Zhang, Tao; Cai, Baigen; Lu, Debiao; Wang, Jian; Xiao, Yu

Train Localization Environmental Scenario Identification Using Features Extracted from Historical Data

Published in:
China Satellite Navigation Conference, CSNC 2021, Proceedings

DOI:
10.1007/978-981-16-3138-2_2

Published: 01/01/2021

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Zhang, T., Cai, B., Lu, D., Wang, J., & Xiao, Y. (2021). Train Localization Environmental Scenario Identification Using Features Extracted from Historical Data. In C. Yang, & J. Xie (Eds.), *China Satellite Navigation Conference, CSNC 2021, Proceedings* (pp. 12-21). (Lecture Notes in Electrical Engineering; Vol. 772 LNEE). Springer. https://doi.org/10.1007/978-981-16-3138-2_2

Train Localization Environmental Scenario Identification Using Features Extracted from Historical Data

Tao Zhang[1(✉)], Baigen Cai, Debiao Lu, Jian Wang, Yu Xiao

Abstract. The application of Global Navigation Satellite System (GNSS) on the railway greatly reduces the cost on train localization. However, the railway environment is complex and changes with the train movement, buildings, trees, railroad cuts and mountains will block and reflect the GNSS signals, which will bring errors to the GNSS-based train position estimation. This paper proposes a railway scenario identification method based on historical GNSS receiver observation data to identify scenarios along the railway. Firstly, a railway environment scenario parameter model library is established according to Feature of Sky Occlusion (FSO) of typical scenarios, apply historical GNSS observation data along the railway to establish the FSO models of scenario segments, and generate FSO feature sequences. The dynamic time warping algorithm (DTW) is used to match the FSO parameter model of the scenario segment with the FSO model library. This paper collected data from field experiments at Beijing Sanjiadian station to verify the algorithm. The scenario identification results showed that the scenario identification method based on DTW can effectively identify the railway scenarios.

Keywords: train localization; GNSS; feature of sky occlusion; dynamic time warping algorithm; scenarios identification.

---

[1] Tao Zhang (✉), Debiao Lu, Jian Wang
Beijing Engineering Research Center of EMC and GNSS Technology for Rail Transport,
School of Electronic and Information Engineering,
Beijing Jiaotong University, 100044 Beijing, China
e-mail: 19120277@bjtu.edu.cn

Baigen Cai
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, 100044 Beijing, China

Yu Xiao
Department of Communications and Networking, School of Electrical Engineering,
Aalto University, Espoo, Finland

1 Introduction

China railway network covers a wide range of the country and the train operation environment is complex. Real-time and accuracy train location performances are the guarantee for safe train operation. With the modern development of train operation control system, higher requirements are put forward for the positioning performance. GNSS has high accuracy / reliability and has been introduced into the railway field for train positioning, timing, and track survey etc. [1].

In open environments, satellite navigation signal reception is good, GNSS positioning can meet the railway positioning requirement. However, during the operation of the train, through various scenarios, including forests, valleys, road cuts, urban canyons, tunnels, etc., which lead to limited satellite navigation signal propagation, resulting in different degrees of degradation or even failure of GNSS-based train positioning [2]. Different degradation degrees of satellite positioning can be described by quantitative methods through parametric modelling of railway scenarios, and establishing environmental scenario models based on Feature of Sky Occlusion (FSO), using satellite azimuth, altitude and signal observation parameters such as intensity describe the occlusion boundary of the sky [3]. Identify the railway scenario according to the FSO parameter model and evaluate the satellite positioning performance.

This paper classifies railway scenarios according to FSO, proposes a satellite positioning environment scenario identification method in train operation scenarios, establishes a railway environment scenario FSO parameter model library and a railway scenario segment parameter model, and uses historical railway receiver observation data to this method authenticating.

2 FSO Model of Railway Scenarios

2.1 Railway Scenarios Classification

The environmental scenarios along the railway mainly include plains, mountains, cities, forests, etc. Due to varying degrees of occlusion and reflection of satellites, the positioning performance are different in each type of environmental scenario. According to the FSO of the railway environment scenario, the satellite positioning scenarios along the railway can be categorized into 5 kinds:

Unobstructed scenario (S1): No objects on both sides of the track block the satellite signals, and the positioning performance is good.

Shallow occlusion on both sides scenario (S2): There are low mountains, buildings, and other occlusions on both sides of the track, which will occlude satellites with low elevations, and will reduce the number of visible satellites and reduce the satellite positioning performance.

One side is deep and the other side is shallowly occluded scenario (S3): One side of the track is blocked by mountains or tall buildings, and the other side of the track is obscured by a low target.

