2016 Fourth International Workshop on Earth Observation and Remote Sensing Applications

Landsat Time Series Clustering under Modified Dynamic Time Warping

Yao Zhao
College of Information and Electrical Engineering, China Agricultural University
CIEE,CAU
Beijing,China
zhaoyaod@163.com

Lei Lin
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences
IRSDE,CAS
Beijing,China
linlei@radi.com

Wei Lu
College of Forestry, Northeast Forestry University
NEFU
Harbin,China
sanpangzi1228@126.com

Yu Meng
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences
IRSDE,CAS
Beijing,China
mengyu@radi.com

***Abstract*—Compared with the single remote sensing image, the time series images provide more information of ground objects, which can greatly improve the clustering accuracy. But time series clustering also has many difficulties, such as the impacts of cloud and other sharp noise. The sharp noise impacts time series clustering by affecting the calculation of the similarity measure of sequences.Therefore, this paper proposed a CD-DTW method (Canberra Distance-Dynamic Time Warping) based on the classic dynamic time warping distance and canberradistance. The CD-DTW method helps to construct a more reliable distance measure for remote sensing time series clustering. The CD-DTW is applied to Landsat time series clustering, and the overall performance of CD-DTW distance (overall accuracy 91.52%; kappa coefficient 0.89) was considerably better than that of classic DTW distance.**

Keywords—Dynamic Time Warping, Canberra Distance, Time Series Clustering，Remote Sensing

I. INTRODUCTION

Remote sensing image data, having macro and real-time characteristics, has been an important means of land cover detection. Earth's surface events occur and ground objects evolve both along with the advance of time. Therefore, the classification of remote sensing images can not only rely on the spatial characteristics of ground objects, but also should take appropriate method to analyze theirtemporal characteristics to extract the relevant information and knowledge. Clustering results with high accuracy could be obtained by mining abundant time information of Landsat time series in clustering process.

DTW (Dynamic Time Warping) algorithm, based on dynamic programming, is a typical optimization method. It uses the time warping function that satisfies certain conditions to describe the time corresponding relation between the input sample and the reference template and could solve corresponding time warping function to minimize the total distance of the two time series sequences. DTW was first developed from editing distance, and then Japanese scholars introduced DTW into the field of speech recognition to solve speech recognition problem in the 70's [1, 2].

At present, DTW algorithm has been successfully used to solve various practical application problems such as string search, handwritten character recognition, video retrieval, gesture recognition, network, circuit element, and traveling salesman problem [3-7]. DTW is also introduced into the remote sensing image time series processing for image clustering. Zhang Zheng [8] applied the DTW algorithm to the NDVI Modis time series clustering, and achieved better results than the ordinary Euclidean distance. Petitjean [9, 10] presented an approach to image time series analysis which is able to deal with irregularly sampled series and which also allows the comparison of pairs of time series where each element of the pair has a different number of samples. In the approach, K-Means algorithm was used for Image clustering based on DTW distance. On the basis of the previous work of Petitjean et al, we modified DTW algorithm aiming at the defects of DTW distance in Landsat image time series similarity comparison and verified the effectiveness of the modified DTW by evaluating the accuracy of land cover classification in the study area.

II. STUDY AREA AND DATA

A. Study area

Changping District Beijing, between N40°2′18″-40°23′13″ and E115°50′17″-116°29′49″, is located in the northwest of Beijing, covering an area of approximately

National Science and Technology Major Project (30-Y20A03-9003-15/16).
978-1-5090-1479-8/16/$31.00 ©2016 IEEE

1352 km^2 (Fig.1). Changping District Beijing belongs to a temperate zone with a continental monsoon climate.The climate in the study area has four distinct seasons with hot and humid summers and cold,windy, and dry winters.The average annual temperature is approximately 11.8℃ and the average annual precipitation is approximately 550.3 mm. Regional terrains from the northwest to the southeast gradually formed a gentle slope,and thin brown earth formed by weathering of rocks is the main soil type. The land cover typesare abundant, including forest, grass, cropland, construction land and water.

