Crowd Group Detection by Using Collective Transition Prior with K-Means Clustering

Naw Naw  
Department of Information Science, University of Technology (Yatanarpon Cyber City)  
nawnaw@utycc.edu.mm

Pwint Theingi Aung Win  
Department of Information Science, University of Technology (Yatanarpon Cyber City)  
pwinttheingiaungwin91097@gmail.com

Abstract

In this era, behavioral detection is mostly vital for every surveillance system. The group of people sometime can be dangerous in some situation. There is a solution to detect the group of people and their directions for various places with association or not, which is a KLT feature tracker with collective transition prior. The video frames are used for doing the crowd analysis purpose. This system aims to detect the number of group from crowd video. KLT (Kanade-Lucas-Tomasi) feature tracker is used to generate tracklets from crowd video. Collective Transition is used for group detection. K-means clustering is used for finding dynamics of K in individuals for consecutive time in coherent filtering.

Keywords: KLT, K-means clustering, collective transition, behavioral detection, surveillance system

1. Introduction

The numbers of people group who are similar in physical nature are group together in coherent motion system. From bacterial provinces and bug swarms to fish reef, aggregate movements broadly exist in various group frameworks and reflect the arranged perceptible practices of constituent people. Behavioral studies demonstrate that mind boggling swarm practices may result from repeated basic communications among its constituent people, i.e., people locally arrange their practices with their neighbors then the group is self-sorted out into aggregate movement without outside control. Collectiveness depicts the level of people going about as an association in aggregate movement. It relies upon various components, for example, the basic leadership of people and group thickness. Quantitatively estimating this all inclusive property and looking at it crosswise over various group frameworks are significant so as to comprehend the general standards of different group practices. Collectiveness depicts the level of people going about as an association in aggregate movement. It relies upon various factors, for example the decision making of individuals and crowd density [5].

This paper is intended to develop an automatic human tracking and group detection which is based on human crowd video. It is an important feature for future video surveillance systems. It can improve a system’s performance in fields such as security, safety, human activity monitoring etc. Human Group detection systems can have different goals such as detecting the presence of humans, recognition of abnormal behavior, crowd management, public space design and others etc.

The design model of group detection system can be subdivided into three parts. The first part is generating tracklets using KLT feature tracker. In the second stage, coherently motion patterns are detected using Coherent Filtering. Finally, the number of group is manipulated from coherent motion patterns by filtering two or more motion people.

Precise group detection in crowd is challenging due to complex interaction among pedestrians. This system is proposed to generate the number of groups from crowd system using Coherent Transition Prior. This system will help to characterize people behaviors, predict and prevent potentially dangerous situations and improve the well-being of communities in future. Using the crowd video from stations, supermarkets and other crowd places, the system firstly generates the tracklets by KLT feature tracker. Coherent motion patterns are generated using these tracklets as input by Coherent Transition Prior. And then the number of group is manipulated by filtering two or more motion people.

2. Related Works

Bolei Zhou and corresponding authors proposed a collective merging algorithm for detecting collective motions form random moving groups [2]. While Interactive Projected Clustering algorithm (IPCLUS) [1] performed the clustering in high dimensional data, the algorithm can also detect the cluster from the low dimensional data. Mixture model of Dynamic pedestrian-Agents (MDA) method was designed for catching the crowd from their behavior patterns [3]. This method
could relate between people past behavior and future one. The crowd group could also be identified by applying bottom-up hierarchical clustering which used symmetric Hausdorff distance approach [4].

3. Consideration of the Groups

A group is in excess of a social affair of people. In specific situations, people in a group are sorted out into solidarity with various degrees of aggregate movements. Group should be determined by the collectiveness of its constituent individuals, which reflects the similarity of the individual’s behavior to others in the same crowd system. Thus, a group is considered as a set of members with a common goal and collective behaviors.

3.1. KLT Feature Tracker

KLT feature tracker is used to track the motion path of people in crowd video. KLT can extract the features by automatically changing its parameters. The tracker can trace the lost features from the new features of the new images. The tracker may abandon the features with have large distance value between the previous and current frames [8].

3.2. Group Detection

There are two methods in group detection which are collection transition and coherent neighbour invariance.

Collection Transition: The coherent motion patterns are found by using the tracklets which are resulted from KLT Feature Tracker by Collective Transition Prior. Firstly, Coherent Filtering method is proposed for detecting real time motion. In figure 2, the system evaluate as representing as one color in one crowd group. As a second step, the system picks the random cluster firstly and then it also finds its anchor cluster with long duration and low variance. After the system can also discover the tracklets for learning CT priors, the Markov Chain method can perform the group refinement [5].

