Vessel Trajectory Data Compression based on Course Alteration Recognition

Le Qi [1], Zhongyi Zheng [1,*]

1 Navigation College
Dalian Maritime University
Dalian 116026, China

Abstract **— The vessel trajectory data from automatic identification system (AIS) is massive and contains much redundant information, which makes it difficult to use the data in research and applications. To solve the problem, a data compression algorithm for vessel trajectory was proposed based on vessel course alteration recognition. To recognize the vessel course alteration, the characteristics of trajectory during vessel course alteration were analyzed. It was found that the vessel course alteration on trajectory was similar to the corner on line, so the corner detection algorithm was referenced. Then the representative point of course alteration on trajectory was determined and kept. Other points on trajectories were deleted as redundant points, thus the data of vessel trajectory was compressed. In order to compare the performance of this algorithm with traditional compression algorithms, experiments were implemented base on actual trajectories of vessels from AIS. The results show that the data compression ratio of this algorithm is higher than traditional compression algorithms in vessel trajectory data compression. And this algorithm has a strong adaptability to different voyages of trajectories. The experiments demonstrate that the proposed algorithm is effective and promising for the study of maritime traffic based on AIS.**

Keywords - Vessel trajectory; Data compression; Vessel course alteration recognition; Corner detection algorithm

I. INTRODUCTION

With the development of big data, knowledge mining and machine learning theory, it is research hotspot to use vessel trajectory data effectively for improving the safety and efficiency of maritime traffic [1-3]. Vessel trajectory, which is the route of a vessel and contains motion features, is the vital data source in maritime traffic study. With the improvement of maritime traffic monitoring systems, especially the automatic identification system (AIS), a large number of vessel trajectories are recorded. However, the information frequency of AIS is very high, so vast amounts of data and redundant information are produced, which makes the data from AIS is difficult to be used directly. Therefore it is necessary to compress the data of vessel trajectory from AIS. Moreover, in practical application, e.g., AIS data importing, analysis and display in electronic chart display and information system (ECDIS), the vessel trajectory data also needs to be compressed.

II. STATE OF THE ART

Trajectory is a kind of vector data. The traditional algorithms for vector data compression are choosing interval points algorithm, light bar algorithm, vertical distance algorithm, Douglas Peucker algorithm and offset angle algorithm [4, 5]. Other algorithms mostly base on the five algorithms and are improved according to the practical application. At present, the study of vessel trajectory compression mainly concentrates on solving issues in practical application based on Douglas Peucker algorithm. However, Douglas Peucker algorithm and other traditional algorithms only consider the features of vector data, the vessel trajectory contains other features, e.g., steering, directly shipping etc., which can not be effectively dealt with by them. To solve the problem, a new trajectory data compression algorithm was proposed based on vessel course alteration recognition in this paper. By analyzing the characteristics of trajectory during vessel course alteration, it is found that corners on line are similar to the point on trajectory where vessel altered its course. So the course alteration recognition is achieved based on corner detection algorithm and characteristics of course alteration. Corner detection algorithm is widely used in computer vision [6-10], intelligent identification [11, 12], retrieval, etc. Corner detection algorithm has many advantages, e.g., stronger adaptive performance, not affected by curve space scale, etc. [13-17] The algorithm suitable for image edge corner detection, object contour, etc. [18-23], but not suitable for course alteration recognition absolutely. So a improvement of corner algorithm is studied. Base on real vessel trajectory data from AIS, the new algorithm is compared with traditional compression algorithms through experiment. And the results show that the compression algorithm based on course alteration recognition has better performance.

The remainder of this paper is organized as follows. Section 3 describes the methodology of the data compression algorithm. Section 4 presents the experiments in which the performance of the algorithm is tested and compared with five traditional algorithms, and the results are analyzed and discussed. Conclusions are summarized in Section 5.

III. METHODOLOGY

(1). Vessel Course Alteration Extraction and Smooth

A vessel trajectory, trajectory A , in the Qiongzhou Strait is shown in Fig. 1. The coordinate of the i_{th} point of trajectory A is $(x_A(i), y_A(i))$. $x_A(i)$ is the longitude and $y_A(i)$ is the latitude. $i = 1, 2, \ldots, n$. Let θ_A stand for the course over ground of the vessel. $\theta_A(i)$ is the course from point $(x_A(i), y_A(i))$ to point $(x_A(i+1), y_A(i+1))$. Let the set $\theta_A = \{\theta_A(i) \mid i = 1, 2, \ldots, n-1\}$ stand for the course of the points on trajectory A.

