_i, y_i) $i = 1, 2, ..., n$ is roughly in a straight line. Although the fitted straight line cannot be required to pass through all the data points (x_i, y_i) strictly, it is always hoped that it passes through the attachment of the given data point as much as possible, that is to say, the requirement is approximately established:

$$y_i \approx a + bx_i, i = 1, 2, ..., N \tag{7}$$

Here, the number of data points is usually much larger than the number of undetermined coefficients. Therefore, the construction of a fitted line is essentially an algebraic problem for solving overdetermined equations. Assume:

$$\bar{y}_i \approx a + bx_i, i = 1, 2, ..., N \tag{8}$$

Represents the approximate value obtained according to the fitted line $y = a + bx$. Generally speaking, it is different from the measured value yi, and the difference between the two is called the residual. Obviously, the size of the residual is an important indicator of the quality of the fit. Specifically, one of the following three criteria can be used to construct a fitting curve:

1. Make the maximum absolute value of the residual minimum, namely

$$\max |e_i| = \min \tag{9}$$

2. Make the sum of the absolute values of the residuals the smallest, namely

$$\sum |e_i| = \min \tag{10}$$

3. Minimize the sum of squares of the residuals, namely

$$\sum e_i^2 = \min \tag{11}$$

Analyzing the above three criteria, the first two formulations are relatively natural, but they are not convenient for practical application due to the absolute value calculation; the method of selecting the fitting curve based on the third criterion is called the least squares method of curve fitting.

4. Steps and Results of Principle Curve Generation

4.1 Steps to Generate Principle Curve

During the experiment, MATLAB R2019a was used to program the local master curve algorithm, and run this program in win10 (64bit) system, 2.7GHz single-core processor/16G memory environment to process the trajectory data and generate its master curve . Use the Qt Creator compiler, use C++ to design the interface, mark points on the map, and display various points, such as the main curve, the points after adding noise, and the curve after fitting, for the most direct display effect.

First, set the method of generating random data. The data generated here is similar to a sine curve. After the data is generated, the adjacency matrix of the data is calculated to find the neighborhood of each data point. After obtaining the adjacency matrix, iteratively search for the local first principal component, calculate the mean sample point of each neighborhood, and finally connect the neighborhoods through cubic splines. To find the mean sample point of each neighborhood, first set the local sample weight initially, and traverse each sample. When the position difference between each old sample point and the new sample point is greater than 0.001, iteratively update the weight of each sample in the neighborhood, and then obtain the coordinates of the new neighborhood sample

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

point according to the updated weight, and then calculate the neighborhood set covariance matrix to find The first principal component line is to find the reconstruction of all points in the neighborhood. Finally, record the mean value of each neighborhood (stored in a new matrix, which is the data of the center point, and the new point after each data point is updated).

4.2 The Result of Generating the Principle Curve

In the result graph below, the blue points represent the initial data, and the red points represent the main curve.

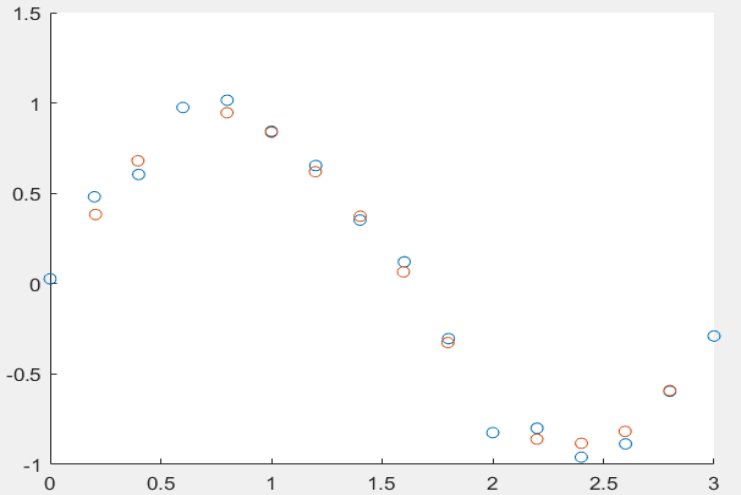

Fig. 2 Local principal curves

5. Noise Processing and Fitting Results

5.1 The Principle and Result of Adding Noise

5.1.1 Principle of Adding Noise

The sensor will be affected by many errors during detection, such as atmospheric ducting, suppression interference, and errors caused by the radar itself.

