MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

Learning Multi-Feature Human Motion Patterns by Automated Near-Optimal Constrained Gravitational Clustering

Zhuo Chen
Department of Electrical and Electronic Engineering, The University of Hong Kong
Pokfulam Road, Hong Kong, H.K.S.A.R.
zchen@eee.hku.hk

N.H.C. Yung
Department of Electrical and Electronic Engineering, The University of Hong Kong
Pokfulam Road, Hong Kong, H.K.S.A.R.
nyung@eee.hku.hk

Abstract

This paper proposes an automated near-optimal gravitational clustering method for learning multi-feature human motion patterns (HMPs). Based on the distance distribution of all observed human trajectories in a scenario, an automated criterion is proposed to determine the similarity threshold. Based on the threshold, clustering is automatically performed and multi-feature HMPs are accordingly learned. The proposed method defines a human trajectory in a better manner by representing the physical state of human as well as the motion characteristic of human. Furthermore, the main advantages of the proposed method are that (1) to derive an automated decision criterion instead of setting a similarity threshold manually for clustering; (2) to directly determine a near-optimal number of clusters without exhaustively searching for a global optimum. The proposed method has been tested in real-world experiments and the results show that the proposed method performs effectively in analyzing real-world HMPs.

1. Introduction

Understanding human motion patterns (HMPs) has been a major focus in recent years. There are a number of related applications, such as anomaly detection [1] and autonomous navigation robotics [2]. In real-world environments, it is extremely difficult to manually analyze HMPs since of an enormous amount of human trajectories collected over period of time. Therefore, it necessitates the use of unsupervised clustering methods for carrying out the learning task, such as direct clustering, agglomerative hierarchical clustering and graph cut [3]. Whichever clustering method is adopted, they are expected to generate optimal HMPs with the most reasonable number of clusters. Some popular clustering methods, such as fuzzy *k*-means clustering [4], require a desired number of clusters to be an input parameter. However, since any priori knowledge of HMPs is likely unavailable in a real-world environment, the number of clusters could be difficult to decide. Other clustering methods [5, 6] use a subjectively or empirically determined similarity threshold to control obtaining final clusters without an investigation of the clustering result. As such, some researchers have recently attempted to find optimal clusters [7-9]. Despite of the different criteria being relied upon, they share the same general concept that finding optimal clusters is equivalent to searching the minimum of a global cost function, which specifically evaluates the clustering result. In theory, minimum of a global cost function can be obtained if there is sufficient time. For *NP*-complete problems of this kind, such time may extend to infinity and makes it difficult to attain a viable solution within a certain time window.

In this paper, we propose an automated near-optimal constrained gravitational clustering method for learning HMPs. The advantages are mainly focused on two issues: (1) to derive an automated decision criterion instead of setting a similarity threshold manually for clustering; (2) to directly determine a near-optimal number of clusters without exhaustively searching for a global minimum. In the proposed method, the observed human trajectories are first represented by feature vectors that incorporate spatial location, velocity and change of heading angle. An automated criterion is proposed based on the distance distribution of all feature vectors for determining the similarity threshold. Then, clustering is automatically performed to obtain multi-feature HMPs using the similarity threshold. The proposed method has been tested in real-world experiments. By making both qualitative evaluation and quantitative analysis, the results show that the proposed method effectively generates near-optimal clusters in real-world scenes.

The rest of this paper is organized as follows. In Section 2, the proposed automated near-optimal constrained gravitational clustering method is presented. Section 3 depicts the result of a real-world experiment produced by the proposed method and Section 4 concludes the paper with a brief discussion of future research work.

2. Automated Near-optimal Constrained Gravitational Clustering

2.1. Constrained gravitational clustering concept

The concept of the constrained gravitational clustering (CGC) algorithm [10] imposes a constraint per iteration to control the clustering process, without a need to assign a termination condition. At the beginning of the clustering process, each human trajectory is regarded as the initial mean vector of a cluster. Here, $\mathbf{t}_k$ is defined as the feature vector of the k^{th} human trajectory which is given as:

$$\mathbf{t}_k = \{o_x^{(k)}, o_y^{(k)}, d_x^{(k)}, d_y^{(k)}, v_m^{(k)}, v_d^{(k)}, c^{(k)}, n_l^{(k)}, n_r^{(k)}\}, \quad (1)$$

where $(o_x^{(k)}, o_y^{(k)})$ and $(d_x^{(k)}, d_y^{(k)})$ denote the origin and the destination of the trajectory, respectively. $v_m^{(k)}$ and $v_d^{(k)}$ denote the mean and the standard deviation of velocity value at all time steps of the trajectory, respectively. $c^{(k)}$ denotes the curvature of trajectory. $n_l^{(k)}$ and $n_r^{(k)}$ denote the number of time steps of turning left and right when compared with the moving direction at the previous time step, respectively.