Depth occlusion on both sides scenario (S4): There are high slopes, mountains, tall buildings and other targets on both sides of the strand, he receiver can only receive high-elevation satellites.

Fully obscured scenario (S5): The tunnel scenario, in which the satellite signal is fully occluded, and the satellite positioning is invalid.

As shown in Fig. 1, skyplot of the S1~S5 scenarios covered by the occlusion degree is presented. The grey part represents the occluded sky.

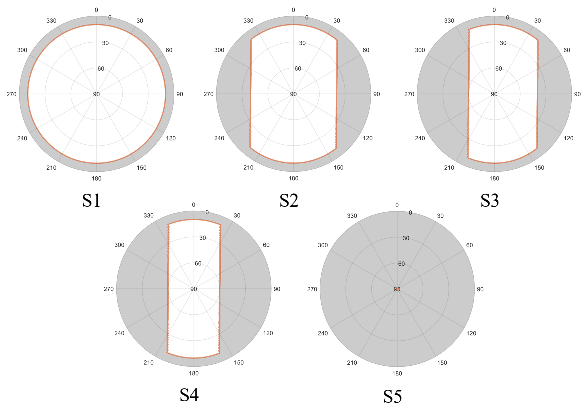

Fig. 1. Skyplot with occlusion for designated 5 railway scenarios

2.2 Construct FSO Model Library

The FSO parameter model is a 360-dimensional discrete sequence that parameterizes the sky occlusion boundary of the railway environment scenario [4]. The FSO of different types of railway environment scenarios have their own distinct characteristics. To accurately identify the railway environment scenarios, this paper idealizes the FSO of the S1~S5 scenarios. The sky occlusion boundary of scenario S2~S4 contains linear boundary as an assumption, so the parameters of geometric model are ideally set as linear boundary for more compliable model libraries. As shown in Fig. 2, l represents the distance from point O to boundary AB, R represents the radius of sky plot, and θ represents Azimuth. The scenario parameter model sequence is sampled with the azimuth resolution at 1°, and the formula (1) represents the relationship between elevation and azimuth of the straight-line boundary of the sampling sequence. The value range of θ is in the range of 0° to

90°, and FSO parameter models of S2~S4 scenarios are constructed through symmetry.

$$ele_i = 90 \times (1 - \frac{x}{R \cdot \cos\theta_i}) \tag{1}$$

Where, $\theta_i = (0, 1, \cdots, [\arccos(1 - \frac{x}{R})])$

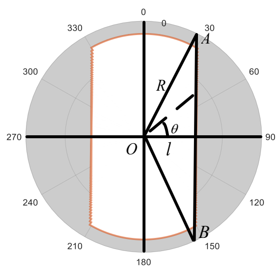

Fig. 2 Model of sky occlusion line boundary description

The FSO parameter models of the S1~S5 scenarios constitute the FSO parameter model library $\Upsilon = (\Theta_1, \Theta_2, \cdots, \Theta_5)$, where $\Theta = (\theta_1, \theta_2, \cdots, \theta_{360})$. The FSO parameter of each scenario is a 360-dimensional discrete sequence, representing the elevation of the sky occlusion boundary corresponding to 360 azimuths.

2.3 FSO Modeling of Scenarios Along the Railway based on Observation Data

The scenario along the railway can be ideally regarded as composed by S1~S5 scenario segments alternately. The idea of differentiation between S1 to S5 can be used to divide the railway scenario into several short scenario segments, and the scenario type of each scenario segment is identified, finally the scenario composition is determined for the entire trajectory. According to the digital track map (DTM) the scenarios along the railway are segmented into scenario segments at equal intervals $S = \{seg_1, seg_2, \cdots, seg_n\}$, as shown in Fig. 3.

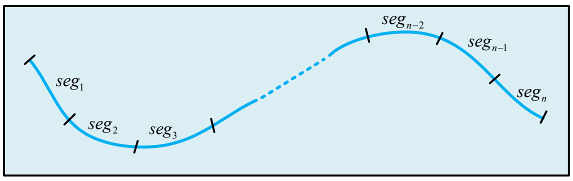

Fig. 3. Schematic diagram of the division of scenarios along the railway

According to the observation data, the satellite information of all epochs in each scenario segment is extracted [5]. Extract the total number of visible satellites in the current epoch, visible satellite pseudo-random noise (Pseudo-Random Noise, PRN), elevation, azimuth, signal-to-noise ratio. Assuming that a total of n scenario segments is divided, i denote the i^{th} scenario segment, and j denote the j^{th} visible satellite in the i^{th} scenario segment. In the i^{th} scenario segment, the number of visible satellites is SV_{ij}, and the elevation EL_{ij}, azimuth AZ_{ij}, and signal-to-noise ratio SNR_{ij} corresponding to the j^{th} visible satellites are extracted. Each scenario segment can be described using a set of parameter vector as $\Omega_i = \{SV_i, \{EL_{ij}, AZ_{ij}, SNR_{ij} \mid j = 1, 2, \cdots SV_i\}\}$.

