Vision Trajectory Database

Tiger Hou
Peking University
Beijing, China
houtiger@pku.edu.cn

Sheng Wang
New York University
NY, USA
swang@nyu.edu

ABSTRACT

This paper is a survey on visual trajectory, which is defined as trajectory extracted from visual objects like video and image. We will introduce the application scenes of visual trajectory and point out the limitation of previous works. And then, we will explain why it's essential to build a trajectory database for these tasks and show some new tasks can be supplied by such a database. Challenges of normalization and similarity measurement would be discussed. Available datasets would also be presented in the end.

CCS CONCEPTS

• **Information systems → Database design and models**.

KEYWORDS

database management, handwriting trajectory, drawing trajectory

ACM Reference Format:
Tiger Hou and Sheng Wang. 2019. Vision Trajectory Database. In *Summer '19: A Summer Research Internship on Database, Jun. 11–Sept. 10, 2019, Singapore, SG.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Nowadays, many works focus on vehicle trajectory management have been proposed. However, for another series of trajectory, like handwriting trajectory and drawing trajectory, which we call *visual trajectory*, they haven't been fully studied. There are four main differences between visual and vehicle trajectory: 1) Freely distributed, unlike vehicles are constrained to road network, handwriting and drawing trajectory exists everywhere on the 2D ground. So the traditional map matching method [21] cannot be applied to this new problem; 2) Accurately recorded, vehicles trajectory recording relies on unstable GPS accuracy, but for visual trajectory, it is recorded by human-machine interaction equipment accurately [19], which makes it easier for similarity search; 3) Same simple rates, due to the same pen-based recording devices; 4) Dense and frequently crossing, visual trajectory extracted from image or video is restrained to the fixed image size, and there are a lot of crossing between trajectories, which means we can not treat it as a set matching problem in similarity search but have to consider temporal order information.

So, we would like to propose a visual trajectory database which supports both whole trajectory and sub-trajectory Top-k similarity search, with a balanced efficiency and effectiveness, i.e. the high similarity search accuracy and retrieval speed.

The possible application scenes include:

1) Handwriting recognition and user identification. As a classic problem, handwriting user identification and recognition has been well studied in computer vision area [20, 23]. While in handwriting user identification, lack of training data has always been a tough issue for Convolutional Neural Network (CNN) training. But if a handwriting trajectory search database was built, it would solve this problem easily because it can work well even in one-shot condition, which means we only have gotten one sample for each handwritten character of a certain user. For character recognition, the idea that visually similar handwriting characters can be distinguished by writing trajectories has been commonly accepted, as Geoffrey E. Hinton mentioned in [13]. And trajectory reconstruction methods based on combination of CNN and RNN or Generative Adversarial Network (GAN) have been mature enough [2, 6]. So when it comes to a messy classification problem that two handwritten characters are visually similar, our database can provide essential help in trajectory similarity search for recognition, on condition that trajectories have been reconstructed.

2) Drawing recommendation. Sketch-based image retrieval has been a hot topic for a long time [17, 24]. It is more convenient for users to describe fine-grained features of their target object in drawings than natural language. Because it could be a long sentence to describe the object explicitly and still unable to get a satisfactory result. Besides, the current vector embedding based search method doesn't support sub-trajectory search, because it's hard to map whole trajectory and sub-trajectory into the same space while ensure high similarity between them. Sub-trajectory search is an important kind of query because sometimes people do not have an explicit description of their target in mind, only a vague description of size, pose, shape instead. That's where sub-trajectory search may help. People could draw a draft with several strokes, and based on the recommended drawings to decide how their final targets exactly look like. Besides, sketch-based image can also be a new kind of emoji used in instant messaging tools, e.g. Imessage Digital Touch on iphone.

3) Sports play retrieval. Sports play retrieval by trajectory searching have been studied in [18, 22]. But the current solution is usually embedding trajectories into vectors in a space, and measure the distance through cosine similarity. As we have already mentioned previously, the distance between embedded sub-trajectory and whole trajectory is large, which means the previous methods

Summer '19, Jun. 11–Sept. 10, 2019, Singapore, SG

ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

do not support sub-trajectory search, either. And our database can also be applied to address this problem.
4) Human-computer interaction. At the moment, a lot of human computer interaction methods are based on finger moving trajectory similarity search [19]. But the searching is too time consuming, an enumerate DTW-based trajectory similarity search in a dataset with 160 trajectories would take roughly 10 seconds, which is so time-consuming. A visual trajectory database support efficient retrieval would surely be useful in this task.