In principle, the whole clustering process is completely controlled by the attraction $F_{\mathbf{t}_k\mathbf{t}_i}$ between the k^{th} and i^{th} clusters, which is defined as:

$$F_{\mathbf{t}_k\mathbf{t}_i} = -G\frac{m_k \times m_i}{\left|\bar{\mathbf{t}}_k - \bar{\mathbf{t}}_i\right|^3}\left(\left|\bar{\mathbf{t}}_k - \bar{\mathbf{t}}_i\right|\right), \tag{2}$$

where G is the gravitational constant, m_k and m_i are the masses represented by the number values of human trajectories in the k^{th} and i^{th} clusters respectively, and $\bar{\mathbf{t}}_k$ and $\bar{\mathbf{t}}_i$ are centroids of the k^{th} and i^{th} clusters, respectively, as depicted in Figure 1.

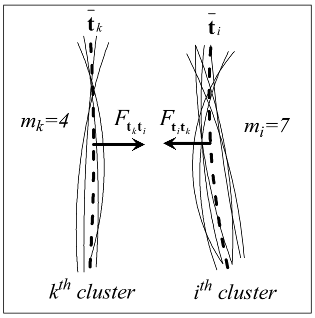

Figure 1. Parameters for gravitational attraction.

The CGC algorithm further proposed a net force acting on a trajectory that controls the movement of the trajectory and the formation of final clusters. The net force on a trajectory $\mathbf{t}_k$ is defined as:

$$F_{\mathbf{t}_k} = \sum_{\substack{i=1 \\ k \neq i}}^{n} F_{\mathbf{t}_k\mathbf{t}_i} \times W(\mathbf{t}_k, \mathbf{t}_i)\cdot \tag{3}$$

In (3), $W(\cdot)$ is a force effective function (FEF) that governs the effectiveness of the attractive force between trajectories, and only those trajectories that satisfy the FEF constraint contribute to the calculation of the net force on $\mathbf{t}_k$. $W(\cdot)$ can take the form of a step function which is given as:

$$W(\mathbf{t}_k, \mathbf{t}_i) = \begin{cases} 1 & \left|\bar{\mathbf{t}}_k - \bar{\mathbf{t}}_i\right| \leq D_F \\ 0 & \left|\bar{\mathbf{t}}_k - \bar{\mathbf{t}}_i\right| > D_F \end{cases}, \tag{4}$$

where D_F is the similarity threshold that controls the clustering. More generally, $W(\cdot)$ can take other form of decreasing functions versus the distance between trajectories, where D_F decides the zero-effectiveness of the attractive force between trajectories. With a non-zero net force, the trajectory moves and the clustering process continues until the iteratively calculated new net force on each trajectory is zero. Thus, final clusters are formed.

2.2. Automated decision criterion for near-optimal clusters

As the above description, D_F is a critical factor for the formation of final clusters. In one extreme, if D_F is equals to the maximum distance value of all pairs of trajectories, all the trajectories are clustered together. In the other extreme, if D_F=0, each trajectory represents a cluster. When D_F is in between the two extremes, it determines how clustering is to be performed as well as the final clusters. Unfortunately, the CGC algorithm did not provide a solution for deciding D_F.

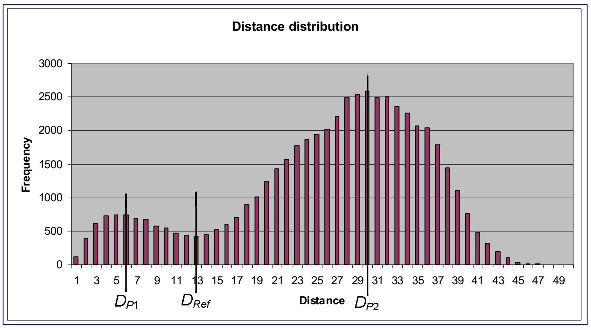

Figure 2. Distance distribution of all human trajectories.

