Effective Trajectory Similarity Measure for Moving Objects in Real-world Scene

Moonsoo Ra, Chiawei Lim, Yong Ho Song, Jechang Jung, and Whoi-Yul Kim

Hanyang University, Electronics and Computer Engineering,
222 Wangsimni-ro, Seongdong-gu, Seoul, 133-791, Korea
{msna,chiawei}@vision.hanyang.ac.kr,{yhsong,jjeong,wykim}@hanyang.ac.kr
http://hanyang.ac.kr

Abstract. Trajectories of moving objects provide fruitful information for analyzing activities of the moving objects; therefore, numerous researches have tried to obtain semantic information from the trajectories by using clustering algorithms. In order to cluster the trajectories, similarity measure of the trajectories should be defined first. Most of existing methods have utilized dynamic programming (DP) based similarity measures to cope with different lengths of trajectories. However, DP based similarity measures do not have enough discriminative power to properly cluster trajectories from the real-world environment. In this paper, an effective trajectory similarity measure is proposed, and the proposed measure is based on the geographic and semantic similarities which have a same scale. Therefore, importance of the geographic and semantic information can be easily controlled by a weighted sum of the two similarities. Through experiments on a challenging real-world dataset, the the proposed measure was proved to have a better discriminative power than the existing method.

Keywords: Video surveillance, trajectory clustering, moving objects

1 Introduction

Trajectories of moving objects are frequently used metadata to analyze behaviors of the moving objects and there are several trajectory clustering methods [1–4] have been introduced in recent years. Most of trajectory clustering methods utilize dynamic programming (DP) based similarity measures to cope with different lengths of the trajectories. However, in order to properly cluster trajectories, using only DP based similarity measures is not desirable, since trajectories acquired from real-world scene have a large shape variation, and may have missing data or noises from inaccurate measurement. Therefore, Liu and Schneider [2] developed a trajectory similarity which combines a geographic similarity with a semantic similarity. As the geographic similarity, [2] used a center of mass for the trajectory and a displacement vector which represents an approximated direction of the trajectory. In addition, as a semantic similarity, Longest Common Subsequence (LCSS) algorithm [5] is used to penalize a similarity between trajectories which have different shapes.

K.J. Kim (ed.), *Information Science and Applications,*
Lecture Notes in Electrical Engineering 339, DOI 10.1007/978-3-662-46578-3_75

However, the similarity measure introduced in [2] is not well applied on the trajectories acquired from a real world scene because, LCSS algorithm has a limited ability to capture semantic relationship between the trajectories. Furthermore, geographic and semantic similarities cannot equally contribute to a total similarity, since the total similarity is defined as a ratio of geographic and semantic similarities. In order to overcome the limitations, this paper proposed an effective trajectory similarity measure. The proposed measure is similar to the existing measure, in terms of using the concept of geographic and semantic similarities; however, it utilizes the starting point and angle difference of the displacement vector to capture the geographic relationship. Furthermore, Hausdorff distance [6] is applied on the normalized trajectory to capture the semantic relationship. One big advantage of the proposed measure is that both geographic and semantic similarities have a same scale; thus, both similarities can equally contribute to the total similarity. Through challenging experiments, the proposed measure is proved to have improved performance on the moving object trajectories acquired from the real-world scene.

The rest of the paper is organized as follows. In Section 2, trajectory obtaining process is explained in detail. The proposed trajectory similarity measure is presented in Section 3. An improved performance of the proposed measure is evaluated using the real world trajectories in Section 4. Finally, this paper is concluded in Section 5.

2 Trajectory Obtaining Process

Trajectory obtaining process consists of three stages: moving object detection, moving object association, and pruning. In the moving object detection stage, Gaussian Mixture Model (GMM) is used to separate foreground and background from an input surveillance video; then labeling algorithm is utilized to obtain blobs of the moving objects. As an implementation of GMM, algorithm introduced in [7] is used and as a labeling algorithm, simple grassfire algorithm [8] is utilized.