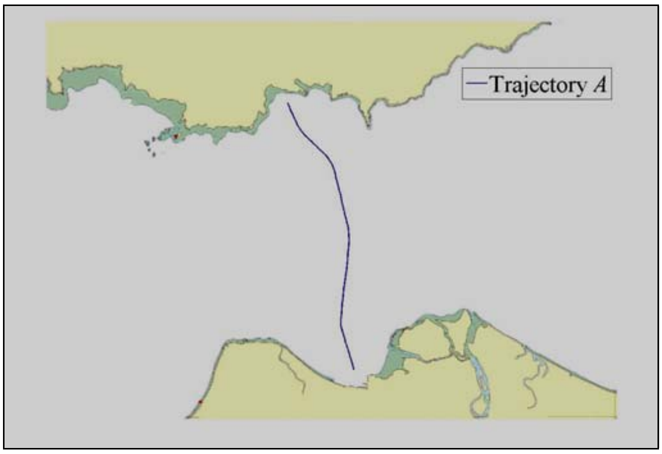


Figure 1. Trajectory A.

Let $\Delta\theta_A$ stand for the vessel course alteration. $\Delta\theta_A(i)$ is the course alteration from $\Delta\theta_A(i-1)$ to $\Delta\theta_A(i)$. If $\theta_A(i) > 0$, the vessel turned right on the point $(x_A(i), y_A(i))$. On the contrary, the vessel turned left on the point $(x_A(i), y_A(i))$. There is no course alteration information at the start and end points of trajectory. The course alteration of trajectory A can be expressed by the set $\Delta\theta_A = \{\Delta\theta_A(i) \mid i = 1, 2, \ldots, n-1\}$. The curve of course alteration, $\Delta\theta_A$, is shown in Fig. 2.

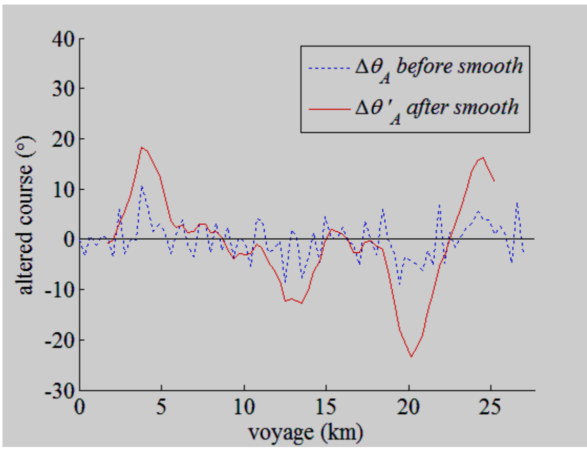


Figure 2. Curve smoothing of course alteration of trajectory A.

In order to get the macroscopic changing of course alteration curve and remove local variation and noise which may be caused by equipment error, environmental influence, etc., $\Delta\theta_A$ should be smoothed before the next step. So the purpose of this step is to smooth the curve of $\Delta\theta_A$. Let symbol $\Delta\theta_A'$ stand for the smoothing result.

$$\Delta\theta_A'(i) = \Delta\theta_A(i) \otimes \varphi(i) \tag{1}$$

Where $\otimes$ stands for convolution and $\varphi(i)$ is a low pass filter function. There are many types of filters, like triangle filter, Hanning filter, Gauss filter, Butterworth filter, Chebyshev wave record, etc. According to the motion feature of vessels, the course alteration is a slow process. Therefore, the course alteration on each point is greatly influenced by the points which are near to it. But the points which are far away from the points have little influence to its course alteration. And the degree of the influence is

DOI 10.5013/IJSSST.a.17.01.05

ISSN: 1473-804x online, 1473-8031 print

declining along with the distance growing. The Gauss filter and triangle filter can be chosen as $\varphi(i)$. In this paper the Gaussian function is taken as an example. There are two main smoothing parameters in $\varphi(i)$, which are Gaussian smoothing-scale σ [24] and q [18, 23]. And the result is $\Delta\theta_A{}' = \left\{\Delta\theta_A{}'(i) \mid i = 2+q, 3+q, \ldots, n-1-q\right\}$, as shown in Fig. 2.