2 RELATED WORK

Trajectory Reconstruction from Videos and Images. Reconstructing trajectory from image and video has been a fully studied problem in computer vision area, like handwriting character trajectory, drawing trajectory and pedestrian motion. Usually, a combination of Convolution Neural Network (CNN) and Recursive Neural Network (RNN) is preferred for image trajectory reconstruction [2]. And there are also works focus on feature designing and computing [10]. When it comes to video trajectory, Generative Adversarial Network (GAN) is applied in these tasks [6].
Trajectory Similarity Measures. In measuring trajectory similarity, Dynamic Time Warping (DTW) is a well accepted method. Different from Euclidean Distance, DTW consider matching points within a certain time window, that's also how the name "warping" comes. But the expense of computation is very large, and several optimizations are proposed in [4, 9, 16], including lower bound pruning, early abandoning Z-Normalization, reordering early abandoning, reversing the Query/Document role in LB_{keogh} and cascading lower bounds. Besides, Longest Common Subsequence (LCSS) and Edit Distance on Real Sequence (EDRS) are also widely accepted trajectory similarity measures. Even though, the approximate computing complexity of current popular measures is still $O(mn)$, where m and n represent the length of query and candidate sequences separately.

3 CHALLENGE

3.1 Preprocessing of Trajectories

We separate the preprocessing of trajectory into two problems: resizing and normalization. For example, the length of trajectory extracted from a 1000×1000 image and a 25×25 image would be definitely different. And the coordinates of points that consist the trajectory would also be changing if the handwritten character is moved from lower right corner to upper left corner of the image. Image resizing has been a fully studied problem in computer vision area [3]. And resizing character into our desired size is an easy job. And for image normalization, many methods have been proposed for medical image normalization [15], which is also applicable to our problem.

3.2 Supporting Sub-trajectory Search

As we are supporting sub-trajectory search, then the vector embedding method is not applicable. The distance between whole trajectory and sub-trajectory in the mapped space is very large as the vector only posses general feature, while neglecting local feature at the same time. So we have to implement the sub-trajectory search based on the time sequence similarity measurements like DTW or curve similarity measurements like Hausdorff distance and Fréchet distance [1, 14].

Table 1: Statistics of CASIA Database

dataset	writers	character samples		
		total	symbol	Chinese/class
OLHWDB1.0	420	1,694,741	71,806	1,622,935/3,866
OLHWDB1.1	300	1,174,364	51,232	1,123,132/3,755
OLHWDB1.2	300	1,042,912	51,181	991,731/3,319
total	1,020	3,912,017	174,219	3,737,798/7,185

3.3 Similarity Measures

The similarity measurement determines the efficiency of our database directly. Although measures like Dynamic TIme Warping (DTW), Longest Common Subsequence (LCSS), and Edit Distance on Real sequence (EDR) are robust to local time shifting and noise, but they all require a lot of computations with $O(n^2)$ complexity. In our task, the size of trajectory dataset usually ranges between no less than 1 million and 50 million, which means we have to figure out how to reduce the computations and in the mean while, ensure accuracy.

4 DATASETS

There are datasets of vision trajectories of considerable size. For handwriting characters, the CASIA database of Chinese character handwriting trajectories [11] consists of over 1.6 million trajectories of more than 3000 classes. And trajectory databases of other languages are also available, like English [12], Japanese [8] and Vietnamese [7], etc. We show the statistics of CASIA database in **Table 1**.

Besides, a drawing trajectory dataset [5] is provided by Google, which contains 50 million drawings across 345 categories. Another drawing trajectory dataset is provided by Georgia Institute of Technology [17] with over 7000 trajectories. In the Georgia Tech dataset, each sketch drawing is labeled with an original image, from which it was generated.

REFERENCES

[1] Pankaj K. Agarwal, Rinat Ben Avraham, Haim Kaplan, and Micha Sharir. 2014. Computing the Discrete Fréchet Distance in Subquadratic Time. *SIAM J. Comput.* 43, 2 (2014), 429–449.
[2] Ayan Kumar Bhunia, Abir Bhowmick, Ankan Kumar Bhunia, Aishik Konwer, Prithaj Banerjee, Partha Pratim Roy, and Umapada Pal. 2018. Handwriting Trajectory Recovery using End-to-End Deep Encoder-Decoder Network. In *24th International Conference on Pattern Recognition, ICPR 2018*. 3639–3644.
[3] Che-Han Chang and Yung-Yu Chuang. 2012. A line-structure-preserving approach to image resizing. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 1075–1082.
[4] Ada Wai-Chee Fu, Eamonn J. Keogh, Leo Yung Hang Lau, Chotirat (Ann) Ratanamahatana, and Raymond Chi-Wing Wong. 2008. Scaling and time warping in time series querying. *VLDB J.* 17, 4 (2008), 899–921.
[5] Google. 2017. The Quick, Draw! Dataset. Retrieved Sep. 3, 2019 from https://github.com/googlecreativelab/quickdraw-dataset
[6] Agrim Gupta, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. 2018. Social GAN: Socially Acceptable Trajectories With Generative Adversarial Networks. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018*. 2255–2264.
[7] Nguyen Tuan Hung. 2016. TUAT Nakagawa Lab. HANDS - Vietnamese Online Handwriting Database (HANDS-VNOnDB). Retrieved August 26, 2019 from http://tc11.cvc.uab.es/datasets/HANDS-VNOnDB_1

[8] Stefan Jäger and Masaki Nakagawa. 2001. Two On-Line Japanese Character Databases in Unipen Format. In *6th International Conference on Document Analysis and Recognition, ICDAR 2001*. 566–571.