In order to associate the moving objects, similarity measure between moving objects should be defined first. Assume that i^{th} moving object in the previous frame is denoted as O^i_{prev} and j^{th} moving objects in the current frame are denoted as O^j_{curr}. Then, similarity S_{ij} between O^i_{prev} and O^j_{curr} are defined through following equations:

$$S_{ij} = s^{HSV}_{ij} \times exp(-dist(\mathbf{p}^i_{prev}, \mathbf{p}^j_{curr})/\lambda), \tag{1}$$

$$s^{HSV}_{ij} = \sum \min(H^i_{prev}, H^j_{curr}) / \sum \max(H^i_{prev}, H^j_{curr}), \tag{2}$$

where $\mathbf{p}^i_{prev}$ and $\mathbf{p}^j_{curr}$ are center of masses for O^i_{prev} and O^j_{curr}; then, H^i_{prev} and H^j_{curr} are HSV color histograms of O^i_{prev} and O^j_{curr}, respectively. Using S_{ij} for all i and j, bipartite graph B is constructed to associate moving objects from previous and current frames. In order to solve the bipartite graph association

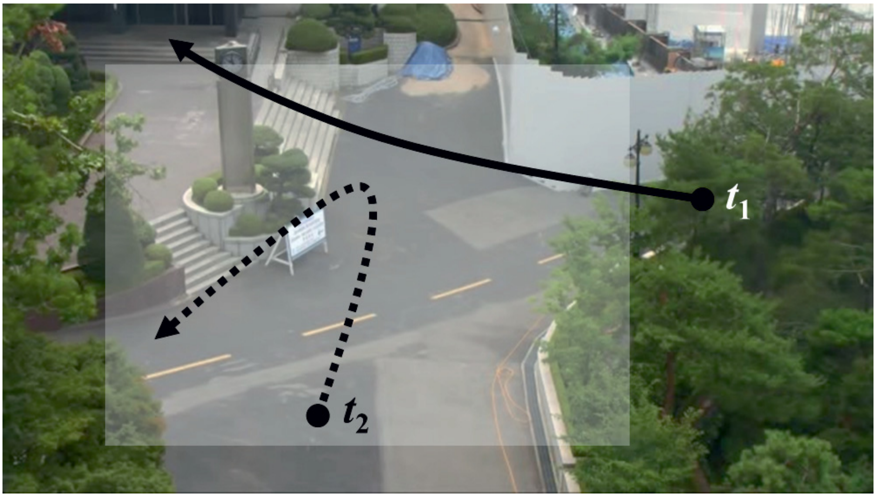

Fig. 1. Example of the second pruning condition. Assume that t_1 and t_2 are obtained from the trajectory obtaining process. However, dotted trajectory, t_2, is not started or ended at boundaries of the image (not shaded region). Trajectories like t_2 are removed from the set of trajectories during the pruning process.

problem, traditional Hungarian algorithm [9] is applied on B. As a result of moving object association stage, set of moving object trajectories $T = \{t_1, t_2, ..., t_N\}$ are obtained.

Finally, inappropriately segmented or associated moving objects are pruned to improve quality of T. In order to determine pruned moving objects, following two conditions are used. First condition is that moving objects should not be disappeared at least 20 frames. Second condition is that moving objects should be started and ended at boundaries of the image. Through the pruning stage, subset of moving object trajectories $T_{sub} = \{t_1, t_2, ..., t_K\}$ is obtained. In order to give the readers better understanding of the pruning conditions, example of the second pruning condition is depicted in Fig. 1.

3 Trajectory Similarity Measure

In this paper, the trajectory $t_i \in T_{sub}$ of a moving object is defined as a sequence of 2D points, $t_i = \{(x_1^i, y_1^i), (x_2^i, y_2^i), ..., (x_L^i, y_L^i)\}$, where L is a length of the trajectory. The proposed similarity measure is defined on the two trajectories, t_i and t_j, and the proposed measure consists of two distinctive similarities: geographic similarity and semantic similarity. Geographic similarity captures spatial adjacency of the trajectories and semantic similarity captures shape difference of the trajectories.

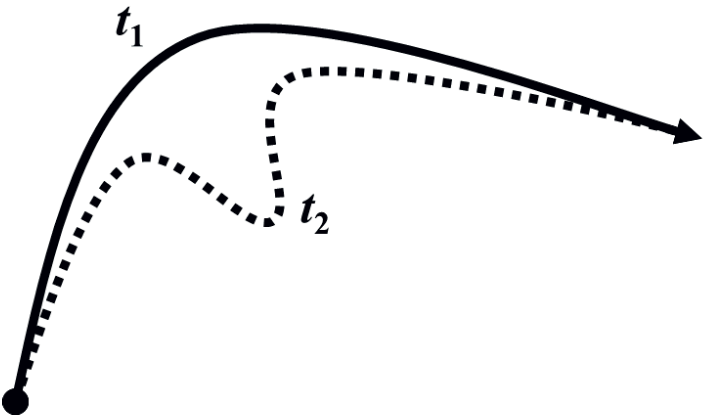

Fig. 2. Example of the trajectories which cannot distinguish by the geographic similarity, $s_{geo}(t_1, t_2)$, since start and end points of solid trajectory and dotted trajectory are same. In order to overcome such problem, semantic similarity will be introduced.