Due to the inclination of the satellite trajectory, the receiver can only receive satellites with high elevations in the range of 0°~30° azimuth and 330°~360°, so only satellites within 30°~330° are considered for identification in this paper. The algorithm flow of FSO model sequence construction is shown in Table 1.

Table 1. Railway environmental feature parameters construction algorithm

Input: a set of segment information of a given line scenario,
$\Gamma = \{\Omega_1, \Omega_2, \cdots, \Omega_n\}$
Output: FSO parameter model collection of scenario segment,
$\Psi = \{\varphi_1, \varphi_2, \cdots, \varphi_n\}$

1. Set the azimuth division range;
2. Divide the azimuth range of the skyplot from 30° to 330° into α regions;;
3. For i=1 to n **do**
4. Eliminate low elevation and low signal-to-noise ratio satellites to get the set Γ' Γ';
5. For m=1 to α **do**
6. Traverse Γ' and get the satellite elevation set $\{EL_{im}\}$ according to $(m-1)\cdot 300/\alpha \le AZ < m\cdot 300/\alpha$;
7. End for
8. For m=1 to α **do**
9. If $\{EL_{im}\}$ is not empty then
10. Take the lowest elevation as the threshold of the visible elevation: $\beta_{im} = \min\{EL_{im}\}$;
11. Else
    There are no visible satellites in the range of this direction angle, $\beta_{im} = (\beta_{i(m-1)} + \beta_{i(m+1)})/2$;
12. End if
    Each azimuth threshold is divided at an interval of 1°, and the threshold of the elevation corresponding to all azimuths is equal to β_{im};
13. End for
14. Construct the environment feature vector of the i-th scenario segment: $\varphi_i = \{(\beta_{i1}, \cdots \beta_{i1}), (\beta_{i2}, \cdots, \beta_{i2}), \cdots, (\beta_{i\alpha}, \cdots \beta_{i\alpha})\}$;
15. End for
16. Construction of the FSO collection of all scenario segments

$\Psi = \{\varphi_1, \varphi_2, \cdots, \varphi_n\}$。

The FSO parameter model of each scenario segment is a 300/λ-dimensional elevation sequence, λ is the azimuth resolution.

3 Railway Scenario Identification based on FSO Model Matching

The FSO parameter model of the scenario segment is a 300/λ-dimensional discrete sequence. The model in the FSO parameter model library is a 360-dimensional discrete sequence. When the lengths of the discrete sequences are not equal, the Euclidean distance cannot be used to calculate the sequence similarity. This paper uses the DTW algorithm to match the FSO parameter model of the scenario segment with the FSO parameter model library to identify the scenario along the railway.

3.1 DTW Algorithm

Dynamic Time Warping (DTW) can measure the similarity of two unequal sequences [6]. The main ideas of DTW are as follows:

1. Assume there are two discrete sequences $A = \{a_1, a_2, a_3, \cdots, a_m\}$ and $B = \{b_1, b_2, b_3, \cdots, b_n\}$.
2. Euclidean distance to calculate the distance between two sequence points: $D(a_i, b_j)$, where $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, as shown in Table 2.

Table 2. Point-to-point Euclidean distance table of two discrete sequences

$D(a_1, b_1)$	$D(a_1, b_2)$	$\cdots$	$D(a_1, b_n)$
$D(a_2, b_1)$	$D(a_2, b_2)$	$\cdots$	$D(a_2, b_n)$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
$D(a_m, b_1)$	$D(a_m, b_2)$	$\cdots$	$D(a_m, b_n)$

3. Find the shortest path according to Table 2. Follow a certain path from $D(a_1, a_2)$ to $D(a_m, b_n)$. Find the path satisfies: if the current node is $D(a_i, b_j)$, then the next node must be at $D(a_{i+1}, b_j)$, $D(a_i, b_{j+1})$ or $D(a_{i+1}, b_{j+1})$, and the path must be the shortest. The calculation is based on the idea of dynamic programming.
4. Find the best output path back from the final shortest distance, from $D(a_1, b_1)$ to $D(a_m, b_n)$. The sum of them is the required DTW distance.