B. Landsat data and preprocess

Fourteen landsate8 images covering the Changping District of Beijing City were selected as the research data, shot on July 31, 2013, September 1, 2013, October 3, 2013,

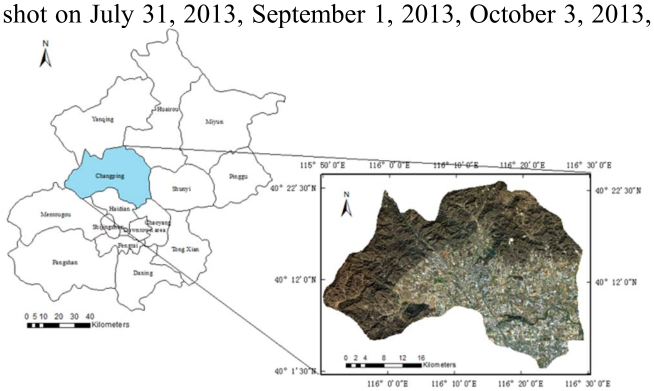

Fig. 1. The geographical region of the study area and the background information isthe true color image (R: red, G: green, B: blue) of Landsat8 data acquired on October3, 2013.

October 19, 2013, November 4, 2013, November 20, 2013, December 6, 2013, February 8, 2014, April 13, 2014, April 29, 2014, May 15, 2014, September 4, 2014, September 20, 2014, October 6, 2014 respectively.

The multi temporal Landsat8 data processing mainly included relative radiometric calibration, geometric precision correction, and image subset. NDVI of all fourteen temporal images are calculated by using the red and near infrared spectral bands of the images after the above correction and subset. In view of the fact that the green, red and near infrared bands of remote sensing images contain rich information of ground features, and NDVI can be used to well distinguished vegetation from other various objects, multivariate time series data for land cover classification consist of time series green bands, time series red bands, time series near infrared bands and time series NDVI.

C. Classification system and validating data

By computing the distance between the center sequences of the categories obtained by unsupervised clustering, merge similar categories, and the final classification image is generated. Based on the characteristics of land cover type inthe region, six classes were identified as the final class types ofthe regional land cover classification experiment, which includedcrop, water,forest, grass,impervious, and bare land.Sub categories contained in each of the major class types is shown in the TABLE1. These validating samples were easilyidentified on the G-F1 image and Google Earth map by visualobservation.

TABLEI. Land cover classification system

Class types	Sub categories
crop	The paddy field, irrigated land, dry land
water	Rivers, lakes, reservoirs, ponds,
forest	Forest, orchard, shrub woodland, other garden land
grassland	Natural pasture, artificial pasture, other grassland
impervious	Urban land, rural residentialland, transportation land and other construction sites
bare land	Other land, saline land, bare sand, swamp, beaches, and other land types not included in the above 5 class types

III. METHOD

A. The flowchart of clustering

A flowchart of Landsat time series clustering based on CD-DTW distance was presented in Fig. 2. The selected fourteen landsate8 images covering the Changping District of Beijing City were firstly preprocessed to have good quality. Then the preprocessed images were used to derive time series NDVI. Secondly, multivariate time series data, consisting of time series green bands, time series red bands, time series near infrared bands and time series NDVI, was prepared for land cover clustering. Thirdly, K-means clustering was conducted based on CD-DTW method. In the process of clustering, 1) based on multivariate time series data, Randomly select the clustering number and the clustering center sequence; 2) Calculate the CD-DTW distances between the other sequences and the clustering center sequence; 3) Cluster based on CD-DTW; 4) Judge whether the clustering center is convergent, if the answer is Yes, then end the clustering, otherwise Update the clustering center sequence by DBA [9]method and returned back to 2) until the answer is Yes. Finally, Classification images were obtained by post classifying the clustering results.

B. The classic dynamic time warping algorithm

DTW (Dynamic Time Warping) is a kind of nonlinear warping algorithm which combines time warping and distance measurement. DTW transforms a complex global optimization problem into a number of local optimization subproblems based on the dynamic programming principle, step by step for decision making. It uses a time warping function under the certain condition to describe the corresponding alignment between the input sequence and the reference sequence.

Suppose there are two sequences $A=(a_1,a_2,...,a_m)$, $B=(b_1,b_2,...,b_n)$, The DTW distance of A and B can be recursively computed by:

$$D(A_i,B_j)=\delta(a_i,b_j)+\min\begin{cases}D(A_{i-1},B_{j-1})\\D(A_{i-1},B_j)\\D(A_i,B_{j-1})\end{cases}$$