Coherent Neighbor Invariance: In this method, the central point of time t is firstly defined. If the neighbors have the spatiotemporal and velocity correlations with the central point, they remain invariant over time. In this system, the coherent neighbor invariance method is used for crowd group detection. The detail algorithm can be found in [6].

Figure 2. Coherent motion patterns

Figure 3 shows the details system flow of the detecting coherent motion patterns process. Firstly, we extract image sequences from crowd video for generating tracklets process. Then, we input crowd video to extract tracklets into the KLT feature tracker. Then, we input the extracted tracklets into Coherent Filtering Algorithm. In the detection of coherent motion patterns, the detected coherent motion patterns are inclined to be divided into parts as K groups. The different results of K group are discussed in section 4. We use k-mean clustering as a great solution for pre-clustering, reducing the space into disjoint smaller sub-spaces and also its simplest implementation reasons.
4. Analysis of the System

In this system, crowd video that has frames between 100 and 200. The image sequences are split from crowd video by 20 frames per second. We use two datasets in video detection which are cross road dataset and grand central station dataset of Japan.

In Coherent Filtering, the scale of coherent motion pattern is generated according to the size of cluster K that used in the algorithm. The cluster size can be decided the size of K. While the noise can be found in large K, the motion pattern will be induced in small K. However, the scale of the coherent motion patterns of a video is highly infected to K value.

The resulted number of coherent motion patterns is related to the size of the random k value. We can compare the detection rate over different size of K. We can formulate the detection rate (DR) using the following formula. We can define the number of undetected group as False Positive (FP) and the number of detected group as True Positive. By using the False Rate (FR) of equation (1) is useful for calculating the detection rate of equation (2).

\[ FR = \frac{FP}{TP+FP} \times 100\% \quad (1) \]

\[ DR=100\% - FR \quad (2) \]

The method has been tested on videos of some scenes captured in Japan road cross point. In the test dataset shows, a group of people heads towards the opposite side. Seen from the larger group, it is difficult to identify individual from each other and much harder to track with them. By using video particle tracking algorithm, crowd are tracked and particle trajectories are presented as shown in Figure 4 and 5, 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 tested with various k size.

The clustered groups result can be obviously noticed which can be different in accuracy in group detection. As in figure 4, although we have large number of k size, the system can detect the a few correct groups. The false rate is greater than the detection rate in this system obviously.

More results are shown in Figure 5. There are five 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. We ignore the other clusters as wrongly detected in figure 4. The accuracy result is increased apparently over 16% than large k value. The gap of k value is over double between figure 4 and 5. However,
the real problem of accuracy decreasing in this problem is the dataset. The numbers of groups are randomly walking in this point and the data cannot be detected easily.

In the detection of coherent motion patterns, the detected coherent motion patterns are inclined to be divided into cluster. While the detection rate is increased in small k value, however, when K is too large, some noise might be included. We can detect this propose by using another dataset (Grand Central Station dataset). As shown in figure 6, the false rate is less than the detection rate. However, our predicted output group number is still less the correct group number on ground truth data.

![Figure 6. Detection coherent motion patterns using k=10](image)

However, the accuracy rate is sharply increased when we reduce the cluster size. The detection rate of the figure 7 is 80% which is greater than the k=10. That result can prove significantly about the size of the K value is changing based on the crowd size. The Grand Central Station dataset have the simpler prediction on the crowd based detection system than previous dataset.

![Figure 7. Detection coherent motion patterns using k=4](image)

**4. Conclusion**

In this system, a Collective Transition Prior is used to detect the coherent motion clusters from video. Number of tracklets is generated from particular video using KLT feature tracker. Using these tracklets, coherent motion patterns is resulted using Coherent Filtering. Finally, number of group is manipulated from coherent motion patterns by filtering two or more motion people. This system is designed to generate only the number of group from crowd video. The detection rate can’t be accurate sometimes as the number of resulted group is related to the size of cluster, K which is a parameter of Coherent Filtering. It cannot output the number of people in each group and the exact position of group. This system can be used as a basis prior to develop crowd management system, crowd group behavior analysis, public area visual surveillance and other crowd based computer vision systems. As another method, we can use bottom-up hierarchical clustering method for determining small groups in pedestrian crowds.

**References**