Clustering Textures with EHG Algorithm for Modelling Video

U.Saravanakumar
PG Scholar, Department of Computer Science and Engineering,
PRIST University, Trichy District, Tamilnadu, India.
saravanakumar23051988@gmail.com

Abstract: In this paper we present a novel approach for common recognition of group activities for video surveillance applications. We propose a Energetic-based approach for detecting abnormal events in surveillance video. It requires the appropriate definition of similarity between events. Human pose estimation via motion tracking systems can be considered as a regression problem within a discriminative framework. We defined the overfitting problem was handled by Hidden Markov Model based similarity. We propose in this paper a multi model-based similarity measure. In this measure, the Hidden Markov Model training and distance measuring are based on multiple samples. The novel Energetic Hierarchical Group (EHG) method acquired the multiple training data. By iteratively reclassifying and retraining the data groups at different clustering levels, the initial training and clustering errors due to overfitting will be sequentially corrected in later steps. Experimental results on real surveillance video show an improvement of the proposed method over a stand column method that uses single sample- based similarity measure and spectral clustering.

Keywords: *pose estimation, group event detection, clustering, group representative, surveillance, motion tracking systems.*

1. INTRODUCTION

Identifying human behavior or human interactions has attracted increasing the research interests [1-6]. The following events are group events.

- people fighting
- people walking together
- people being followed
- group conversations in a party
- terrorist launching attacks in groups

In this paper we propose a multi model-based similarity measure to hold back the overfitting problem, where Hidden Markov Model representation is based on several similar samples. The acquisition of these several training data is by hierarchically collect and iteratively retraining the whole dataset, which is summarized as Energetic Hierarchical Group (EHG) algorithm. This algorithm can animatedly correct initial overfitting errors as the numbers of training samples increase (i.e. data clusters become bigger).

In addition, it is not sensitive to the absolute values of similarity, because simple comparison operation instead of eigenvalue decomposition is needed in the proposed approach.

In real videos, the suspicious events are rare, difficult to describe, hard to predict and can be subtle. However, based on the assumption that an abnormal event is associated with the distinctness of the activity. (e.g., a running person where everybody walks is interpreted as abnormal as well as a walking person where the rest run) and a normal event indicates the commonality. (e.g., a path that most people walk on)In this paper, we address the following issues for cluster incident discovery.

1.1 Cluster incident discovery with supple or unreliable number of group members

Most previous cluster event detection researches [1-2] use a Hidden Markov Model or its variation to model the human interactions. Some people try to recognize human interactions based on a content-independent semantic set [3-4]. However, most of these works are designed to recognize group activities with a fixed number of group members, where the input feature vector length is fixed.

They cannot handle cases where the number of group members is supple or even unreliable, which is often the case in our daily life (e.g., people may leave or join a group activity). In this case, the input feature vector length may vary with different number of group members.

1.2 Cluster incident discovery with a Hierarchical Activity Structure

In many scenarios, interacting people form subgroups. However, these subgroups are not independent to each other and they may further interact to form a hierarchical structure. For example, in Fig.1, three people fighting form a subgroup of fighting.

Fig 1 – Group activity

At the same time, another person is approaching the three fighting people and these four people form a larger group of approaching. This is an example of hierarchical activity structure with the cluster of approaching at a higher level than the group of fighting. Some algorithms [1-2] could be extended to deal with the problem of hierarchical structure event discovery when the number of group members is fixed.

However, to the best of our knowledge, our work is the first to address the problem of cluster incident discovery with a varying number of group members under a hierarchical activity structure.

1.3 Clustering with an Abnormal space Metric

Most previous clustering algorithms [6,10] perform clustering based on a symmetric distance metric (i.e. the distance between two people is symmetric regardless of the relationship of the people). In the group event detection, some activities such as “following” are asymmetric (e.g. person A following person B is not the same as person B following person A). Defining a suitable asymmetric distance metric and performing clustering under the asymmetric distance metric is an important issue.