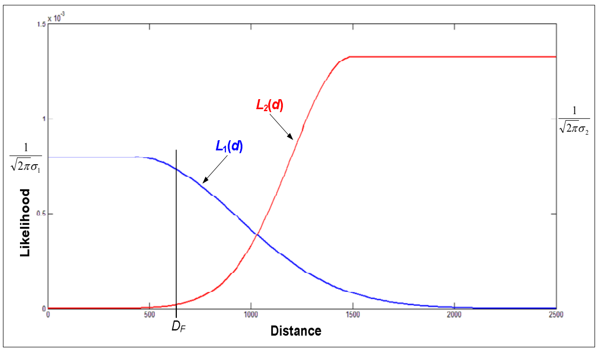

Figure 3. Determination of D_F.

In this paper, we propose an automated criterion for determining D_F which is based on the distance distribution of all human trajectories as depicted in Figure 2. The distance of a pair of trajectories is calculated as Euclidean distance between the corresponding feature vectors. A set of values for the number of bins have been tried for comparing different obtained distribution. It is found that the main characteristics of different distribution could be consistently depicted except when the number of bins is extremely large or small. In this paper, we evenly divide the range of distance values into 50 bins, then the distance distribution is obtained by calculating the average of all distance values that fall into a bin as the representative distance value of the bin, and counting the number of the distance values that fall into a bin as the corresponding frequency value of the representative distance value. The two main characteristics of the distribution are that: (1) a set of similar distance values fall into a bin that forms a frequency peak (marked at D_{P1} in Figure 2), which refers to the distance values of most trajectory pairs inside the same cluster; and (2) another set of similar distance values converge to be another frequency peak (marked at D_{P2} in Figure 2), which refers to the distance values of the most trajectory pairs in the different clusters. Thus, we can separate all distance values into two groups based on a boundary that locates at D_{Ref} as shown in Figure 2, which refers to a frequency minimum between two peaks. The left-hand group depicts the range of distance values of that two trajectories may be in the same cluster, and the right-hand group represents the range of distance values of that two trajectories may belong to different clusters.

Based on the distance values in each group, we calculate the mean and the standard deviation, where μ_1 and σ_1 are for the left-hand group and μ_2 and σ_2 are for the right-hand group. Then, two likelihood functions $L_1(d)$ and $L_2(d)$ are proposed based on (μ_1,σ_1) and (μ_2,σ_2), which are given as:

$$L_1(d)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma_1} & d\le D_{P1} \\ \dfrac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}\left(\frac{d-\mu_1}{\sigma_1}\right)^2} & d>D_{P1}\end{cases}, \quad (5)$$

$$L_2(d)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\left(\frac{d-\mu_2}{\sigma_2}\right)^2} & d\le D_{P2} \\ \dfrac{1}{\sqrt{2\pi}\sigma_2} & d>D_{P2}\end{cases}. \quad (6)$$

Here, d is the distance value between two human trajectories which is also calculated as Euclidean distance between corresponding feature vectors.

As depicted in Figure 3, $L_1(d)$ depicts a decreasing likelihood that two trajectories are in the same cluster versus an increasing distance between them, and $L_2(d)$ represents an increasing likelihood that two trajectories belong to different clusters versus an increasing distance between them. Therefore, the decision criterion for D_F is given as:

$$D_F=\arg\min_d\{CB(d)\,|\,d:L_2(d)\le L_1(d)\}. \quad (7)$$

In (7), $CB(d)$ is the clustering balance which is defined as the summation of intra-cluster error sum Λ_d and inter-cluster error sum Γ_d for the clustering result obtained based on each d value ($CB(d)=\Lambda_d+\Gamma_d$), where Λ_d and Γ_d are defined as:

$$\Lambda_d=\sum_{j=1}^{m}\sum_{i=1}^{n_j}\left\|\mathbf{t}_i^{(j)}-\bar{\mathbf{t}}_j\right\|_2^2, \quad (8)$$

$$\Gamma_d=\sum_{j=1}^{m}\left\|\bar{\mathbf{t}}_j-\bar{\mathbf{t}}_0\right\|_2^2, \quad (9)$$

