http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (1): 174-179, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.174.179

Detecting Group-level Crowd Using Spectral Clustering Analysis on Particle Trajectories

Dongping Zhang, Feiyu Chen and Huiliang Peng
College of Information Engineering, China Jiliang University, Hangzhou 310018, China

Abstract: Analyzing human crowds is becoming an important issue in video surveillance and one challenging task is to detect group-level crowd due to their non-rigid shapes nature. This study presents a novel method which synergistically combining two state-of-the-art methodologies to identify groups in crowds. The first is the ability to track crowd trajectories using particle video technology and the second is a new class of novelty clustering algorithms based on spectral analysis of graph. Simultaneity, the social science principle of human collective behavior, such as the similarity of location, velocity, appearance, is also inspired to cluster crowd trajectories. Experimental results demonstrate that our method is effective in tracking and identifying group-level crowds for public surveillance videos.

Key words: Group-level crowd, particle trajectory, spectral clustering analysis

INTRODUCTION

Now-a-days there is an increasing interest in human behavior analysis from video surveillance data. The crowd recognition technology has received some attention in recent years but most research has focus on multi-target tracking proposed by Zhao and Nevatia (2003) and Haritaoglu *et al.* (2000). And multi-target tracking in crowded scenes have appeared is computation intensive and can perform badly with even medium crowding situation. Other research in vision addresses in high-level crowd flow analysis. Although these works, corner feature trajectories or optical flow estimation, are sufficient to generate predictive model, they do not address the problem of groups crowd and their focus isn't on a basis for further midlevel analysis of events.

Therefore, researchers are particularly interested in seeking to address the problem of detecting and tracking about the evolution of crowd structures so as to understand group-level activities of the crowd for further research. In computer vision, typical grouping strategies Lau *et al.* (2010) relied on a bottom-up, agglomerative clustering of individuals in order to find the group structure. Ge *et al.* (2009) identified small group structure of a crowd in such a typical grouping strategies by iteratively merging sub-groups with the strongest intergroup closeness, utilizing a measure based on the symmetric Hausdorff distance. This is appropriate in environments where observed person-to-person distances follow standard social norms but not in environments where rapid changes in interaction distances occur. Lau *et al.* (2010) posed the group modeling problem as a recursive multi-hypothesis model selection, the merging of two groups is justified using a Mahalanobis distance between closest contour points. Such grouping schemes require a threshold to determine the final grouping. Hard thresholds on the blob size are used to identify whether a blob is a person or a group of people. As opposed to grouping individual tracks, there are also algorithms that identify grouping without necessarily identifying individuals in the groups. Chen and Huang (2011) used optical flow estimation without identifying individuals for a clue to cluster human crowds into groups.

This study builds upon state-of-the-art detection technology of group-level crowd which is synergistically combining two state-of-the-art methodologies. The first is the ability to detect and track crowd using particle video technology which is advance a long-range motion estimation to overcome that drawback in optical flow. The second is a new class of novelty clustering algorithms based on spectral analysis of graph. In this clustering algorithm, the particle trajectories with similar motion and appearance information are clustered as group by the psychological concept.

DETECTION AND TRACKING CROWD

This section describes the novel detection and tracking approach that can be effectively used to group pedestrians in public surveillance. Initially, crowd detection technique are using a method which is the

Corresponding Author: Dongping Zhang, College of Information Engineering, China Jiliang University, China Tel: 18888970656

combination of successive frame edge comparison technique and background maintain-update technique as described in detection section. Then detected crowd in detection section are tracked with long-range particle video technique. In tracking section our tracking approach will describe in more detail.

Detection: This part of section will provide a brief outline of our segment method which is mainly consisted of two parts of detection techniques. One of them is about the difference between the current frame and background frame. In this processing step, the up-to-date background information from the video sequence is constructed and maintained with Gaussian mixture model proposed by Kim and Hwang (2002). When comparing each frame with the background, it is assumed to be in object region, if the region of any pixel is significantly different from the background. Due to compare with RGB and YUV color space, HSV is more appropriable to compute the difference between pixel colors. So the RGB space is converted to HSV space and the difference between current frame and background frame is going to be presented by the threshold of pixel.