The radar error mainly comes from the range azimuth and elevation angle.The deviation of the distance is like the mean square error σ_r of the distance measurement, and the approximate value is $\tau/3$. The average delay error due to atmospheric refraction is:

$$\Delta R_{rf} = \frac{0.0072N_s(h+0.33h^2)}{15+h+0.33h^2} \tag{12}$$

The typical value of N_s is 313, the distance R is between 200~300 km, and the target height h is 10~20km. Due to the distance, there is also an internal noise error, which is:

$$\sigma_{rn} = \frac{\tau}{\sqrt{2E/N_0}} \tag{13}$$

Where τ is the radar pulse width and $2E/N_0$ is the total received signal-to-noise energy ratio.

The bearing deviation is:

$$\sigma_{\theta_n} = \frac{0.5\theta_{a3}}{\sqrt{n(S/N)_m}} \tag{14}$$

Where θ_{a3} is the 3dB width of the antenna horizontal beam; n is the number of pulses within the 3dB width of the beam when the antenna sweeps over the target; $(S/N)_m$ is the output signal-to-noise ratio when aiming at the target (the general value is about 0.2°).

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

Elevation angle deviation: θ_e is the elevation angle to the target, θ_e is usually less than 20°; the total mean square value of the elevation angle error $\sigma_{\theta_e} \le 0.078^o$;Mean square value of height measurement error σ_H^2 , simplified as:

$$\sigma_H = R_t \cos\theta_e \sigma_{\theta_e} \tag{15}$$

Suppress interference: According to relevant literature, it can be known that the detection range of radar under interference conditions is:

$$R_{\max J} = \sqrt[4]{\frac{P_t G_r^2 \lambda^2 \sigma_\tau R_J^2 L_1 B_J}{4\pi (S/N)_{min} L P_J G_J G_1 B_r}} \tag{16}$$

P_t is the transmit power, generally $7*10^6 W$; G_t is the radar transmit and receive antenna gain, $G_r = G_t = 10^4$; λ is the radar signal wavelength, with a value of 0.056 cm; σ_τ is the effective reflection area of the target, with a value of $5\,m^2$; R_J is the distance between the radar and the jammer is $106\,km$; L_1 is the transmission loss of the jamming signal, the value is 4; B_J is the jammer jamming signal bandwidth, the value is $106\,Hz$; $(\frac{S}{N})_{min}$ is the minimum detectable signal noise ratio, the value is 11; L is the power loss factor of the radar ($L > 1$), the value is 2.5; P_J is the transmission power of the jammer, the value is 2000W; G_J is the gain of the jammer antenna to the radar direction, the value is 7; G_1 is the gain of the radar antenna in the direction of the jammer, with a value of 60; B_r is the bandwidth of the radar receiver, with a value of $6*105\,Hz$; R is the distance between the radar and the target.

The atmospheric duct mainly affects the detection range of the radar, and the minimum detectable signal power under the atmospheric duct is:

$$s_{i\min} = kT_0 B_n F_0 D_0 = 10\lg(kT_0) + 10\lg B_n + F_0 + 10\lg(\frac{S}{N})_{min} \tag{17}$$

Where B_n is the radar receiver bandwidth (MHz), the value is $6*105\,Hz$; F_0 is the receiver noise figure (dB), the value is 3 dB ; T_0 is the equivalent noise temperature of the receiver (k), T_0 =290K at standard room temperature, assuming 600K; $(\frac{S}{N})_{min}$ is the minimum detectable signal-to-noise ratio, the value is 10.