3.1 Geographic Similarity

In this section, geographic similarity is defined to satisfy following two properties: the similarity has a higher value when start points between the trajectories are spatially adjacent and approximated directions of the trajectories are similar. Spatial similarity of the start points is calculated as a traditional Euclidean distance, $d(\mathbf{s}_i, \mathbf{s}_j) = \|\mathbf{s}_i - \mathbf{s}_j\|$, where $\mathbf{s}_i$ and $\mathbf{s}_j$ are start points of t_i and t_j, respectively. The approximated direction is defined as a displacement vector, $\mathbf{d} = \mathbf{e} - \mathbf{s}$, where $\mathbf{e}$ is an end point of the trajectory; then similarity $s_{disp}(\mathbf{d}_i, \mathbf{d}_j)$ between $\mathbf{d}_i$ and $\mathbf{d}_j$ is calculated as an angle difference of $\mathbf{d}_i$ and $\mathbf{d}_j$. Using $d(\mathbf{s}_i, \mathbf{s}_j)$ and $s_{disp}(\mathbf{d}_i, \mathbf{d}_j)$, the proposed geographic similarity s_{geo} is defined as:

$$s_{geo}(t_i, t_j) = d(\mathbf{s}_i, \mathbf{s}_j) + s_{disp}(\mathbf{d}_i, \mathbf{d}_j), \tag{3}$$

$$s_{disp}(\mathbf{d}_i, \mathbf{d}_j) = \max(\|\mathbf{d}_i\| + \|\mathbf{d}_j\|) \exp(\lambda|\theta_i - \theta_j|/\pi), \tag{4}$$

where θ_i and θ_j are angles of $\mathbf{d}_i$ and $\mathbf{d}_j$, respectively, and λ is a parameter for controlling a decreasing rate of the exponential function, and $\|\mathbf{d}\|$ indicates a magnitude of the displacement vector $\mathbf{d}$.

As denoted in equation (3) and (4), geographic similarity does not consider shape of the trajectories; therefore, it cannot distinguish trajectories which have same start and end points but different trajectory shapes as illustrated in Fig. 2.

3.2 Semantic Similarity

The proposed semantic similarity is designed to have a higher value when shapes of the trajectories are similar. As a semantic similarity $s_{sem}(t_i, t_j)$, Hausdorff

Fig. 3. Sample images from the challenging real-world dataset. Left image is a sample image from video A, and right image is a sample image from video B. Camera view of the left image is including a parking lot; therefore, vehicles are appearing frequently. While, camera view of the right image is targeted on the entrance of a subway station; therefore, a lot of people can be observed even in short duration of time.

distance is adopted, since it is known as having a good performance in comparing shapes of objects [6, 10, 11]. Hausdorff distance of two trajectories, $d_H(t_i, t_j)$, is defined as

$$s_{sem}(t_i, t_j) = d_H(t_i, t_j) = \max(h(t_i, t_j), h(t_j, t_i)), \tag{5}$$

$$h(t_i, t_j) = \max_{\mathbf{u} \in t_i} \min_{\mathbf{v} \in t_j} \|a - b\|, \tag{6}$$

where $\mathbf{u}$ and $\mathbf{v}$ are 2D points belong to t_i and t_j, respectively.

Since Hausdorff distance is calculated based on the Euclidean distance as denoted in equation (6), s_{sem} and s_{geo} have a same scale; thus, contributions of each similarity can be easily controlled by weighted sum of $s_{sem}(t_i, t_j)$ and $s_{geo}(t_i, t_j)$.

3.3 Proposed Similarity Measure

The proposed similarity measure s_{total} is defined as a weighted sum of $s_{geo}(t_i, t_j)$ and $s_{sem}(t_i, t_j)$ as following:

$$s_{total}(t_i, t_j) = \alpha s_{geo}(t_i, t_j) + (1 - \alpha) s_{sem}(t_i, t_j), \tag{7}$$

where α is a parameter for controlling an importance between geographic and semantic similarities. In this paper, α is empirically set as 0.4.

In order to cluster moving object trajectories by using the proposed measure, Affinity propagation algorithm [12] is utilized. Advantage of using Affinity propagation is that it automatically selects a number of clusters; therefore, parameter optimization process is not necessary throughout the experiments.