(2) Course Alteration Estimation

It is known that, if a vessel alters its course, the course of the vessel keeps on increasing or decreasing during the course alteration, which is similar to the corner on line. Therefore, if θ_A keeps on increasing or decreasing during a continuous period of time, it can be considered that there may be an alteration action. Let symbol $\Delta\Theta_A$ stand for the altered course of each action, which can be calculated from $\Delta\theta_A$. Because $\Delta\theta_A$ has been smoothed to get $\Delta\theta_A{}'$. Therefore, correspondingly, $\Delta\Theta_A{}'$ can be calculated from $\Delta\theta_A{}'$.

Let the set of $\Delta\theta_A{}'$ be divided into m subsets. If the values of two adjacent elements are of same sign, they should be assigned to a same subset. Each subset is a candidate of course alteration. Let the j_{th} subset be $\left\{\Delta\theta_A{}'(h), \Delta\theta_A{}'(h+1), \ldots, \Delta\theta_A{}'(h+k)\right\}$, $2 \le h \le k \le n-1-q$.

$$\Delta\Theta_A{}'(j) = \sum_{i=h}^{h+k} \Delta\theta_A{}'(i) \tag{2}$$

$\Delta\Theta_A{}' = \left\{\Delta\Theta_A{}'(j) \mid j = 1, 2, \ldots, m\right\}$. Although $\Delta\Theta_A{}'$ is not the altered degree of each course alteration action, $\Delta\Theta_A{}'$ can reflect the deference of each course alteration, as $\Delta\theta_A{}'$ is the smoothing result of $\Delta\theta_A$. Therefore $\Delta\Theta_A{}'$ can also be used to measure the significance of altered degree of each course alteration. Table 1 is the calculation results of $\Delta\Theta_A{}'$.

TABLE I CALCULATION RESULTS OF $\Delta\Theta_A{}'$

j	$\Delta\Theta_A{}'$ (°)
1	-0.8746
2	124.9515
3	-107.3235
4	4.3802
5	-164.5168
6	90.4166

To determine the course alteration from the candidates, a threshold of $\Delta\Theta_A{}'$ is needed [25]. This threshold is used to remove the weak and insignificant course alteration from candidates. Let $\Delta\Theta_{A,threshold}{}'$ stand for the threshold. If $\left|\Delta\Theta_A{}'\right| > \Delta\Theta_{A,threshold}{}'$, there is a point where vessel altered course in the subset $\left\{\left(x_A(i), y_A(i)\right) \mid i = h, h+1, \ldots, h+k\right\}$. If the course alteration is too small, this subset is a straight line.

$$\pm\Delta\Theta_{A,threshold}{}' = mean(|\Delta\Theta_A{}'|) \pm K \cdot std(|\Delta\Theta_A{}'|), K \ge 0 \tag{3}$$

Where $mean(|\Delta\Theta_A{}'|)$ is the mean of all elements of set $|\Delta\Theta_A{}'|$, $std(|\Delta\Theta_A{}'|)$ is the deviation of all elements of set $|\Delta\Theta_A{}'|$, K is adjusting coefficient. The larger the value of K is, the less the number of course alteration is. The smaller the value of K is, the more the number of course alteration is. $K = 0$, $\Delta\Theta_{A,threshold}{}' = 82.0772$. $j = 2$, $j = 3$, $j = 5$ and $j = 6$ are the four subsets which correspond four vessel course alterations.

(3) Course Alteration Point Determination and Data Compression

The purpose of this subsection is determining where the course alteration has taken place, which means determining the point on trajectory where vessel altered course. In corner detection algorithm, the determination of corner location is a difficult and import part. In different situation, the methods of the determination are always different. Similarly, the location of course alteration should be determined by considering the motion regularity of vessels, physical significance of the course alteration, the distribution characteristics of the coordinate points, etc. Therefore, an appropriate method is proposed in this subsection.

Because the course alteration on trajectory is a process, usually there are no less than a candidate points on trajectory for a course alteration. In order to compress the data volume, only one point is kept as the representative point of course alteration. All of the other points are deleted as the redundant points. Let the j_{th} subset be $\left\{\left(x_A(i), y_A(i)\right) \mid i = h, h+1, \ldots, h+k\right\}$, $\left|\Delta\Theta_A{}'(j)\right| > \Delta\Theta_{A,threshold}{}'$. The method of locating the point where vessel altered course in this paper is to add $\Delta\theta_A{}'$ from $\Delta\theta_A{}'(h)$ to $\Delta\theta_A{}'(h+k)$ and choose i when the accumulative results is near to half of $\Delta\Theta_A{}'(j)$ as shown in Eq. (4).

$$I_j = \underset{h \le i \le h+k}{\arg\min}\left(\left|\left|\sum_{l=h}^{i} \Delta\theta_A{}'(l)\right| - \frac{\left|\Delta\Theta_A{}'(j)\right|}{2}\right|\right) \tag{4}$$

Let $\left(X_j, Y_j\right)$ be the course alteration point of j_{th} subset.

$$\begin{cases} X_j = x\left(I_j\right) \\ Y_j = y\left(I_j\right) \end{cases} \tag{5}$$

According to Eq. (4), the four course alteration points on trajectory A are recognized, as shown in Fig. 3.