2. GROUP-BASED APPROACH

2.1 HMM representation of video events

In many existing work on surveillance video analysis [2,3,6,7], video events are represented as object trajectories or time evolutions of certain frame features, which can be further modeled by HMM.

2.2 Detection of abnormal events

Based on the models of normal groups, detection of abnormal events can be performed to new video data. Specifically, given an unseen object trajectory *i*, the likelihood of observing *i* given any Hidden Markov Model of normal events.

2.3 Energetic Hierarchical Group (EHG)

Hidden Markov Modeling based on multiple samples provides a better representation of the trajectory data. However, this is a "chicken-and-egg" problem.

1). Space measurements: calculate distances between two groups *i* and *j* in the dataset.
2). Reclassifying : *mi* and *mj* are replaced by U ; then based on the *N*-1 HMMs, all the data are classified into *N*-1 groups by the maximum likelihood criterion;
3). Retraining : the *N*-1 HMMs are retrained based on the updated *N*-1 data groups;
4) Integration : the two groups *i* and *j* with smallest *dij* are integrated into one if the above criterion is satisfied.

3. Experimental Results

In this section, we show experimental results for our proposed methods and compare our results with other methods. We perform experiments based on the BEHAVE dataset [9]. Six long sequences are selected in our experiments with each sequence including 7000 to 11000 frames.

We try to detect seven group activities: Approach, WalkTogether, Split, Ignore, Chase, Fight, and RunTogether. The definitions of these seven activities are listed in Table I. We classify these seven activities into two categories with WalkTogether, Ignore, Fight, and RunTogether as normal activities, and Approach, Split and Chase as abnormal activities.

It should be noted that we extended the definition of activity Ignore. The two people will ignore each other if they do not have other activity correlation. Furthermore, Ignore will also be used to model the non interaction case between two normal groups. We also add a single activity into the normal activity list for those people that cannot be clustered into any normal group.

TABLE I - Normal Activities

Activity	Definition
WalkTogether	People walking together
Approach	Two people or groups with one approaching the other
RunTogether	The group is running together
Ignore	Ignoring of one another

TABLE III - Abnormal Activities

Activity	Definition
Fight	Two or more groups fighting
Split	Two or more people splitting from one another
Chase	One group chasing another

3.1 Cartridge selection

The video signal input can be receive through the following 3 ways:

1. From Local Hard drive
2. Live video url from internet.
3. Capture Devices (Web camera, TV tuner card etc..,)

3.2 Investigate cartridge

Avi media Library in .net Framework 2.0 is used. There are many inner classifications are available in avi format. Before extracting frames support for Tracking is fixed first.

3.3 Take out edges

Every video is converted as frames for object tracking. In live video internet urls there is no need to frame extraction. Because they are already available as Frames.

3.4 Track the items

Frames are like an image. Pixels are classified in an array. Horizontal and vertical Object matching is taken to track the variations in a pixels are identified. They are noted in a new array.

3.5 Rebuild the Frames with motion identifiers

Finally, based on a new array value Frames are constructed with Motion identifying red marks. From the frames new video is reconstructed.

3.6 Alarm & Sent Message

If the Abnormal event is detected, then the Alarm is set, at the same time the Alert Message is sent into nearest Police station.

4. CONCLUSION & FUTURE WORK

The Hidden Markov Model version of object line enables the measure of comparison between video events by cross likelihood but endure from the overfitting problem due to data shortage. We proposed in this paper a novel Energetic Hierarchical Group (EHG) approach, where the Hidden Markov Models are trained on many samples and the opening clustering errors caused by overfit are corrected in the iterative process and which is capable of improving the recognition accuracy. Experimental results demonstrate the effectiveness of our proposed algorithm. In the future work, we will explore the automatic switch mechanism to deal with the videos.

ACKNOWLEDGEMENT

I sincerely thanks to all authors in reference section. All papers in the reference section are very useful for my research. Their concepts, algorithms and techniques are very useful for my research.