3.2 Scenario Identification Based on DTW Algorithm

The FSO parameter model library is $\Upsilon = \{\Theta_1, \Theta_2, \cdots, \Theta_5\}$, where Υ_i represents the parameter model of the S1~S5 scenario, and $\Theta_i = \{\theta_1, \theta_2, \cdots, \theta_{360} | i=1, \cdots, 5\}$ is the discrete elevation sequence, and the sequence length is 360. The set of FSO parameter models of the scenario segment $\Psi = \{\varphi_1, \varphi_2, \cdots, \varphi_K\}$, K is the total number of scenario segments along the railway, and the elevation sequence of the FSO parameter model of each scenario segment is $\varphi = \{(\beta_1, \cdots, \beta_1), (\beta_2, \cdots, \beta_2), \cdots, (\beta_\alpha, \cdots, \beta_\alpha)\}$, the sequence length is 300/λ.

The sequence length of the FSO parameter model of the scenario segment is different from that of the FSO parameter model library. The DTW algorithm is used to match the discrete sequence of the FSO parameter model of unequal length to identify the type of each scenario segment. The scenario identification algorithm is shown in Table 3.

Table 3. Process of railway scenarios identification algorithm

Input: FSO parameter model library for typical scenarios,
$\Upsilon = (\gamma_1, \gamma_2, \cdots, \gamma_5)$
FSO parameter model of the scenario segment to be tested, $\Phi = (\varphi_1, \varphi_2, \cdots, \varphi_K)$
Output: scenario identification result
1. For i =1 to K **do**
2. Select the features of different categories of scenario classification results φ_i;
3. For j =1 to 5 **do**
4. Use Euclidean distance to calculate the point-to-point distance table $D(\theta_i, \beta_j)$ of sequence φ_i and sequence γ_j;
5. Find the shortest path in the Euclidean distance table $D(\theta_i, \beta_j)$;
6. Calculate the DTW distance dis (φ_i, γ_j) of two discrete sequences based on the shortest path;
7. End for
8. Find the best matching reference feature Min $\{\text{dis}(\varphi_i, \gamma_j) \mid j=1,2\cdots,5\}$;
9. End for
10. Output scenario identification results.

4 Experiment Results and Analysis

4.1 FSO Parameter Modeling Library

The FSO parameter model library includes S1~S5 scenario parameter model. Based on skyplot observations, the lowest threshold of the elevation of the 5 scenarios is set to 10°. The sky occlusion boundary of the S2~S4 scenario contains an ideal linear boundary. Set the maximum elevation of the linear occlusion boundary. The maximum elevation of the two linear boundaries of the S2 scenario is 40°, and the maximum of the 2 linear boundaries of the S3 scenario. The elevations are 40° and 60°, respectively, and the maximum elevation of the two straight line boundaries of the S4 scenario are both 60°. The sky occlusion boundary is sampled at 1° azimuth intervals, and the parameter model of each scenario is $\Theta_i = (\theta_1, \theta_2, \cdots, \theta_{360})$.

4.2 Experimental Data Collection and Scenario modeling

The experiment was carried out at Beijing Sanjiadian station, using Spirent GSS6450 to record 4 satellite constellation RF signals. Two receivers are used for the comparison of the historical data collected by GSS 6450. U-blox M8N receiver receives BDS and GPS signals, and Septentrio AsteRx-m2a receives signals from BDS, GPS, GLONASS, and Galileo navigation systems.

According to the digital track map, the scenario along the segment of the railway is equally segmented into 5 meters, and a total of 326 segments are divided. According to two pieces of data, FSO parameter modelling corresponding to each scenario segment. Excluding the satellites whose elevation is less than 10° and SNR less than 20dbm, the azimuth of the skyplot is divided into 6 areas, each area is 30°, and then the FSO parameter model is constructed by interpolation, and the interpolation interval is 1° azimuth resolution. The sequence length is 300.

4.3 Scenario Identification

Since there is no positioning result output for the S5 full occlusion scenario, the S5 scenario segment can be identified according to the missing positioning segment, so scenario identification algorithm is performed on S1~S4, and S5 is labelled according to this rule. The FSO parameter model of each scenario segment is matched with the S1~S4 scenario models in the FSO parameter model library. When the DTW distance is the shortest, the type of the scenario is identified.