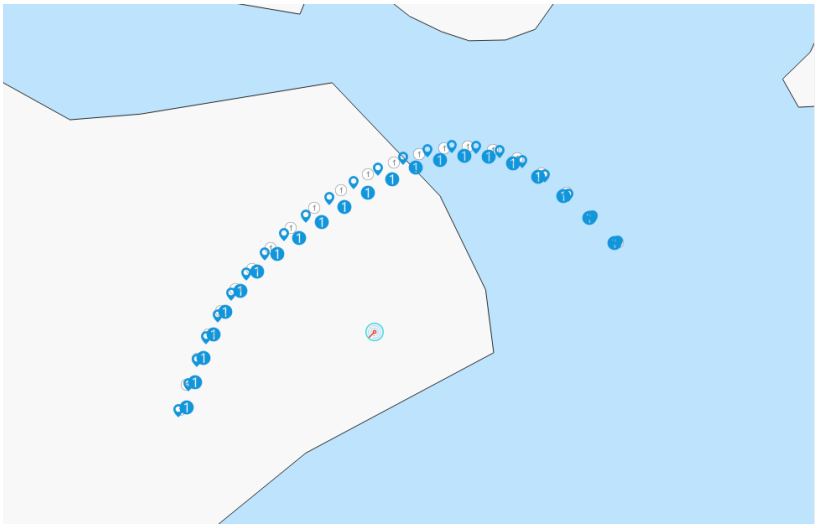

Fig. 3 The result after adding noise

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

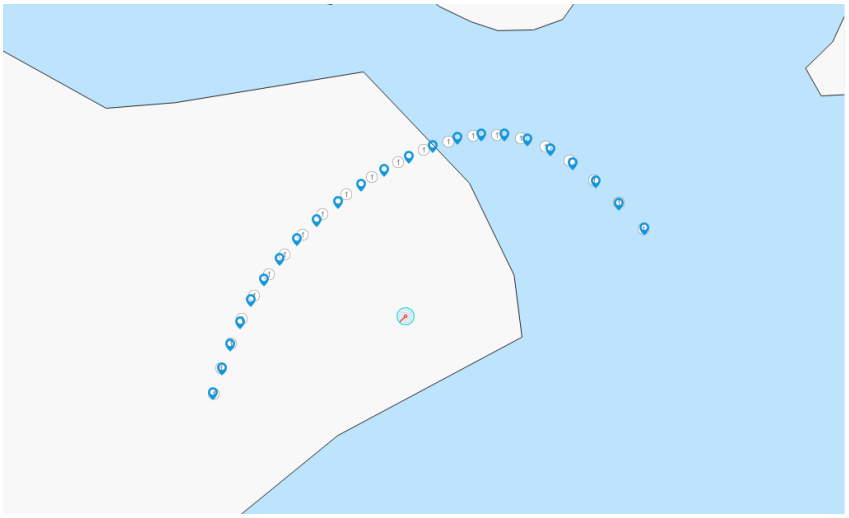

Fig. 4 The result after fitting

5.1.2 Noise and Fitting Results

After the noise is added, the noise points are fitted. The solid point No. 1 represents the result after noise, the hollow points represent the fitted curve, and the white points represent the main curve.

6. Error Analysis

6.1 Methods and Principles of Error Analysis

6.1.1 KL Dispersion Analysis

In probability theory or information theory, KL divergence (Kullback–Leibler divergence), also known as relative entropy (relative entropy), is a method to describe the difference between two probability distributions P and Q. It is asymmetric, which means D(P||Q) ≠ D(Q||P). In particular, in information theory, D(P||Q) represents the information loss that occurs when the probability distribution Q is used to fit the true distribution P, where P represents the true distribution, and Q represents the fitted distribution of P. Some people call KL divergence KL distance, but in fact, KL divergence does not satisfy the concept of distance. It should be: 1) KL divergence is not symmetric; 2) KL divergence does not satisfy the triangle inequality. For a discrete random variable or two probability distributions P and Q of a continuous random variable, the definitions of KL divergence are as follows.

$$D(P\,||\,Q)=\sum_{i\in X} P(i)*[\log(\frac{P(i)}{Q(i)})] \tag{18}$$

$$D(P\,||\,Q)=\int_x P(x)*[\log(\frac{P(i)}{Q(i)})]dx \tag{19}$$

KL divergence has its own clear physical meaning in information theory. It is used to measure the number of extra Bits required to encode samples from P distribution on average using Q-distribution-based coding. And its physical meaning in the field of machine learning is to measure the similarity or similarity of two functions, and it is also frequently used in functional analysis [2]. In Shannon Information Theory, P-based encoding is used to encode samples from P. The average number of bits required for the optimal encoding (that is, the entropy of this character set) is:

$$H(x)=\sum_{x\in X} P(x)*\log(\frac{1}{P(x)}) \tag{20}$$

Using P-based coding to encode samples from Q, the number of bits required becomes:

$$H'(x)=\sum_{x\in X} P(x)*\log(\frac{1}{Q(x)}) \tag{21}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

Thus, we can get the KL divergence of P and Q:

$$D(P \| Q) = H^{'}(x) - H(x) = \sum_{x \in X} P(x) * [\log(\frac{P(x)}{Q(x)})] \tag{22}$$

6.1.2 Canonical Correlation Analysis

Correlation component analysis is extended from the correlation coefficient in probability theory. In probability theory, when studying the linear correlation between two variables, the concept of correlation coefficient is proposed. To generalize, if the linear correlation between a variable and multiple random variables is studied, the concept of full correlation coefficient (or multiple correlation coefficient) is proposed. Then, in 1936, there was a mathematician named Hotelling, who made further promotion, studying the linear correlation between multiple random variables and multiple random variables, and proposed the theory of classical correlation analysis.