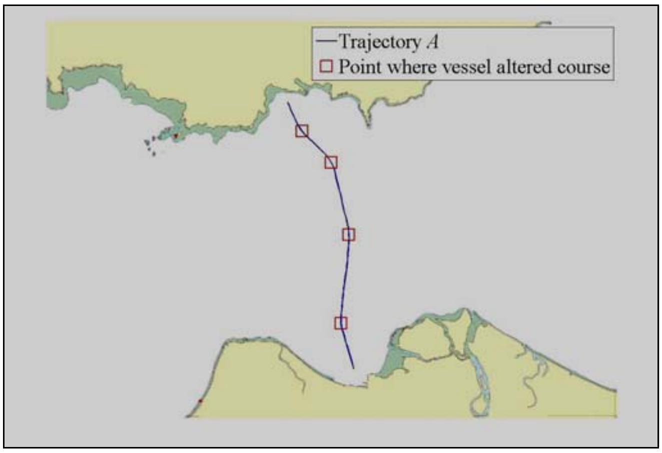


Figure 3. The points where vessel altered course on trajectory A .

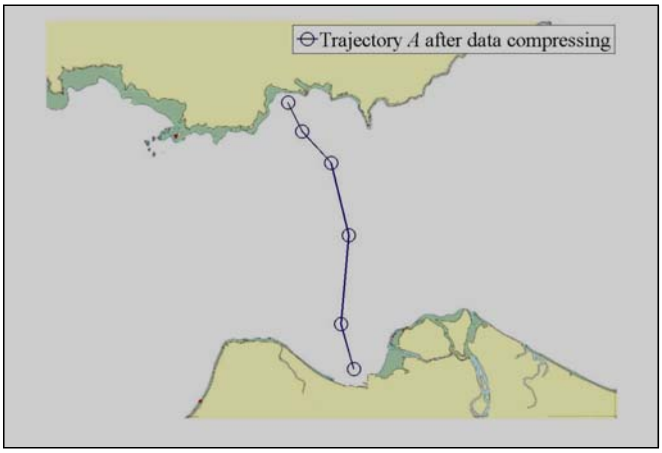


Figure 4. Trajectory A after data compressing.

The origin and destination on trajectory A are known. With the four representative points above, the trajectory can be replaced by the six points and the lines which connect them one by one, as shown in Fig. 4. All of the other points on trajectory A are deleted. Then, the data volume of trajectory A has been compressed.

IV. RESULT ANALYSIS AND DISCUSSION

To compress the data volume, the characteristics of trajectory have been analyzed. First, the voyages of trajectories are always different. Therefore the algorithm must has a strong adaptation to deal with trajectories of different size. Secondly, the course of a vessel could alter a lot in a short local segment. Generally this is caused by the collision avoidance. This phenomenon is very important for the study of maritime traffic. Therefore the local points need to be kept after the data compression. Trajectory B is taken as example to test the algorithm proposed in this paper and five traditional algorithms, which are choosing interval points algorithm, light bar algorithm, vertical distance algorithm, Douglas Peucker algorithm and offset angle algorithm. As shown in Fig. 5, seven segments in the trajectory corresponding seven course alterations. And the other parts of trajectory B are straight lines. Thus, seven representative points in the seven short local segments should be kept after the data compression according to each algorithm. The parameters in the algorithms need to be adjusted to fit the requests.

Fig. 6 (a)~(f) show the compression results of trajectory B according to the six algorithms, when the compression radio is the largest under the requests above. Nine points are kept after the compression according to course alteration recognition algorithm. Seven of the nine points are the representative points mentioned above. The other two points are the start and end points. It is obvious that the compression ratio of course alteration recognition algorithm is higher than the other traditional algorithms.