where $\mathbf{t}_i^{(j)}$ and $\bar{\mathbf{t}}_j$ denote the i^{th} human trajectory in the j^{th} cluster and the mean trajectory of the j^{th} cluster, respectively; and $\bar{\mathbf{t}}_0$ denotes the global mean trajectory of all human trajectories. n_j is the number of human trajectories in the j^{th} cluster, and m is the number of final clusters, respectively. The criterion of optimal clustering proposed in [7] seeks the global minimum of clustering balance values. The d value with the minimal $CB(d)$ is accordingly determined as D_F. In order to obtain D_F more efficiently, we further add an optional condition ($L_2(d)\le L_1(d)$) in (7) to reduce the search range of d. Since $L_1(d)$ and $L_2(d)$ represent the likelihood of two trajectories being in the same cluster and in the different clusters, respectively, the d values that do not satisfy $L_2(d)\le L_1(d)$ are not effective candidates for D_F. When D_F is determined, a zero-effectiveness of the attractive force between two trajectories is decided by (4), and no movement exists between the two trajectories accordingly. Thus, the final clusters are obtained.

In order to further evaluate the criterion for determining D_F, we depict the curve of all $CB(d)$ values in Figure 4, in which each d value is represented by the average of all distance values in each bin as depicted in Figure 2. As shown in Figure 4, we especially mark the determined D_F and the range of d values that satisfy the optional condition in (7). The curve describes how the different similarity thresholds impact corresponding clustering results. It can be seen that the clustering balance of the clustering result by D_F determined by (7) reaches the same minimum (marked as $CB_{min}(d)$ in Figure 4) as the one when searching the entire range of d values. In other words, the D_F determined by the proposed produces near-optimal final clusters efficiently.

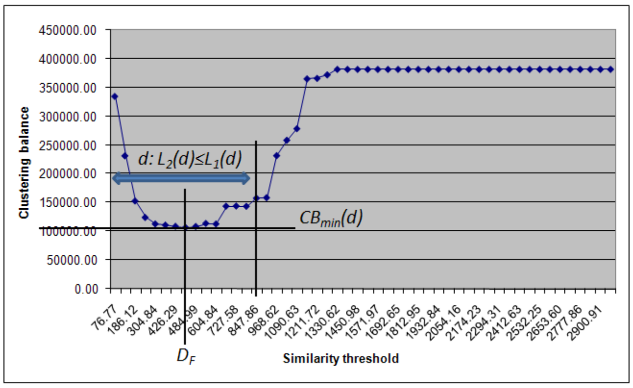

Figure 4. Curve of clustering balance of similarity threshold in the entire range.

3. Experiment

In this section, we demonstrate how the proposed method works in a real-world scene for learning HMPs. The scenario of the experiment is based on people walking freely in a shopping mall. A fixed-background video for this scenario was taken and a total of 326 observed human trajectories were accordingly extracted from the video. Figure 5 depicts all 326 observed human trajectories in which red-curves and green-curves represent double-directional trajectories between each pair of entrances/exits, respectively.

Figure 5. Observed human trajectories.

Figure 6. Learned HMPs.

The distance distribution of the observed human trajectories has been depicted in Figure 2. D_F is accordingly determined by Equation (7). HMPs are then learned by using D_F in the CGC algorithm as presented in Section 2.1. Figure 6 shows the learned HMPs, in which green-curves and red-curves represent the observed training trajectories and the learned HMPs, respectively, and the red circle labels the destination of each HMP. From the point of qualitative evaluation, it can be seen that the learned HMPs by the proposed method are reasonable for the real-world experiment.

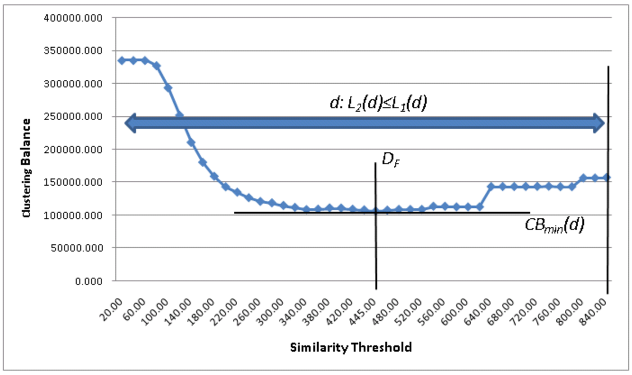

Figure 7. Curve of clustering balance of similarity threshold in the reduced range.