978-1-5090-1479-8/16/$31.00 ©2016 IEEE

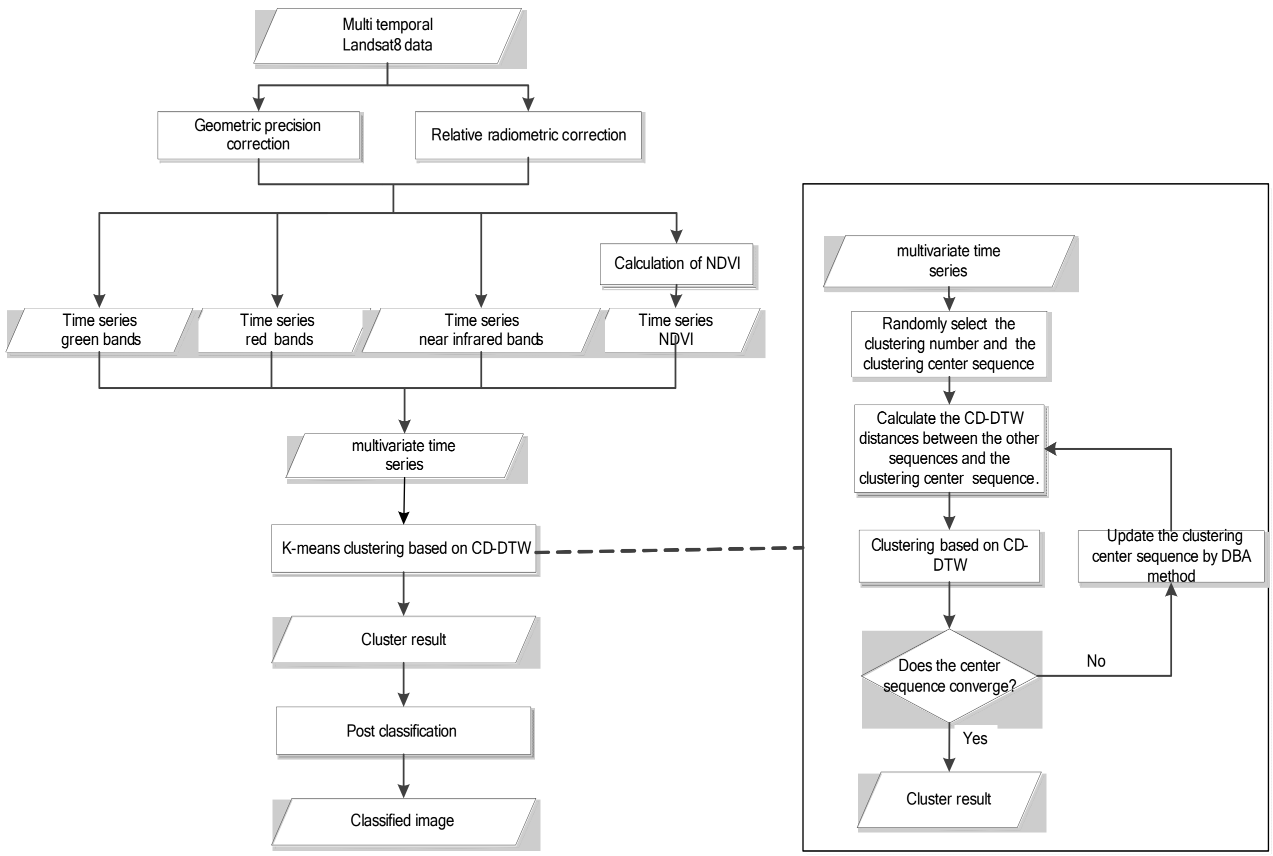

Fig. 2. The flowchart of Landsat Time Series Clustering based on CD-DTW distance

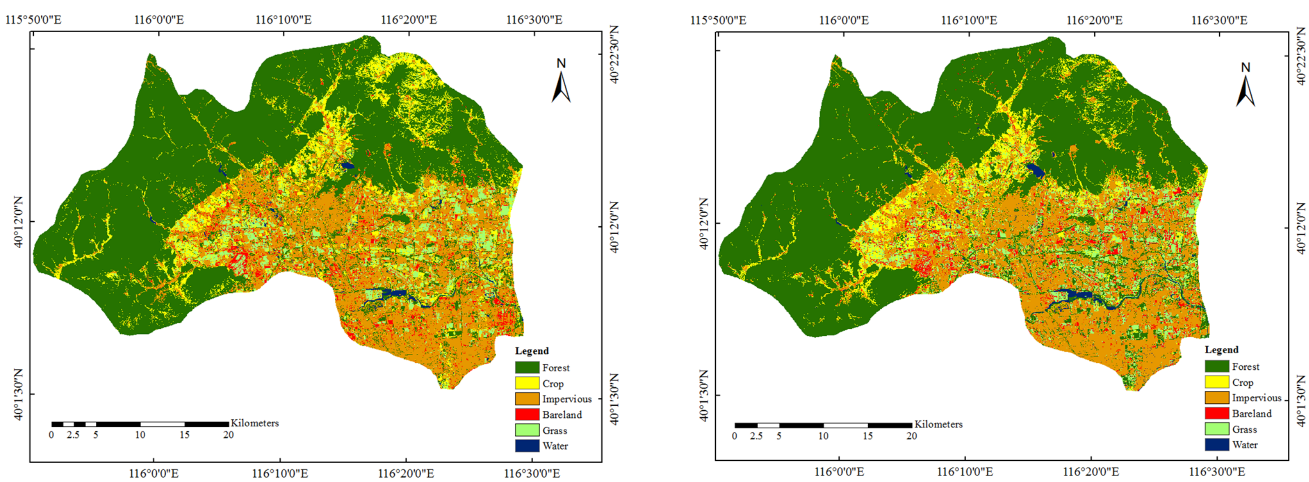

Fig. 3. Land cover classification results of Landsat time series clustering based on EUC-DTW(left) and CD-DTW(right)

978-1-5090-1479-8/16/$31.00 ©2016 IEEE

TABLEII. Confusion matrixes for land cover classification based on EUC-DTW.

Classes	Ground Truth (Pixels)							User accuracy (%)
	Forest	Crop	Impervious	Bare land	Grass	Water	Total	
Forest	3898	120	154	0	8	485	4665	83.56
Crop	132	2102	19	1	130	0	2384	88.17
Impervious	37	42	5750	73	56	3	5961	96.46
Bare land	0	0	227	657	5	0	889	73.9
Grass	0	215	31	451	1112	0	1809	61.47
Water	0	0	15	1	0	968	984	98.37
Total	4067	2479	6196	1183	1311	1456	16692	
Producer accuracy(%)	95.84	84.79	92.8	55.54	84.82	66.48		

TABLEIII. Confusion matrixes for land cover classification based on CD-DTW.

Classes	Ground Truth (Pixels)							User accuracy (%)
	Forest	Crop	Impervious	Bare land	Grass	Water	Total	
Forest	4005	323	91	2	8	31	4460	89.8
Crop	31	2072	38	0	185	0	2326	89.08
Impervious	26	4	5928	310	3	1	6272	94.52
Bare land	0	7	75	843	111	0	1036	81.37
Grass	5	73	59	28	1004	0	1169	85.89
Water	0	0	5	0	0	1424	1429	99.65
Total	4067	2479	6196	1183	1311	1456	16692	
Producer accuracy (%)	98.48	83.58	95.67	71.26	76.58	97.8		

$$\delta(a_i,b_j)=\sqrt[p]{\sum_{k=1}^{N}\left\|a_{ik}-b_{jk}\right\|_p}$$

Where $D(A_i,B_j)$ is the DTW distance between the data point at time i of the sequence A and data point at time j of the sequence B and D is a $m\times n$ matrix. When the time sequence to be processed has multiple features, a_i and b_j are both multi-dimensional vectors, and N is the number of features. $\|\bullet\|_p$ is the norm of p, it stands for Manhattan $(p=1)$ and Euclidean distance $(p=2)$. When $p=2$, we call it EUC-DTW.