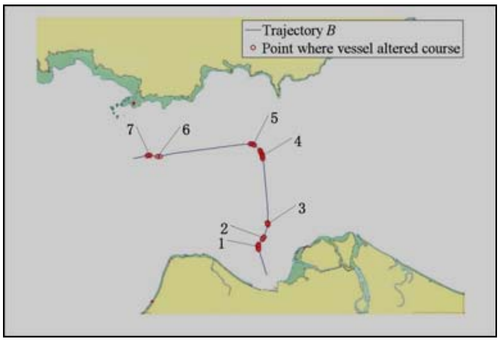
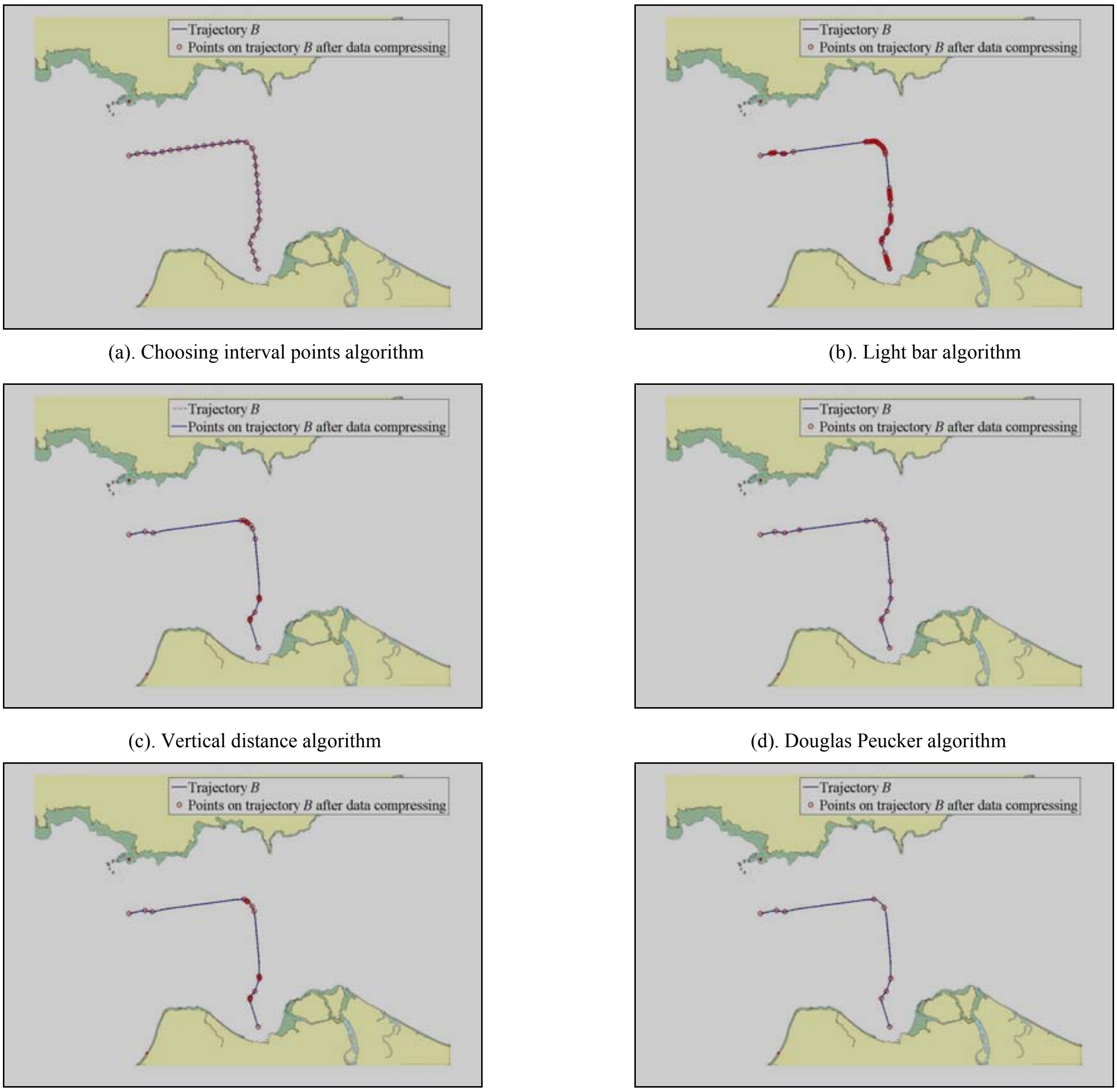


Figure 5. Trajectory B.