4 Experimental Result

In this section, performance of the proposed similarity measure was evaluated by using a challenging real-world dataset. The dataset consists of video A and

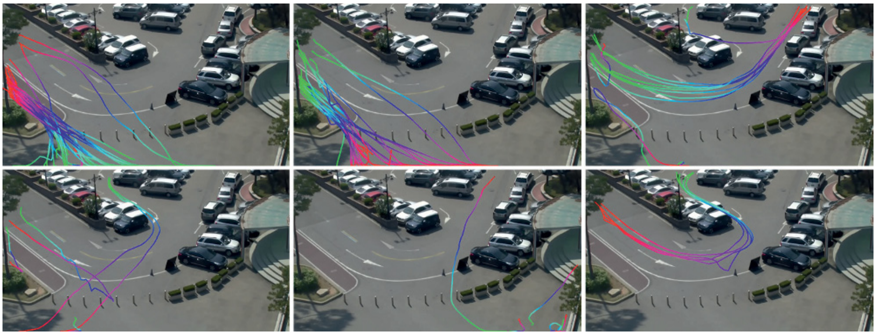

Fig. 4. Result of the clustering algorithm with the existing measure [2]. Video A is used for the source of the trajectories and total six clusters were acquired. As you can see in the top-left or bottom-left images, obviously different trajectories are considered as a single cluster.

B with 1280×720 resolution, and sample images from the dataset are depicted in Fig. 3. Video A and B were captured in Seoul campus of Hanyang university, and had 30 and 60 minutes duration, respectively. The evaluation process was conducted on Intel i5-2500 3.3 GHz computer with 4 GB memories.

Experiments were carried out for comparing performance of the proposed measure with existing measure in [2]. In detail, numbers of moving object trajectories obtained from two videos were 85 and 243, respectively; then, obtained trajectories were clustered by Affinity propagation with different similarity measures. When the proposed measure was utilized for the clustering, 11 and 22 clusters were obtained for video A and B, respectively. On the other hand, for the existing measure, 6 and 14 clusters were obtained. Details of clustering results for the video A is illustrated in Fig. 4 and Fig. 5; while, only subset of the clustering results for video B is depicted in Fig. 6, since number of pages for the paper is limited. In the figures, all trajectories are colored in rainbow, and the color has its own meaning. Points colored in green are closer to the start point; while, points colored in red are closer to the end point.

As shown in Fig. 4 and Fig. 5, there were a lot of people walking through a road which is located on the left side of the video. By using the existing measure, all the detail movements (some of the people are oriented to the left) of the people were grouped to a single cluster; while, clustering algorithm with the proposed measure could discriminate the detail movements.

In Fig. 6, difference of discriminative powers for the similarity measures could be observed more obviously. A cluster grouped by using the proposed measure only contains trajectories started from top-left to bottom-right of the image; however, trajectories clustered by using the existing measure have two distinctive shapes.

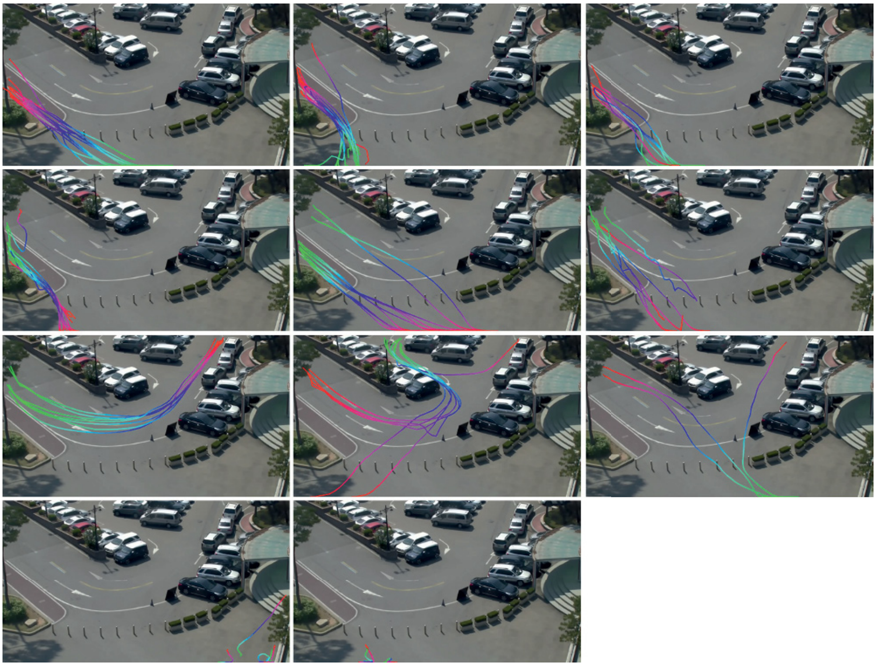

Fig. 5. Result of the clustering algorithm with the proposed measure. Video A is used for the source of the trajectories and total 11 clusters were obtained. Different from the clustering result of the existing measure, the proposed has a higher discriminate power than the existing measure.