The other one is using the moving edge from success frames to segment the regions of motion crowds. In segmentation algorithm, edge information plays a key role in extracting the physical change in a real scene but it is suffering great deal of noise even in stationary background. In algorithm for segregation is more exact. It mainly consist of three steps, the canny edge calculation, motion edge calculation and motion objects obtained. When obtained the edge map, the motion regions are ready to be extracted with the method proposed by Meier and Ngan (1998). It's about that firstly finding both the horizontal candidates and the vertical candidates, then the intersection regions obtained by logical AND operation is further processed by morphological operations. In order to enhance the accuracy of regions it is added plus-minus 45 degree candidates to the method.

The ultimate detection result of motion regions combined with two methods in different frames are shown in Fig. 1.

Tracking: Fundamental to the success of any algorithms for recognizing group activities is the ability to track individuals (or group of individuals) under crowded conditions. Several vision-based techniques for people tracking in video sequences have been proposed in the past years, most of them with surveillance purposes. In these applications, the crowd density in unsubstantiated regions is always sparse, so that the faces of the individuals can be recognized. However, such group-level crowd result in occlusions and the goal of extract trajectories for each individual may not be possible.

This part of section will employ particle video motion estimation by Sand and Teller (2008) to track detected crowd. The key advantage of particle video approach is that it is both spatially dense and temporally long-range. In contrast, feature tracking is long range but spatially sparse and optical flow is dense but temporally short-range. Thus, particle video data is ideal for the purpose of attaching annotations, can estimate the motion of any pixel into any frame by finding a nearby particle. For each particle video data has the same 5 channels, image brightness, green minus red channel, green minus blue channel, x gradient and y gradient as the flow estimation algorithm. The process of computation is shown in Fig. 2, the steps are performed.

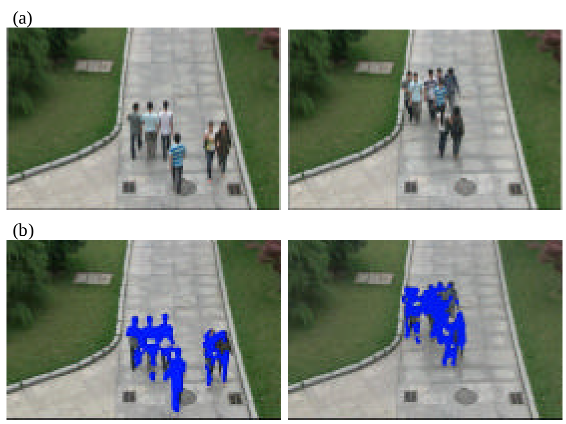

Fig. 1(a-b): Detection result of motion regions, (a) The different original frames and (b) The detected crowds are in blue color

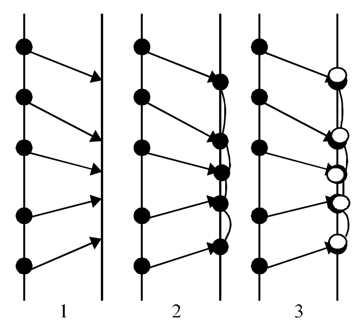

Fig. 2: The process of computing the particle trajectory, 1: Propagate, 2: Link and 3: Optimize using particle video

Propagation: Particles terminating in an adjacent frame are extended into the current frame according to the forward and reverse optical flow fields. To propagate particle i from frame t-1 to t, use the flow field:

$$\begin{aligned} x_i &= x_i(t-1) + u(x_i(t-1), y_i(t-1), t-1) \\ y_i &= y_i(t-1) + v(x_i(t-1), y_i(t-1), t-1) \end{aligned} \quad (1)$$