(a). Choosing interval points algorithm

(b). Light bar algorithm

(c). Vertical distance algorithm

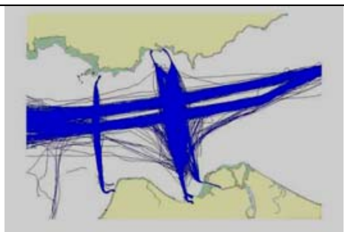
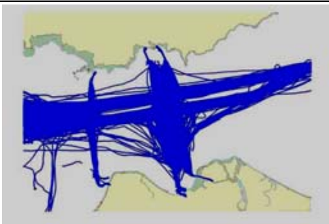
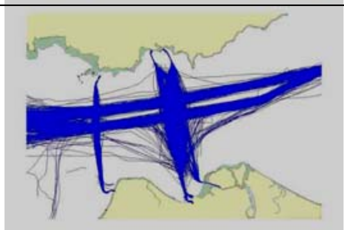
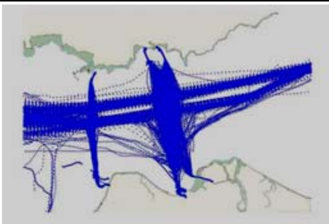
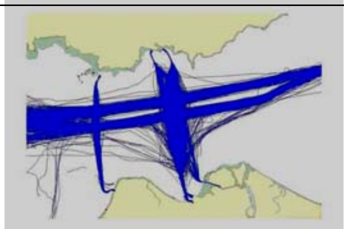
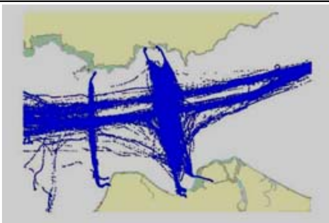
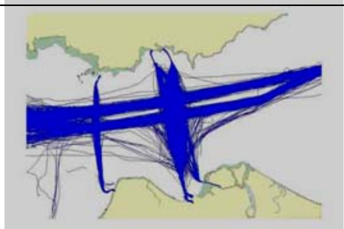
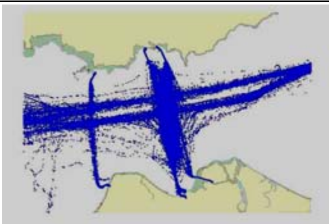
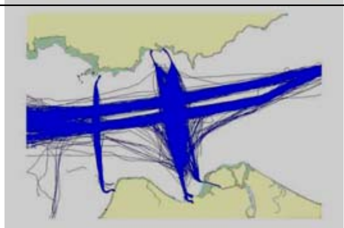
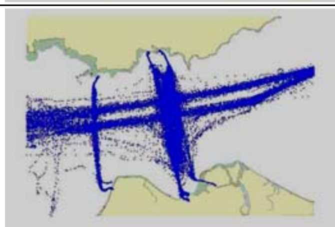
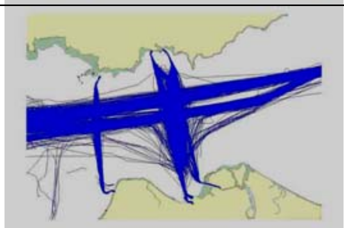
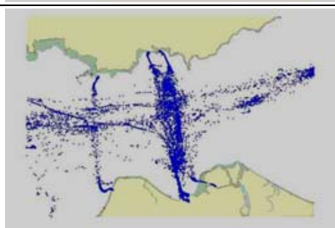
(d). Douglas Peucker algorithm

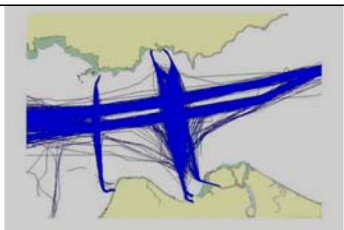
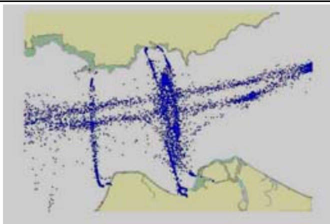
(e). Offset angle algorithm

(f). Course alteration recognition algorithm

Figure 6. Compression results of trajectory B according to six algorithms.

TABLE II COMPRESSION PERFORMANCE COMPARISON

Data compression algorithms	Trajectories	The points on trajectories	Compression ratio	Voyage changing adaptability
None (samples)			--	--
Choosing interval points algorithm			4.34	Strong
Light bar algorithm			3.93	Strong
Vertical algorithm			9.23	Weak
Douglas Peucker algorithm			12.35	Weak
Offset angle algorithm			20.53	Weak

Course alteration recognition algorithm			54.05	Strong

By taking the vessel trajectory data from AIS for example, the performance of course alteration recognition algorithm and the traditional algorithms can be tested and compared. The number of raw trajectories is 1450, and the umber of points on the trajectories is 622069. The samples and test results are shown in table 2.

