INEF12Basketball Dataset and the group behavior recognition issue

Alberto Pozo, Jesús García and Miguel A. Patricio

Applied Artificial Intelligence Group, Universidad Carlos III de Madrid, Avd. de la Universidad Carlos III, 22, 28270, Colmenarejo, Madrid, Spain {alberto.pozo,jesus.garcia, miguelangel.patricio, }@uc3m.es
http://www.giaa.inf.uc3m.es

Abstract. Activity recognition is one of the most prolific fields of research. For this reason, there are new fields of research that expand the possibilities of the activity recognition: Group behavior recognition. This field does not limit the number of elements in the scene, and there are a lot of new elements that must be analyzed. Each group, like each individual element, has its behavior, but this behavior depends on their elements, and the relationships between these elements. All these new elements cause that group behavior recognition was a new field of research, with some similar elements but it must be studied apart. This way, group behavior recognition is a novel field, in which there are not many researches and there are not many datasets that could be used by researchers. This situation causes the slow advance of the science in this field. This paper tries to show a complete description of the problem domain, with all the possible variants, a formal description and show a novel architecture used to solve this issue. Also describes a specific group behavior recognition dataset, and shows how it could be used.

Keywords: Dataset, Group behavior recognition, activity representation, computer vision.

1   Introduction

Behavior recognition, with humans or with other kind of elements, is one of the most prolific fields of the current research. One of the typical restriction in these research is that must be only one element in the scene. The element's behavior is analyzed and the system recognizes the activity that the element is doing. There are a lot of papers with this objective, like (1) and (2).

The one-element restriction is valid for several situations, but in other cases, there are not isolated elements but these elements are joined in several groups. This could be useful in a number of situations, like group sports, animal behavior, coordinated human activities, etc. In these situation the group behavior is not only

the combination of the individual behaviors, however, there are a lot of elements that must be analyzed like the internal group relationships, external group relationships (with others groups) individual behaviors, groups hierarchy, etc.

So there are several commons elements between activity recognition and group behavior recognition, but there are also several new elements so this is a novel field of research that provides a lot of new challenges and difficulties. There is an essential element in any research that cannot be missing, datasets. All the experimentation is based on this element, and it is essential to create a new system, to improve it, and to prove its results.

This new field of research has some papers that manage the problem from very different points of view, like (3) or (4). However, there are not many specific dataset that could be used for these researches, and this is a serious difficulty for the researches. This paper describes a novel representation of this problem domain, showing all the different types of group behavior recognition problems and also describes a new specific dataset, which could be used to improve the current researches and to create new group behavior recognition systems. Also this dataset could be used to prove the existing techniques with a different dataset.

The next section describes related work. Section 3 describes problem domain. Section 4 describes our architecture approach and section 5 describes our dataset, which is followed by the conclusions.

2 Related work

As described above, there is not many dataset specializing in group behaviors for the scientific community. There is extensive literature dedicated to activity recognition, and there are a lot of dataset specializing in this area, which there is only one studied element, like KTH dataset (5) or Weizman (6); which are used in many papers like (7), (2), (8). But these dataset cannot be used to prove the group behavior recognition techniques, so we need a specializing group behavior dataset that provides the behavior of a group, with some raw information like positioning, individual actions, etc.

This is the case of the Nanez Pers work, where he has launched a three parts dataset. Each part related with one sport: Basketball, squash and European handball. Nanez Pers in (9) presents a dataset specializing in computer vision, and one of the three parts (European handball part) is focus on the group behavior recognition.

In this dataset, there is a lot of valuable information for the group behavior recognition issue. CVBASE06 (9) has ten minutes data from a European handball mach, with one team players tracking. So there are annotated all positions of each player (from one team) in each frame of video. Furthermore, CVBASE06 gather information about what are doing every player (of the tracked team) like throwing the ball, for example, and the team performed action, like aggressive attack. All this information has been annotated by an expert (European handball coach) by

several views of the video. So this dataset combines both types of information, individual data like positioning and action, and group data like action performed. So CVBASE06 could be used to group behavior recognition issue, however, there are some others aspect that could be improve.