Since the ground truth cannot be easily obtained in a real dynamically changing environment, it is difficult to evaluate by directly comparing the clustering results with the ground truth. Here, a quantitative evaluation is performed on how close the clustering result by each similarity threshold is to the global optimal clustering. The evaluation is based on comparison of the clustering balance values ($CB(d)$) of different clustering results. From Figure 4, we can observe that all effective d values for determining D_F fall into the reduced range obtained by (7). It can be seen that 15 d values in all 50 similarity threshold values are in the reduced range and the d value with the minimal $CB(d)$ value is determined as D_F. In order to further evaluate the clustering result based on the D_F from (7), we select more similarity threshold values from the reduced range and for comparison with the clustering result by the proposed method. In the same reduced range, 42 d values are re-sampled for performing clustering. Comparing the $CB(d)$ values of the 42 clustering experiments with $CB(D_F)$ by the proposed method as depicted in Figure 7, we can see $CB(D_F)$ is also the minimum, which further proves that the clustering result by proposed method is the nearest to the global optimal clustering.

4. Conclusion

In this paper, we presented an automated near-optimal constrained gravitational clustering method for learning multi-feature HMPs in real-world scenes. An automated criterion is first proposed for determining the similarity threshold based on the distance distribution of all observed human trajectories. Then, clustering is automatically performed using the similarity threshold and multi-feature HMPs are accordingly learned. The main advantages of the proposed method are that (1) to derive an automated decision criterion instead of setting a similarity threshold manually for clustering; (2) to directly determine a near-optimal number of clusters without exhaustively searching for a global optimum. The proposed method also offers a better representation for a human trajectory, which represents the physical state of human as well as the motion characteristic of human. From the real-world experimental results, it is concluded that the proposed method can generate near-optimal clusters to be HMPs. Based on the obtained research results by the proposed method, our future research will focus on two issues: (1) to investigate online learning of HMPs and to further evaluate the learned HMPs in terms of dynamic change in the environment; (2) to interpret the learned HMPs in semantic level and based on which to research human behavior prediction problem.

References

[1] D. Makris and T. Ellis, "Learning semantic scene models from observing activity in visual surveillance," IEEE Transactions on System, Man and Cybernetics, Part B, Vol. 35, No. 3, pp. 397-408, June 2005.

[2] Z. Chen, D. C. K. Ngai and N. H. C. Yung, "Behavior prediction based on obstacle motion patterns in dynamically changing environments," Proceedings of the 2008 IEEE/WIC/ACM Joint Conferences on Web Intelligence and Intelligent Agent Technology, Sydney, Australia, December 2008.

[3] B. T. Morris and M. M. Trivedi, "Learning trajectory patterns by clustering: experimental studies and comparative evaluation," Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, Maimi, Florida, June. 2009.

[4] R. O. Duda, P. E. Hart and D. G. Stork. "Pattern classification," Wiley-Interscience, New York, 2001.

[5] J. Gomez, D. Dasgupta and O. Nasraoui, "A new gravitational clustering algorithm," Proceedings of the SIAM International Conference on Data Mining, San Francisco, U.S.A., May 1-3 2003.

[6] B. T. Morris and M. M. Trivedi, "Learning, modeling, and classification of vehicle track patterns from live video," IEEE Transaction on Intelligent Transportation Systems, Vol. 9, No. 3, pp. 425-437, September 2008.

[7] Y. Jung, H. Park, D. Du and B. L. Drake, "A decision criterion for the optimal number of clusters in hierarchical clustering," Journal of Global Optimization, Vol. 25, pp. 91-111, 2003.

[8] D. Buzan, S. Sclaroff and G. Kollios, "Extraction and clustering of motion trajectories in video," Proceedings of IEEE International Conference on Pattern Recognition, Cambridge, UK, August 2004.

[9] N. Suzuki, K. Hirasawa, K. Tanaka, Y. Kobayashi, Y. Sato and Y. Fujino, "Learning motion patterns and anomaly detection by human trajectory analysis," Proceedings of IEEE International Conference on Systems, Man and Cybernetics, Montreal, Canada, October 7-10, 2007.

[10] N. H. C. Yung and A. H. S. Lai, "Segmentation of color images based on the gravitational clustering concept," Optical Engineering, Vol. 37, No. 3, March 1998.