Linking: To quantify relative particle motion, our algorithm need create links between particles. It simply link each particle to its N nearest neighbors. Using links from adjacent frames reduces temporal linking variability, while still allowing links to appear and disappear as particles pass by one another. For each link (i, j), it computes a squared motion difference according to the flow vectors:

$$D(i,j) = \frac{1}{|T|}\sum_{t\in T}(u_i(t)-u_j(t))^2 + (v_i(t)-v_j(t))^2 \quad (2)$$

Use a Gaussian prior ($\sigma = 1.5$) as a weight to the link on the motion difference:

$$l_{ij} = N(\sqrt{D(i,j)}; \sigma_l) \quad (3)$$

Optimization: The core of the particle video algorithm is optimization process which can remodify the particle position caused by particle drift problem in long-segment diffusion process. The algorithm substantively is to minimize an objective energy function in particle i of the t frame's as follows:

$$E(i,t) = \sum_{k\in K_i(t)} E_{Date}^{[k]}(i,t) + \alpha \sum_{j\in L_i(t)} E_{Distur}(i,j,t) \quad (4)$$

that includes two parts: a data term and a distortion term. The data term measures how well a particle's initial appearance vector matches the current frame and $K_i(t)$ denotes the set of active channels. The distortion term measures relative motion relative motion between the particles. If two linked particles move in different directions, they will have a larger distortion term. If they move in the same direction, they will have a smaller distortion term and $L_i(t)$ denotes the set of particles linked to particle i in frame t.

The algorithm optimizes Eq. 4 in a manner similar to the variational technique. Finding that a = 1.5 provides is a reasonable trade-off between the two terms using a fixed-point loop around a sparse linear system solver. Within the objective function E, $dx_i(t)+x_i(t)$ is substituted for $x_i(t)$ (and instances of y accordingly). A system of equations is obtained by taking partial derivatives which the algorithm solves for $dx_i(t)$ and $dy_i(t)$:

$$\{\frac{\partial E}{\partial dx_i(t)} = 0, \frac{\partial E}{\partial dy_i(t)} = 0 \mid i \in p, t = F\} \quad (5)$$

The $dx_i(t)$ and $dy_i(t)$ values produced by solving this system are added to the current particle positions ($x_i(t)$ and $y_i(t)$).

SMALL GROUPS CLUSTERING

Small groups clustering algorithm is always inspired by McPhail and Wohlstein (1982) hypothesize, who present the only objective measure known of in the social science literature to determine which people are traveling together through the scene. In their literature, groups in crowd are determined by the three test: (1) any two people who are within a fixed distance, (2) any two contiguous people are judged to have the same speed and (3) any two contiguous people traveling at the same speed whose directions of motion are the same. These tests have generally been applied by researchers who analyze video surveillance. McPhail and Wohlstein (1982) hypothesized and present a clustering approach that hypothesize small group traveling together using the notion of group "entitativity", defined in terms of criteria from social cognition: similarity (in appearance or behaviors), proximity (in location, motion velocity). Potential groups are identified in crowd regions using spectral clustering based on robust measurement to particle video trajectories.

Measurement: Taking into account the group-level crowd obstructed by neighboring pedestrian, the individual template tracking is not practical to operate. So the proposed measurement strategy not any longer rigidly adheres to traditional individual trajectory measurement. Particle trajectories similarity is analyzed in tracked crowd to judge the distribution of group people.

To detect group formation, the particles from crowds are tracked to measure the similarity with its neighbor across time. If two or more particles keep consistent similarity among them in a certain period of time, the method considers they belong to the same group. The similarity criterion concerned is not only the space distance and motion velocity distance but also appearance channel distance which is peculiar to video particle trajectory. The appearance channel especially in color and gradient channels have the character about the moderate insensibility to changes in lighting and reflectance which facilitates judging the particle trajectory in outdoor environment. Denote the trajectory of a particle in the scene as a set of tuples (s, v, c, t), where s is the position vector of the particle and the v is the motion velocity at frame t, c is the channel vector about particle appearance. In comparison with McPhail and Wohlstein (1982) frame-based test, the extended method is to an aggregated pairwise similarity measure between two trajectories over time:

$$S_{ij} = \frac{w_1 * \sum_t \left\| s_i^t - s_j^t \right\| + w_2 * w * \sum_t \left\| v_i^t - v_j^t \right\| + w_3 * \sum_t \sum_{k=1}^{N} \left(c_i^k - c_j^k \right)}{\Gamma} \tag{6}$$

$$w = \frac{\Gamma}{(2 \times T_{max} + \Gamma)} \tag{7}$$

where, Γ is be the temporal overlap of the trajectories between particle p_i and p_j. Compute the weight parameter w by overlap time and T_{max} which is set to the value of the longer existing time between particle i and j. Though the scaled weighing parameter w, the average pairwise similarity S_{ij} over time Γ favors representing the particle relationship. The others weight parameters are set along with video size change.

Clustering algorithm: The identify groups algorithm based on a spectral clustering approach that starts with particle trajectory similarity measured in section A. In order to achieve more effectively clustering analysis, a diagonal matrix is constructed to group the particle trajectories, as shown in Table 1. The similarity S_{ij} between particle p_i and p_j is computed by Eq. 7. When S_{ij} satisfies $S_{ij} \leq \sigma$, the two particle are considered positively related but not confirm they are in the same cluster and accordingly the location $p_i p_j$ of adjacency matrix is set the value of S_{ij}. Otherwise, the two clusters are considered negatively related and thus the location $p_i p_j$ of adjacency matrix is set as 0. The adjacency matrix is demonstrated in Table 1.

Then the spectral clustering algorithm is detailed as follows:

- Form the affinity matrix A defined by $A_{ij} = \exp(-d^2_{ij}/2\sigma^2)$, if $i \neq j$ and $A_{ij} = 0$
- Define D to be the diagonal matrix whose (i,j)-element is the sum of A's ith row and construct the matrix $L = D^{-1/2} A D^{-1/2}$

Table 1: From the matrix, the elements except for diagonal elements are set "0" in four locations and the particles between p_3 and p_5, p_5 and p_N are can't be clustered in one group

	p_1	p_2	p_3	p_4	p_5	…	p_{N-1}	p_N
p_1	0	$S_{1,2}$	$S_{1,3}$	$S_{1,4}$	$S_{1,5}$	…	$S_{1,N-1}$	$S_{1,N}$
p_2	$S_{1,2}$	0	$S_{2,3}$	$S_{2,4}$	$S_{2,5}$	…	$S_{2,N-1}$	$S_{2,N}$
p_3	$S_{1,3}$	$S_{2,3}$	0	$S_{3,4}$	0	…	$S_{3,N-1}$	$S_{3,N}$
p_4	$S_{1,4}$	$S_{2,4}$	$S_{3,4}$	0	$S_{4,5}$	…	$S_{4,N-1}$	$S_{4,N}$
p_5	$S_{1,5}$	$S_{2,5}$	0	$S_{4,5}$	0	…	$S_{5,N-1}$	0
…	…	…	…	…	…	0	…	…
p_{N-1}	$S_{1,N-1}$	$S_{2,N-1}$	$S_{3,N-1}$	$S_{4,N-1}$	$S_{5,N-1}$	…	0	$S_{N-1,N}$
p_N	$S_{1,N}$	$S_{2,N}$	$S_{3,N}$	$S_{4,N}$	0	…	$S_{N-1,N}$	0

- Find $x_1, x_2, \ldots, x_k$, the k largest eigenvectors of L (chosen to be orthogonal to each other in the case of repeated eigenvalues) and form the matrix $X = [x_1, x_2, \ldots, x_k] \in R^{m \times k}$ by stacking the eigenvectors in columns
- Form the matrix Y from X by renormalizing each of X's rows to have unit length (i.e. $Y_{ij} = X_{ij}(\Sigma_j X^2_{ij})^{1/2}$)
- Treating each row of Y as a point in R^k, cluster them into k clusters via k-means
- Finally, assign the particle p_i to cluster j if and only if row i of the matrix Y was assigned to cluster j

To summarize, starting from large number of particle trajectories, the similarity matrix of each other from three aspects, location, motion velocity and appearance channel is constructed to include all particle trajectories information. Then the spectral clustering uses the eigenvectors of the matrix to denote original data. In the end, it clusters them into k clusters via k-means. Compare to traditional clustering algorithm, such as K-means, it is added the step of reducing dimensions by calculation of the eigenvectors. It makes the algorithm more robust and better performance than traditional ones and it is much suitable to the complicated trajectory data.