Classical correlation analysis is a multivariate statistical method to study the correlation between two groups of variables. To study the correlation between two sets of variables, there are two methods: 1. List a table, just like studying the covariance matrix, this table contains the correlation between any two variables of the two sets of variables. Then, analyze based on this correlation coefficient table. 2. Like principal component analysis pca, in each group of variables, a number of comprehensive indicators are selected, and these comprehensive indicators are formed by linear combinations of variables. Study the linear relationship between variables by studying the relationship between the two sets of comprehensive indicators.

The specific step is to find the linear combination of each set of variables, and then the linear combination of the two sets of variables has the greatest correlation. This combination is not unique, and it may also cover all the characteristics of the variable.

Continue to seek the linear combination of each set of variables, and this time, the linear combination of the two sets of variables is required to have the greatest correlation, and it should be irrelevant to the combination found the first time. That is, there is a relationship similar to orthogonal.

$$X = w_{x_1} x_1 + ... + w_{x_m} x_m = w_x^T x \tag{23}$$

$$Y = w_{y_1} y_1 + ... + w_{y_m} y_m = w_y^T y \tag{24}$$

Canonical correlation analysis is called canonical because the new variables obtained by the linear combination of each group of variables, X and Y are called canonical variables. The specific calculation process is as follows:

Therefore, what we have to solve is to find the maximum correlation coefficient ρ of X and Y, namely:

$$\max_{w_x, w_y} \rho = \frac{C[X,Y]}{\sqrt{V[X]V[Y]}} \tag{25}$$

Where V is the variance and C is the covariance. Since the variance V and the covariance C are constant for increasing or decreasing a constant term, we can consider their average value to be 0. Then the above formula can be reduced to:

$$\rho = \frac{E[XY]}{\sqrt{E[X^2]E[Y^2]}} = \frac{E[w_x^T x y^T w_y]}{\sqrt{E[w_x^T x x^T w_x]E[w_y^T y y^T w_y]}} = \frac{w_x^T C_{xy} w_y}{\sqrt{w_x^T C_{xx} w_x w_y^T C_{yy} w_y}} \tag{26}$$

Among them: C_{xx} and C_{xy}, C_{yx} are the autocorrelation coefficient matrix and the cross-correlation matrix respectively. After obtaining the partial derivatives of w_x and w_y from the above formula, we get the following formula:

$$C_{xy} w_y = \rho \lambda_x C_{xx} w_x \tag{27}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

$$C_{yx}w_x = \rho\lambda_y C_{yy}w_y \tag{28}$$

Among them:

$$\lambda_x = \lambda_y^{-1} = \sqrt{\frac{w_y C_{yy} w_y}{w_x C_{yy} w_x}} \tag{29}$$

Simply scaling the linear coefficients of Equations 1 and 2 will not affect the correlation coefficient value. Therefore, the norms of w_x and w_y can be freely chosen. Naturally, we choose wx. For equations 5 and 6, you can use the following formula to solve:

$$C_{xx}^{-1}C_{xy}C_{yy}^{-1}C_{yx}w_x = \rho^2 w_x \tag{30}$$

$$C_{yy}^{-1}C_{yx}C_{xx}^{-1}C_{xy}w_y = \rho^2 w_y \tag{31}$$

Where w_x and w_y are the feature vectors of the above two formulas, and ρ is the square of the correlation coefficient. In application, after solving equations 8 and 9 to obtain w_x and ρ^2, substitute them into equations 5 and 6 to obtain w_y. It should be noted that we have obtained several results, and there are several feature vectors. In fact, we have $\min(m,n)$ results, but there is a set of corresponding w_x and w_y corresponding to the largest correlation coefficient. Here we only take the largest correlation coefficient. From this, we obtain the correlation coefficient ρ of the required two sets of data.