Starting from the lower left corner of the warping matrix and walking step-by-step to the upper right corner with only stepping forward and never backward, passing each element of the warping matrix denotes a single time warp (alignment) between the data point at time i of the sequence A and data point at time j of the sequence B, thus we get the whole warping path. We use the warping path to illustrate the alignment process of the two time series sequences. The DTW distance of the last data point $D(A_m,B_n)$, which is affected by the global data point, appears thus to be a measure of the similarity between the two sequences.

C. Canberra Distance –Dynamic Time Warping

Because of the complex of the actual object, the uncertain factors in the imaging process, and the instability of the long time observation, a wide variety of sharp noises often exist in time series images obtained finally. If distance affected by the dimension is calculated to weigh similarity measure of time series sequences, these sharp noises tend to have a larger difference, which is easy to cause the sample sequence to be attributed to the errors due to these occasional distortion points.

In this study, Dynamic Time Warping algorithm combined with Canberra distance is used as the similarity measure of time series sequences of different pixels, which is called CD-DTW distance. Canberra distance is a deformation of Manhattan distance, which can be regarded as the weighted Manhattan distance. Canberra distance is not affected by the dimension, and it can eliminate the effect of the sharp noise on the similarity measure of time series sequences. Therefore, we make full use of the advantages of Canberra distance and the DTW algorithm, and apply the similarity measure to the land cover classification of time series remote sensing images.