EXPERIMENTS RESULT

The method has been tested on videos of some scenes captured in our campus, some are shown in Fig. 1. In the test dataset Fig. 3 shows, a group of people heads towards directly along the path, at the same time four or five students walk together towards opposite the other group. Seen from the larger group, it is difficult to identify individual from each other and much harder to track with them by traditional temple tracking algorithm. By using video particle tracking algorithm, crowd are tracked and particle trajectories are presented as shown in Fig. 3, which is useful to cluster pedestrians that move together over time. In the clustering algorithm, the basic parameters is generally required to tune for a given sequence, depending on the video frame size, the crowd density and the speed of individuals and the distance threshold in this test video sequences both are set to 10. The clustered groups result can be obviously noticed.

More results are shown in Fig. 4. There are three groups walking in the surveillance view and two groups are very close in our view angle. In traditional distance cluster method, they are exceedingly judged to the same group in error. Identifying groups based on a spectral clustering approach is concerned not only the space

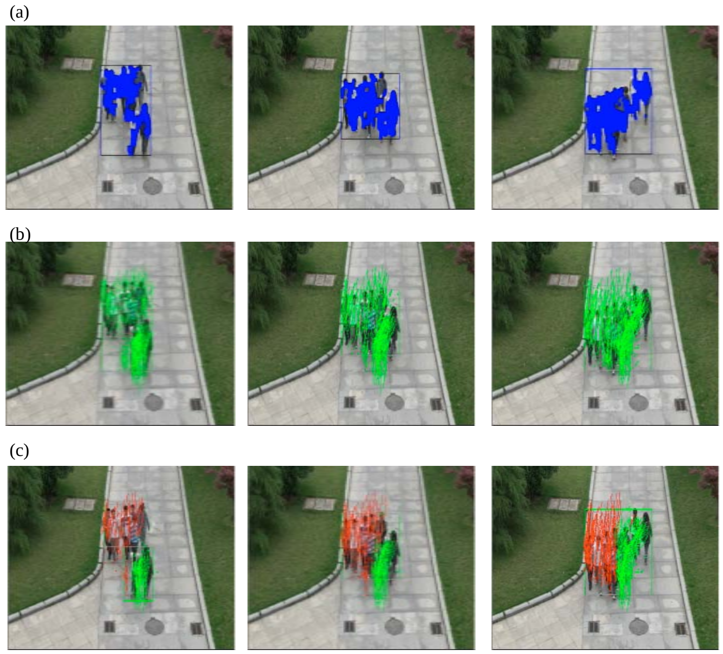

Fig. 3(a-c): The experiment result of the first video sequence, (a) The motion crowd are detected from three different frame in the first video sequence, (b) Trajectories corresponding with (a) and (c) Clustering results show the different groups by two colors even if they are so closed