6.2 Error Analysis Result

After completing the fitting, compare the fitted curve with the points after adding noise, and analyze the errors of the longitude and latitude of the points as shown in the figure below.

The abscissa represents time, and the ordinate represents latitude. The average radar error is -0.000115798 degrees, and the overall average error is -0.000118618 degrees.

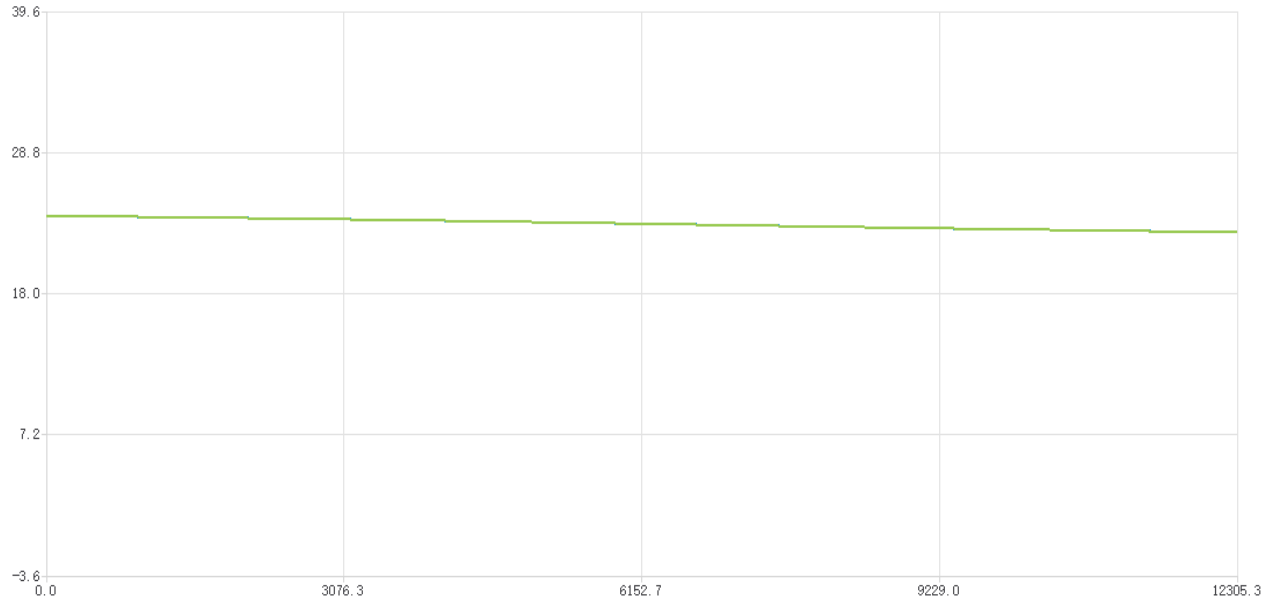

Fig. 5 Ordinate error analysis

The abscissa represents time, and the ordinate represents longitude. The average radar error is -0.000434612 degrees, and the overall average error is -0.000453411 degrees.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