The trajectories are the actual track lines of vessels, which are the fundamental information for the study of maritime traffic. Because the representative points must be kept as described above, the distortion of trajectories is very low after the compression according to each algorithm, as shown in table 2. And it is hard to find the difference of performance between each algorithm by the trajectories maps. But the points on trajectories can reflect the difference of data volume and point distribution after compression. The compression ratio, voyage changing adaptability and computational complexity are compared and analyzed based on the experimental results.

Firstly, the six algorithms ordered by the compression ratio from high to low are course alteration recognition algorithm, offset angle algorithm, Douglas Peucker algorithm, vertical algorithm, choosing interval points algorithm, light bar algorithm. The light bar algorithm is sensitive to the local course alteration of curve. It is suitable for dealing with the vector curve, whose angular is obvious. But for the smooth vector curve, the compression ratio may be decreased. Therefore the light bar algorithm is not suitable for compressing the vessel trajectory. For the choosing interval points algorithm, the point which can be kept after the compression is irregular. Therefore, to make sure all of the points where vessels altered course are kept, the only way is to decrease the compression ratio. So the compression ratio of the choosing interval points algorithm is low too. The theories of vertical algorithm and Douglas Peucker algorithm are perfect and reasonable. The compression ratios are higher than the two algorithms above. The offset angle algorithm is sensitive to the course alteration of curve, which is similar to light bar algorithm. It is sensitive to the global changing. But vessel trajectory is a kind of smooth vector curve. Even so, this algorithm can compress the data better than other traditional algorithms. The course alteration recognition algorithm can track the process of course alteration. It is sensitive to both local and global course alteration of trajectory. This feature can decrease the points which must be kept. Therefore the ratio is increased. The compression ratio of the course alteration recognition algorithm is higher than the traditional algorithms.

Secondly, the vertical algorithm and Douglas Peucker algorithm are sensitive to distance. The compression ratio of them is related to the value of distance threshold. Thus, the two algorithms have weak adaptabilities to the voyage changing. But the other four algorithms, choosing interval points algorithm, light bar algorithm, offset angle algorithm and course alteration recognition algorithm are not sensitive to distance, so they have strong voyage changing adaptabilities.

Finally, the computational complexity of Douglas Peucker algorithm is higher than other five algorithms. The reason is that Douglas Peucker algorithm is a process of recursively calling disintegrating method, which leads to a high computational complexity. But the other five algorithms are implemented by traversing the data, so the computational complexities are low.

From above, the performance of course alteration recognition algorithm is better than the five traditional algorithms. It is suitable for compressing vessel trajectories. The points where vessels altered course can be kept with a high compression ratio and automatically adapt to the voyage changing. Besides, its computational complexity is low.

V. CONCLUSION

In this paper, a new data compression algorithm is proposed to compress the trajectory data of vessels. In this algorithm, the course alteration information was kept and the other information was deleted as redundancy. 1450 trajectories were taken as samples to test the performance of this algorithm and five traditional data compression algorithms. The test results were compared and analyzed. Although the compression ratio of course alteration recognition algorithm is 54.05, which is higher than the five traditional algorithms, the distortion of trajectories after compression is very low. Because the course alteration recognition algorithm is sensitive to both global and local course alteration of trajectory, it can adapt to different voyage of trajectory with high compression ratio. Besides, the course alteration recognition algorithm consumes little computational resource, for its low computational complexity. Therefore, this algorithm can delete redundant data of vessel trajectory from AIS efficiently, which can reduce the data storage spaces and computational resources to expand the availability of AIS data in the development of maritime intelligent traffic system. Further maritime traffic research based on this algorithm is required, e.g., microcosmic traffic flow, vessel domain, trajectory tracking, etc.

REFERENCES

[1] M. J. Osborne, J. Pepper, "Creating Marine Spatial Data Infrastructure for the UK". Edited D.R. Green, Springer Netherlands, Berlin, vol. 13, pp. 51-55, 2010.

[2] J. Pepper, "Unlocking the Marine Data Treasure Chest". Edited D.R. Green, Springer Netherlands, Berlin, vol. 13, pp. 57-64, 2010.

[3] S. Wahyu, G. H. Pramono, B. Purnawan, "Establishment of Marine and Coastal Spatial Data Infrastructure in Indonesia". Edited D.R. Green, Springer Netherlands, Berlin, vol. 13, pp. 97-103, 2010.