REFERENCES

[1] Iveel Jargalsaikhan, Cem Direkoglu, Noel E. O'Connor1, Alan F. Smeaton, Kathy Clawson, Hao Li,Marcos Nieto, Aitor Rodriguez, Pedro Sanchez, Karina Villarroel Paniza, AnaMart´ınez Llorens ,"Interactive Surveillance Event Detection" , SAVASA Project@ TRECVID 2012.

[2] Weiyao Lin, Ming-Ting Sun, Radha Poovendran and Zhengyou Zhang, "Group Event Detection with a Varying Number of Group Members for Video Surveillance" IEEE Trans. Circuits System and Video Technology", VOL. 20, NO. 8, AUGUST 2010

[3] Wanqing Li, Member, Zhengyou Zhang, Fellow and Zicheng Liu, Senior Member, IEEE," Expandable Data-Driven Graphical Modeling of Human Actions Based on Salient Postures", IEEE Trans. Circuits System and Video Technology, VOL. 18, NO. 11, NOVEMBER 2008.

[4] Weiyao Lin, Ming-Ting Sun, Fellow, IEEE, Radha Poovendran, Senior Member, IEEE, and Zhengyou Zhang," Activity Recognition Using a Combination of Category Components and Local Models for Video Surveillance", IEEE Trans. Circuits System and Video Technology, VOL. 18, NO. 8, AUGUST 2008.

[5] Jack M. Wang, David J. Fleet and Aaron Hertzmann, Member, IEEE," Gaussian Process Dynamical Models for Human Motion", IEEE Trans. Pattern analysis and machine intelligence, VOL. 30, NO. 2, FEBRUARY 2008.

[6] Adeel Mumtaz, Emanuele Coviello, Gert. R. G. Lanckriet, Antoni B. Chan," Clustering Dynamic Textures with the Hierarchical EM Algorithm for Modeling Video", JOURNAL OF LATEX CLASS FILES, VOL. 6, NO. 1, JANUARY 2007.

[7] Christ, R.E.R.," Gaussian Hierarchical Bayesian Clustering Algorithm", Seventh International Conference on Intelligent Systems Design and Applications, ISDA 2007.

[8] Zicheng Liu, Senior Member, IEEE, Michael Cohen, Senior Member, IEEE, Deepti Bhatnagar, Ross Cutler, and Zhengyou Zhang,"Head-Size Equalization for Improved Visual Perception in Video Conferencing" , IEEE Trans. Multimedia, VOL. 9, NO. 7, NOVEMBER 2007.

[9] Dong Zhang, Student Member, Daniel Gatica-Perez, Samy Bengio, and Iain McCowan "Modeling Individual and Group Actions in Meetings with Layered HMMs", IEEE Trans. Multimedia, VOL. 8, NO. 3, JUNE 2006.
[10] R. Urtasun, S. Epfl, D. J. Fleet, and P. Fua, "3D people tracking with Gaussian process dynamical models," in Proc. IEEE Conf. Comput. Vis. Pattern Recognition., 2006, vol. 1, pp. 238–245.
[11] Dong Zhang, Student Member, Daniel Gatica-Perez, Samy Bengio, and Iain McCowan "Modeling Individual and Group Actions in Meetings with Layered HMMs", IEEE Trans. Multimedia, VOL. 8, NO. 3, JUNE 2006.
[12] Ying-li Tian, Arun Hampapur, Lisa Brown, Rogerio Feris, Max Lu, Andrew Senior, Chiao-fe Shu, and Yun Zhai," Event detection, query, and retrieval for video surveillance", Book chapter for Artificial Intelligence for Maximizing Content Based Image Retrieval.
[13] Chris Whiten, Robert Lagani`ere, Ehsan Fazl-Ersi, Feng Shi, "Interactive Surveillance Event Detection", VIVA Research Lab University of Ottawa, Ottawa, Canada.