Moreover, there is scientific and educational international joint project called RoboCup (10) that could be used to test and improve the group behavior recognition systems. This competition has many different leagues divided into five main clsasses: RoboCup Soccer, RoboCup@Home, RoboRescue, RoboCupJunior and Demostratios. Inside the Soccer league there is a client-server based architecture simulator that could be gather information useful for the group behavior recognition issue.

On the other hand we could read the works from Ruonan Li and Rama Chellappa (3), (11), and Behjat Siddiquie (12) where they uses an American football dataset called GaTech Football Play. Unfortunately, there are no references to use this dataset by the other researches.

3 Group behavior recognition issue: a general overview

Group behavior recognition is a novel field of research that comes from the elimination of the one-element restriction in activity recognition issue. This field of research has a lot of potential domains like group sports, military intelligence, fauna behavior recognition, video surveillance, etc.

Group behavior recognition is composed by two steps: in the first one the features of the system should be extracted, and in the second one the features are used to recognize the behavior. The system could have a lot of types of features like position, individual action, trajectory, speed, color, etc.

In this paper we are going to focus on the second step, there is only a short description about how was the dataset construction process (feature extraction), and we try to show how this information could be used in the second step (behavior recognition).

3.1 General Description

There are a lot of types of domains and problem where we could use the group behavior recognition, for this reason it is important to define a general description that includes all of them.

In general, we have several elements that are forming groups; these elements are moving around the scene and have relationship with other elements of their group and with the other group. Also, groups have relationship with other groups. So we have a general scene where exist a number of elements (which could be fixed or not). Each element of the system has a set of features (like positioning, color, shape, etc.) The element's features could suffer changes in time. Each element of the system should belong to a group, and could belong to many groups at

the same time. It is important to emphasize that any element of the system must be in a group or not, so there could be isolated elements.

Each group has an internal and an external attitude. Each attitude could be cooperative or competitive. Internal attitude defines the attitude between the members of a group, and external attitude defines the attitude of the group respecting of the rest of the groups.

Different domains could have some different descriptions, in general, there are six types of problems that's could be classify by three main features: If there are one or more groups, if there are one or more behavior on each sequence, and if the number of groups is fixed or is not.

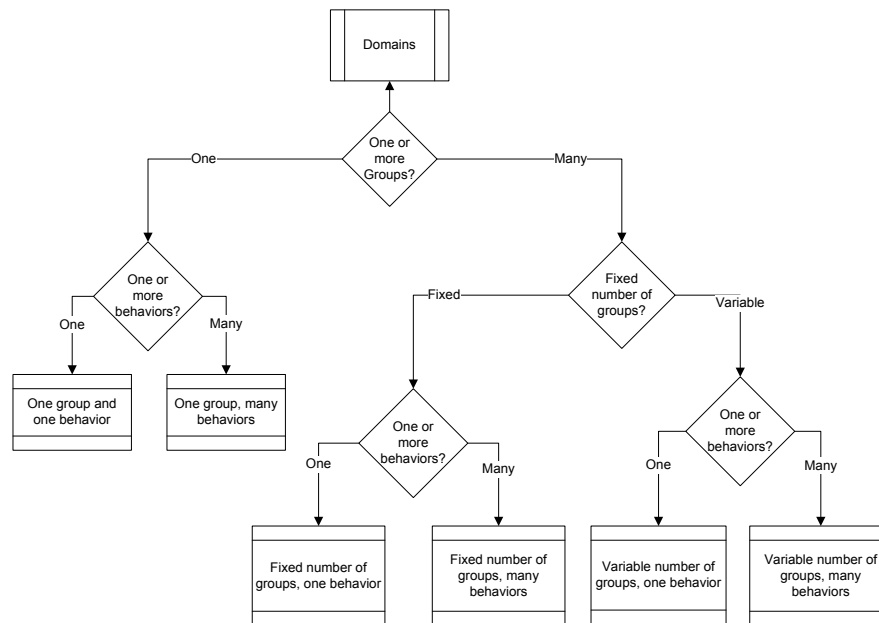The following image shows this classification.

Fig. 1: Problem classification

3.2 Problem description

In the first level the features extractor system analyze the scene and gather all the relevant information to the group behavior recognition system. So in the second level we have one sequence composed by a number of T instants, where are included a number of N elements (this number cannot change in time). The elements of the scene are distributed in a number of G groups. Each element of the scene has a set of features that depends on the specific problem domain. Some of these features could be logical (like boolean), numeric, text, etc. But in the system must be some feature that indicates the location of the element in some way.