5 Conclusion

In this paper, the similarity measure based on geographic and semantic similarities which have a same scale is proposed. Through the experiments on the challenging real-world dataset, the proposed measure is proved to have a better discriminative power than the existing method. Furthermore, a balance between the geographic and semantic similarities can be easily controlled, since they are combined by a form of weighted sum. However, the proposed method has a lack of ability to discriminate unusual trajectories, so that future research direction will be detecting unusual trajectories from the dataset. In addition, performance measure for the trajectory clustering quality is going to be researched to numerically analyze the clustering results.

Acknowledgments "This research was supported by the MSIP (Ministry of Science, ICT & Future Planning), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency)" (NIPA-2014-H0301-14-1018)

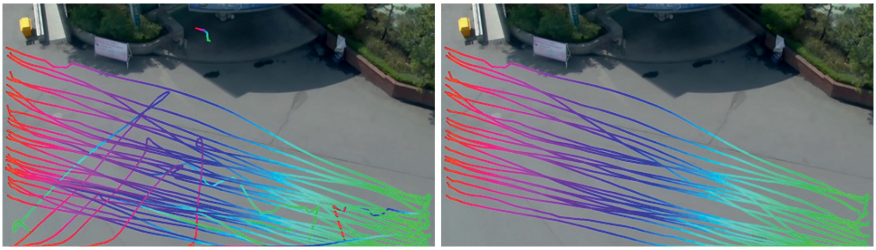

Fig. 6. Subset of the clustering results for video B. Left image is one of the results obtained by using the existing measure, and right image is one of the results acquired by using the proposed measure.

References

1. Zhang, Z., Huang, K., Tan, T.: Comparison of Similarity Measures for Trajectory Clustering in Outdoor Surveillance Scenes. In: 18th International Conference on Pattern Recognition, vol. 3, pp. 1135–1138 (2006)
2. Liu, H. and Schneider, M.: Similarity Measurement of Moving Object Trajectories. In: 3rd ACM SIGSPATIAL International Workshop on GeoStreaming, pp. 19–22 (2012)
3. Wang, H., Su, H., Zheng, Kai., Sadiq, S., Zhou, X.: An Effective Study on Trajectory Similarity Measures. In: 24th Australasian Database Conference, vol. 137, pp. 13–22 (2013)
4. Chen, P., Gu, J., Zhu, D., Shao, F.: A Dynamic Time Warping based Algorithm for Trajectory Matching in LBS. In: International Journal of Database Theory and Application, vol. 6, no. 3, pp. 39–48 (2013)
5. Bergroth, L., Hakonen, H., Raita, T.: A Survey of Longest Common Subsequence Algorithm. In: 7th International Symposium on String Processing and Information Retrieval, pp. 39–48 (2000)
6. Huttenlocher, D. P., Klanderman, G. A., Rucklidge, W. J.: Comparing Images Using the Hausdorff Distance. In: IEEE Transaction on Pattern Recognition and Machine Intelligence, vol. 15, no. 9, pp. 850–863 (1993)
7. Zivkovic, Z.: Improved Adpative Gaussian Mixture model for Background Subtraction. In: 17th International Conference on Pattern Recognition, vol. 2, pp. 28–31 (2004)
8. Gonzalez, R. C., Woods, R. E.: Digital Image Processing. Pearson Education, New Jersey (2010)
9. Munkres, J.: Algorithms for the Assignment and Transportation Problems. In: Journal of the Society for Industrial and Applied Mathematics, vol. 5, pp. 32–38 (1957)
10. Dubuisson, M.-P., Jain, A. K.: A Modified Hausdorff Distance for Object Matching. In: International Conference on Pattern Recognition, pp. 566–568 (1994)
11. Jesorsky, O., Kirchberg, K. J., Frischholz, R. W.: Robust Face Detection Using the Hausdorff Distance. In: Third International Conference on Audio- and Video-based Biometric Person Authentication, pp. 90–95 (2001)
12. Frey, B. J., Dueck, D.: Clustering by Passing Messages Between Data points. In: Science, vol. 315, 972–976 (2007)