The CD-DTW distance of two sequences is calculated as follows：

$$D(A_i,B_j)=\delta(a_i,b_j)+\min\begin{cases}D(A_{i-1},B_{j-1})\\D(A_{i-1},B_j)\\D(A_i,B_{j-1})\end{cases}$$

978-1-5090-1479-8/16/$31.00 ©2016 IEEE

$$\delta(a_i, b_j) = \sum_{k}^{n} \frac{|a_{ik} - b_{jk}|}{|a_{ik}| + |b_{jk}|}$$

Where $|a_{ik}|$ is the absolute value of the data point at time i of the feature K of the sequence A

IV. RESULT

The land cover classification results obtained by post classifying the Landsat Time Series Clustering results based on EUC-DTW and CD-DTW are shown in Figure 3 In order to ensure comparability of the accuracies, the classification accuracies and kappa statistics were estimated based on the same validation samples. The confusion matrix of the Landsat time series classification based on EUC-DTW distance and CD-DTW distance are shown in TABLE2 and TABLE3, respectively. The overall classification accuracy based on EUC-DTW distance was 86.79%, with kappa coefficient of 0.83. The overall performance ofCD-DTW distance (overall accuracy 91.52%; kappa coefficient 0.89) was considerably better than that of EUC-DTW distance. Compared to EUC-DTW, the user accuracy of the grass class based on CD-DTW was improved significantly to 85.89%, but producer accuracy reduced from 84.82% to76.58%. There is a substantial increase for the user accuracy and producer accuracy of bare land class. That bare land was omitted to impervious and grass was commissioned to bare land was the main reason for the low accuracy of the bare land. That grass was omitted to crop and bare land was commissioned to grass was the main reason for the low accuracy of the grass. Other class types all had a higher user and producer accuracies, better separation with other classes.

V. CONCLUSION

In this paper, an effective time series clustering framework was proposed with application to clusterLandsat8 time series. CD-DTW distance was chosen as the similarity measure since it can eliminate the effect of the sharp noises of time series sequences. We compare the EUC-DTW and CD-DTW in the land cover classification of Landsat8 time series and the CD-DTW algorithm performed better. In the future, CD-DTW can be tested for more data in other domains and compared to more variants of DTW.

ACKNOWLEGMENTS

This work was supported by National Science and Technology Major Project (30-Y20A03-9003-15/16)

REFERENCES

[1] H. Sakoe, S. Chiba. A dynamic programming approach to continuous speech recognition [C]//Proceedings of the seventh international congress on acoustics. 1971, 3: 65-69.

[2] H. Sakoe, S. Chiba. Dynamic programming algorithm optimization for spoken word recognition[J]. Acoustics, Speech and Signal Processing, IEEE Transactions on, 1978, 26(1): 43-49

[3] J.B. Kruskal, M. Liberman. The symmetric time-warping problem: from continuous to discrete[J]. Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison, 1983: 125-161.

[4] Aach J, Church G M. Aligning gene expression time series with time warping algorithms[J]. Bioinformatics, 2001, 17(6): 495-508.

[5] Bar-Joseph Z, Gerber G, Gifford D K, et al. A new approach to analyzing gene expression time series data[C]//Proceedings of the sixth annual international conference on Computational biology. ACM, 2002: 39-48.

[6] Gavrila D M, Davis L S. Towards 3-d model-based tracking and recognition of human movement: a multi-view approach[C]// International workshop on automatic face-and gesture-recognition. 1995: 272-277.

[7] Rath T M, Manmatha R. Word image matching using dynamic time warping[C]//Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on. IEEE, 2003, 2: II-521-II-527 vol. 2.

[8] Zhang Z, Tang P, Huo L, et al. MODIS NDVI time series clustering under dynamic time warping[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2014, 12(05): 1461011.

[9] Petitjean F, Ketterlin A, Gançarski P. A global averaging method for dynamic time warping, with applications to clustering[J]. Pattern Recognition, 2011, 44(3): 678-693.

[10] Petitjean F, Inglada J, Gançarski P. Satellite image time series analysis under time warping[J]. Geoscience and Remote Sensing, IEEE Transactions on, 2012, 50(8): 3081-3095.

978-1-5090-1479-8/16/$31.00 ©2016 IEEE