As shown in Table 4, 326 scenario segments are identified, and length of each type of scenario segment is identified based on the receiver observation data. The field-tested railway scenario, the length of each scenario is determined according to the laser point cloud data. The lengths of the S4 and S5 scenario segments identified from the receiver data are the same as the actual scenario. S4 is a scenario with depth occlusion on both sides, with a length of 5 meters. In the actual scenario, there are high walls on both sides of the track, and the identification result is consistent with the actual scenario. S5 is a full occlusion scenario with a length of 20 meters, which is consistent with the actual tunnel scenario. The deviations of S1~S3 scenario segments are 6.12%, 1.86%, 5.17%, respectively, and the degree of difference is small, indicating the effectiveness of the algorithm.

Table 4. Result of railway scenarios segment

Data acquisition receiver	S1	S2	S3	S4	S5
U-blox M8N	38	232	51	1	4
Septentrio AsteRx-m2	49	211	61	1	4
Difference	22.45%	9.95%	16.39%	0.00%	0.00%

As shown in Fig. 4, the DTW distance between the FSO model sequence of the S1~S4 scenario segment set identified from the observation data and the S1~S4 scenario FSO model sequence in the FSO model library.

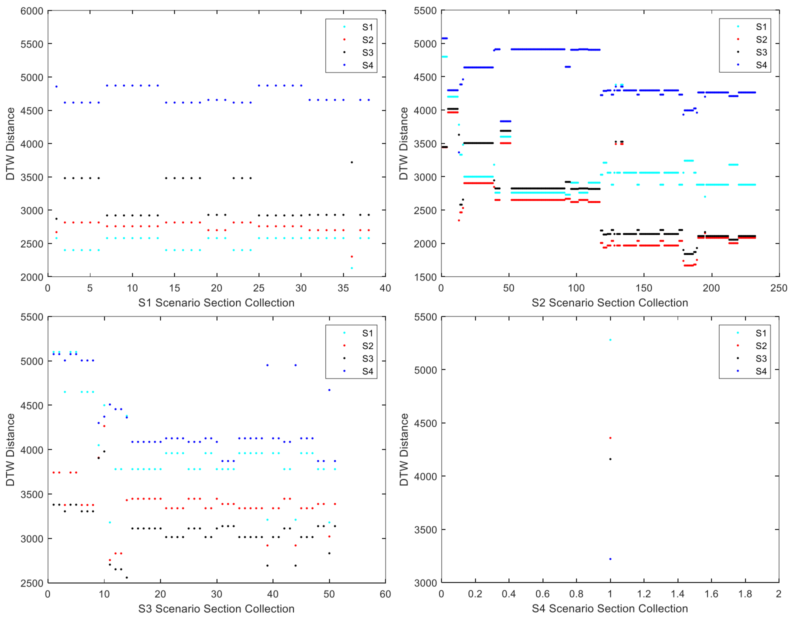

Fig. 4. DTW distance between scenario segment and FSO scenario library

5 Conclusion

The FSO parameter model is established to identify the scenario along the railway is proposed. According to the FSO parameter model library of the scenario along the railway with environmental characteristics, the scenario along the Beijing Sanjiadian railway is divided into scenario segments at equal intervals, and FSO parameter modeling is performed on each scenario segment according to the receiver observation data. The DTW algorithm is used to match the FSO parameter model of the scenario segment with the FSO parameter model library, to identify the type of the scenario segment along the railway and compare it with the actual environmental scenario along the railway to prove the effectiveness of the scenario identification method.

Acknowledgement. This paper is supported by National Key Research and Development Program of China (2018YFB1201500), Beijing Science Program of Beijing Municipal Science and Technology (Z181100001018032), National Natural Science Foundation of China (U1934222, 61873023), Beijing Natural Science Foundation (L191014), and Beijing Nova Program of Science and Technology (Z191100001119066).

References

1. Liu J, Cai B, & Wang J (2014). Status and development of satellite navigation system based train positioning technology. Journal of Central South University, pp. 4033-4041.
2. Tan J (2019). Research on Satellite Positioning Availability Evaluation and Prediction Method in Railway Scenario. Beijing Jiaotong University.
3. Wu B, Lu D, & Cai B (2018). 3D Map-Aided GNSS Train Positioning Integrity Analysis. China Satellite Navigation Conference.
4. Xu H, Hsu L, & Lu D (2020). Sky visibility estimation based on GNSS satellite visibility: an approach of GNSS-based context awareness. GPS Solutions, pp. 1-15.
5. Jiang S (2019). Research on GNSS Signal Propagation Trajectory Error Modeling in Train Dynamic Operation Environment. Beijing Jiaotong University.
6. M. Müller (2007). Dynamic Time Warping, in Information Retrieval for Music. Springer Berlin Heidelberg: Berlin, Heidelberg, pp. 69-84.