[9] Eamonn J. Keogh, Li Wei, Xiaopeng Xi, Michail Vlachos, Sang-Hee Lee, and Pavlos Protopapas. 2009. Supporting exact indexing of arbitrarily rotated shapes and periodic time series under Euclidean and warping distance measures. *VLDB J.* 18, 3 (2009), 611–630.

[10] Laura Leal-Taixé, Michele Fenzi, Alina Kuznetsova, Bodo Rosenhahn, and Silvio Savarese. 2014. Learning an Image-Based Motion Context for Multiple People Tracking. In *2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014*. 3542–3549.

[11] Cheng-Lin Liu, Fei Yin, Da-Han Wang, and Qiu-Feng Wang. 2011. CASIA Online and Offline Chinese Handwriting Databases. In *2011 International Conference on Document Analysis and Recognition, ICDAR 2011*. 37–41.

[12] David Llorens, Federico Prat, Andrés Marzal, Juan Miguel Vilar, María José Castro, Juan-Carlos Amengual, Sergio Barrachina, Antonio Castellanos, Salvador España Boquera, J. A. Gómez, Jorge Gorbe-Moya, Albert Gordo, Vicente Palazón, Guillermo Peris, Rafael Ramos-Garijo, and Francisco Zamora-Martínez. 2008. The UJIpenchars Database: a Pen-Based Database of Isolated Handwritten Characters. In *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2008*.

[13] Vinod Nair and Geoffrey E. Hinton. 2006. Inferring Motor Programs from Images of Handwritten Digits. In *Advances in Neural Information Processing Systems 18, NIPS 2006*. 515–522.

[14] Sarana Nutanong, Edwin H. Jacox, and Hanan Samet. 2011. An Incremental Hausdorff Distance Calculation Algorithm. *PVLDB* 4, 8 (2011), 506–517.

[15] John A. Onofrey, Dana I. Casetti-Dinescu, Andreas D. Lauritzen, Saradwata Sarkar, Rajesh Venkataraman, Richard E. Fan, Geoffrey A. Sonn, Preston C. Sprenkle, Lawrence H. Staib, and Xenophon Papademetris. 2019. Generalizable Multi-Site Training and Testing Of Deep Neural Networks Using Image Normalization. In *16th IEEE International Symposium on Biomedical Imaging, ISBI 2019*. 348–351.

[16] Thanawin Rakthanmanon, Bilson J. L. Campana, Abdullah Mueen, Gustavo E. A. P. A. Batista, M. Brandon Westover, Qiang Zhu, Jesin Zakaria, and Eamonn J. Keogh. 2012. Searching and mining trillions of time series subsequences under dynamic time warping. In *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2012*. 262–270.

[17] Patsorn Sangkloy, Nathan Burnell, Cusuh Ham, and James Hays. 2016. The sketchy database: learning to retrieve badly drawn bunnies. *ACM Trans. Graph.* 35, 4 (2016), 119:1–119:12.

[18] Long Sha, Patrick Lucey, Yisong Yue, Peter Carr, Charlie Rohlf, and Iain A. Matthews. 2016. Chalkboarding: A New Spatiotemporal Query Paradigm for Sports Play Retrieval. In *Proceedings of the 21st International Conference on Intelligent User Interfaces, IUI 2016*. 336–347.

[19] Sharad Vikram, Lei Li, and Stuart J. Russell. 2013. Writing and sketching in the air, recognizing and controlling on the fly. In *2013 ACM SIGCHI Conference on Human Factors in Computing Systems, CHI 2013*. 1179–1184.

[20] Qingqing Wang and Yue Lu. 2017. A Sequence Labeling Convolutional Network and Its Application to Handwritten String Recognition. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017*. 2950–2956.

[21] Sheng Wang, Zhifeng Bao, J. Shane Culpepper, Zizhe Xie, Qizhi Liu, and Xiaolin Qin. 2018. Torch: A search engine for trajectory data. In *41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2018*. 535–544.

[22] Zheng Wang, Cheng Long, Gao Cong, and Ce Ju. 2019. Effective and Efficient Sports Play Retrieval with Deep Representation Learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019*. 499–509.

[23] Curtis Wigington, Chris Tensmeyer, Brian L. Davis, William A. Barrett, Brian L. Price, and Scott Cohen. 2018. Start, Follow, Read: End-to-End Full-Page Handwriting Recognition. In *Computer Vision - ECCV 2018 - 15th European Conference*. 372–388.

[24] Qian Yu, Feng Liu, Yi-Zhe Song, Tao Xiang, Timothy M. Hospedales, and Chen Change Loy. 2016. Sketch Me That Shoe. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*. 799–807.