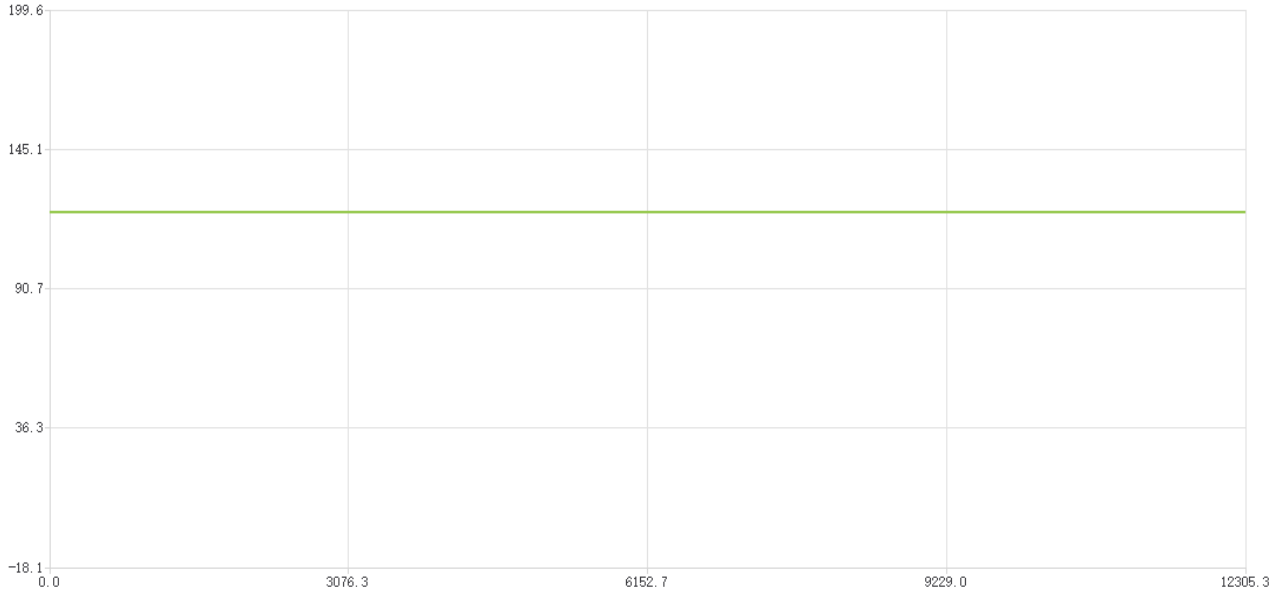

Fig. 6 Longitude error analysis

7. Conclusion

The article uses a method based on the local master curve to fit the trajectory of the maneuvering target detected by the radar. The local master curve proves that the concept of combining the local principal component and the mean shift is a simple and useful tool for calculating the master curve. Compared with most other master curve algorithms, this tool shows superior performance in simulated data sets. The higher the order of curve fitting using the least square method, the higher the goodness of the fitting, and the better the fitting curve can reflect the changing trend of the observed data. The experimental results show that the flight trajectory obtained by the curve fitting method has little error compared with the local master curve obtained from the discrete point set, which proves that the master curve can well reflect the trajectory of the target detected by sensors such as radar.

Acknowledgements

I want to thank everyone who helped me a lot with my thesis. First, I am most grateful to my mentor, who provided me with a wealth of suggestions and criticisms for writing the paper. I would also like to express my heartfelt thanks to the teachers of the School of Information Engineering, Shanghai Maritime University, who together led me into the academic temple. I am very grateful to my parents. Without their intellectual and material support, I could hardly complete the paper on time. Finally, I want to thank all my classmates, because their endless support and enthusiasm makes me feel like living in a big family.

References

[1] HESTIE T,STUETZLE W. Principle curves[J]. Journal of the American Statistical Association, 1989, 84(406): 502-516.

[2] DUCHAMP T, STUETZLE W.Extremal properties of principal curves in the plane[J]. Annals of Statistics, 1996, 24(4):1511-1520.

[3] K'EGL B,KRZYZAK A et al.A polygonal line algorithm for constructing Principal curves.In: proceedings of Neural Information Processing Systems[J].Dever Colorado, USA,1999,501-507.

[4] CHANG K,GHOSH J.Principal curves classufuer-a nonline approach to pattern classification[A].In IEEE International Joint Conference on Neural Networks,1998, 695-700.

[5] TIBSHIRANI R.Principal curves revisited [J].Statistics and Computation,1992,2:183-190.

[6] PIEGL L A,TILLER W. The NURBS book[M]//The NURBS, Book. Springer Berlin Heidelberg, 1997: 47-116.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202103_7(3).0017

[7] LANCASTER P,SALKAUSK AS K. Curve and surface fitting: An introduction[M].London: Academic Press, 1986:87-107.

[8] WILLIAMSON J H. Least-squares fitting of a straight line[J]. Canadian Journal of Physics, 1968, 46(16): 1845-1847.

[9] LIN H W,WANG G G,DONG C S. Constructing iterative non-uniform B-spline curve and surface to fit data points[J].Science in China(Series F:Information Sciences), 2004,47(3):315-331.

[10] LIN H W,BAO H J,WANG G J.Totally positive bases and progressive iteration approximation[J]. Computer & Mathematics with Applications, 2005, 50(3):575-586.

[11] DENG C Y,LIN H W. Progressive and iterative approximation for least squares B-spline curve and surface fitting[J]. Computer-Aided Design, 2014,47(1):32-44.

[12] STOER J, BULIRCH R. Introduction to numerical analysis[M]. New York:Springer Science & Business Media, 2013:570-643.