[4] F. Chen, H. Ren, "Comparison of vector data compression algorithms in Mobile GIS". International Conference on Computer Science and Information Technology, Chengdu, China, 2010.

[5] H. Ji, Y. Wang, "The Research on the Compression Algorithms for Vector Data". International Conference on Multimedia Technology, Ningbo, China, 2010.

[6] K. Chen, X. Zou, J. Xiong, H. Peng, "Fast corner detection algorithm with sub-pixel accuracy imitating Hessian-Laplace". Application Research of Computers, vol. 31, no. 7, pp. 2195-2200, 2014.

[7] W. Liu, L. Zhang, J. Sun, "Improved Method of Corner Detection Based on AP Clustering Image Block". Computer Engineering, vol. 41, no. 4, pp. 217-221, 2015.

[8] B. Sun, J. Tan, "Improved curvature scale space corner detection". Computer Engineering and Applications, vol. 50, no. 19, pp. 161-164, 2014.

[9] M. S. Verkeenko, "Development of an algorithm for fast corner points detection". Journal of Computer and Systems Sciences International, vol. 43, no. 4, pp. 392-401, 2014.

[10] C. Wang, "Corner detection algorithm based on random Hough transformation". Computer Engineering and Design, vol. 36, no. 3, pp. 716-720, 2014.

[11] X. Zhang, H. Wang, A. W. B. Smith, L. Xu, B. C. Lovell, D. Yang, "Corner detection based on gradient correlation matrices of planar curves". Pattern Recognition, vol. 43, no. 4, pp. 1207-1223, 2010.

[12] H. Wang, Z. Zhou, J. Cao, S. Yan, "Corner detection via multi-scale autocorrelation matrix on edge contours". Systems Engineering and Electronics, vol. 21, no. 9, pp. 1220-1224, 2014.

[13] M. Awrangjeb, G. Lu, C. S. Fraser, "Performance Comparisons of Contour-Based Corner Detectors". IEEE Transactions on Image Processing, vol. 21, no. 9, pp. 4167-4179, 2012.

[14] W. Zhang, F. Wang, L. Zhu, Z. Zhou, "Corner detection using Gabor filters". IET Image Processing, vol. 8, no. 11, pp. 639-646, 2014.

[15] H. Wang, "Study of fast adaptive harris corner detection method". Video Engineering, vol. 37, no. 19, pp. 208-212.

[16] M. Wang, Y. Dai, "Local Robust Feature Based on FAST Corner Detection". Transactions of Beijing Institute of Technology, vol. 33, no. 10, pp. 1045-1050, 2013.

[17] A. M. G. Pinheiro, M. Ghanbari, "Piecewise Approximation of Contours Through Scale-Space Selection of Dominant Points". IEEE Transactions on Image Processing, vol. 19, no. 6, pp. 1442-1450, 2010.

[18] A. Rosenfeld, "An Improved Method of Angle Detection on Digital Curves". IEEE Transactions on Computers, vol. 24, no. 9, pp. 940-941, 1975.

[19] Y. Xie, X. Han, Y. He, "Information Technology in Multi-object Tracking based on Bilateral Structure Tensor Corner Detection for Mutual Occlusion". Advanced Materials Research, vol. 977, pp. 502-506, 2014.

[20] C. Zhi, H. Tang, "Vehicle Speed Detection Method Based on Video Corner Feature Matching". Computer Engineering, vol. 39, no. 12, pp. 176-180, 2013.

[21] A. Masood, M. Sarfraz, "Corner detection by sliding rectangles along planar curves". Computers & Graphics, vol. 31, no. 3, pp. 440-448, 2007.

[22] X. Zhang, H. Wang, M. Hong, L. Xu, D. Yang, B. C. Lovell, "Robust image corner detection based on scale evolution difference of planar curves". vol. 30, no. 4, pp. 449-455, 2009.

[23] V. Pedrosa Glauco, A. Z. Barcelos Celia, "Anisotropic diffusion for effective shape corner point detection". Pattern Recognition Letters, vol. 31, no. 12, pp. 1658-1664, 2010.

[24] M. Farzin, M. Farahnaz, "Performance evaluation of corner detectors using consistency and accuracy measures". Computer Vision and Image Understanding, vol. 102, no. 1, pp. 81-94. , 2006

[25] S. Izchak, S. Assaf, K. Daniel, "A Geometric Approach to Monitoring Threshold Functions over Distributed Data Streams". Edited M. May and L. Saitta, Springer Berlin Heidelberg, Berlin, vol. 6202, pp. 163-186, 2010.