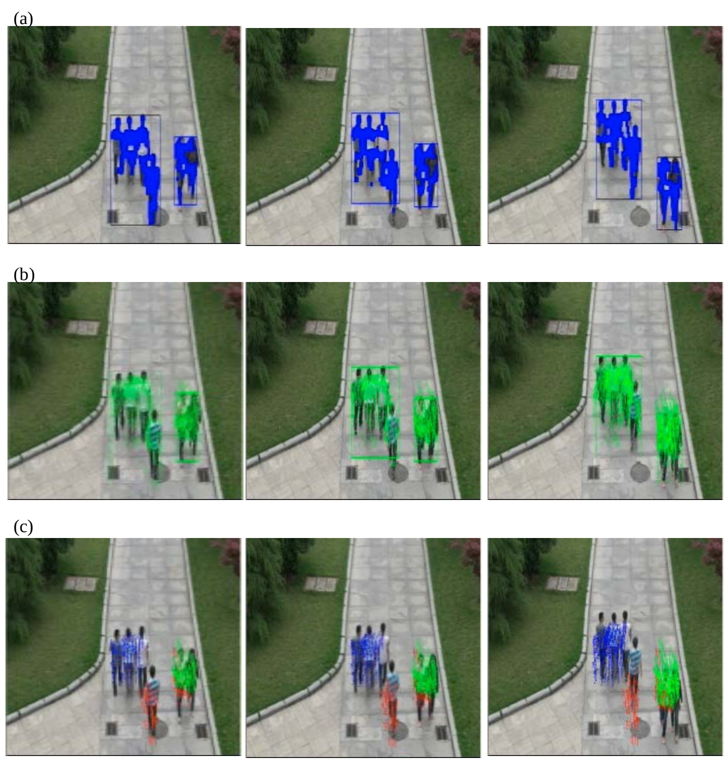

Fig. 4(a-c): Another video sequences are tested with our algorithm, (a) Detection results, (b) Corresponding trajectories and (c) Clustering results show the three different groups by three different colors distance and motion velocity distance but also appearance channel distance. To measure the performance of our method, compare the results of the method with the ground truth on the test data and the correct detection percent of groups has reached to 90%.

CONCLUSION

This study describes a novel approach for detecting human groups in videos. The framework of detection and tracking algorithm can serve as a prior step for group clustering algorithm analysis. The proposed algorithm involves crowd detection and tracking. Finding that extracting accurate and continuous trajectories in obstruct crowd is the key point for successful detection of human groups. It was found that particle trajectories in the algorithm of particle motion estimation in tracking crowd are much reliable. Finally, analysis the character of trajectories in human crowd to cluster them based on spectral clustering algorithm. The clustering algorithm is inspired by the social science literature to determine which people are traveling together, while the particle trajectory's peculiarity is also considered. Experimental results obtained by using our test video sequences have shown that our system is effective in detect small group people in crowd for uncontrolled environment of surveillance videos.

ACKNOWLEDGMENTS

This study was supported by Zhejiang Provincial Science and Technology Research Program of Zhejiang Province under grant 2010C1010 and Zhejiang Provincial NSF (Grant No. Y1110506). The study was also funded by the Opening Project of Shanghai Key Laboratory of Integrate Administration Technologies for Information Security.

REFERENCES

Chen, D.Y. and P.C. Huang, 2011. Motion-based unusual event detection in human crowds. J. Vis. Commun. Image R., 22: 178-186.

Ge, W., R.T. Collins and B. Ruback, 2009. Automatically detecting the small group structure of a crowd. Proceedings of the IEEE Workshop on Applications of Computer Vision, December 7-8, 2009, Snowbird, UT, pp: 1-8.

Haritaoglu, I., D. Harwood and L.S. Davis, 2000. W4: Real-time surveillance of people and their activities. IEEE Trans. Pattern Anal. Machine Intell., 22: 809-830.

Kim, C. and J.N. Hwang, 2002. Object-based video abstraction for video surveillance systems. IEEE Trans. Circuits Syst. Video Technol., 12: 1128-1138.

Lau, B., K.O. Arras and W. Burgard, 2010. Multi-model hypothesis group tracking and group size estimation. Int. J. Social Rob., 2: 19-30.

McPhail, C. and R. Wohlstein, 1982. Using film to analyze pedestrian behavior. Sociol. Methods Res., 10: 347-375.

Meier, T. and K.N. Ngan, 1998. Automatic segmentation of moving objects for video object plane generation. IEEE Trans. Circuits Syst. Video Technol., 8: 525-538.

Sand, P. and S. Teller, 2008. Particle video: Long-range motion estimation using point trajectories. Int. J. Comput. Vision, 80: 72-91.

Zhao, T. and R. Nevatia, 2003. Bayesian human segmentation in crowded situations. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, June 18-20, 2003, Madison, Wisconsin, pp: 459-466.