This is because we are focused in the problems where the location and the trajectories of the elements are important, but it is important to remark that this restriction is carry out by most of the problem's domains.

The elements include in some group could not change, and the number of groups and elements could be fixed or could be not. Moreover, all the elements must be in a group, so there is not any isolated element.

So for each moment t, we have the scene information $_t = \{G_t^1, G_t^2, ..., G_t^m, ..., G_t^M\}, 1 \leq m \leq M$ (where M is the group amount) and $G_t^m = \{I_t^1, I_t^2, ..., I_t^n, ..., I_t^{Nm},\}, 1 < n \leq N_m$ (where N_m is the number of elements of the group m), and $I_t^n = \{f_t^1, f_t^2, ..., f_t^c, ..., f_t^C ,\}, 1 \leq c \leq C$ (where C is the features amount).

There are four types of features describe above: *Positive, boolean, relative* and *enum*.

Boolean: this type of feature could be one or zero, **positive**: feature could have some value between zero and infinite, **relative**: could have some value between zero and one, and indicates how much the element fulfill with something, and **enum**: this type of features could be one of a list of possible values determinates by the problem domain.

4 Architecture

Group behavior recognition is a general research field that includes a lot of problem domains like was described above. The problem could be divided in two steps, first level must extract the features of the elements and second level must use this features to recognize the behavior of the groups.

There are two different situation depending on the number of behaviors per sequence, if there is only one, the system only have to recognize it, but if there are many behaviors the system have to cut the sequence before recognize them.

The architecture shows below tries to manage with the second situation, and it is a general approach that could be implemented with many different specific algorithms.
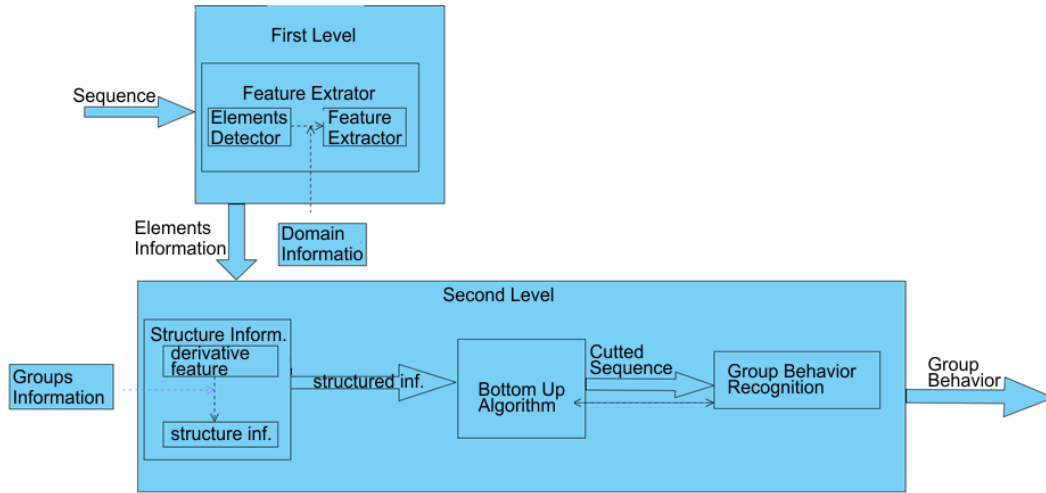
Fig. 2: Architecture

This architecture approach manages the problem of the many behaviors on one sequence with an algorithm to cut the primitive sequence, like buttom up.

The makeup group information is obtained by a external element of the system, in most case an expert.

5 INEF12Basketball Dataset.

Within the field of artificial intelligence, especially in the field of activity recognition, there are plenty of papers from different authors trying to create autonomous systems that were be able to detect, classify, or simulate the insolate individual behavior.

All these papers have a lot of experiments based on a several activity recognition datasets, which allow this research field advances quickly. However, there is a novel field that emerges from the elimination of one restriction of the domain: It could be more than one element, joined in a group. This new field of research has some resemblances but also have a lot of differences so it must be manage with new approaches and techniques.

This new field of research is focus on a group of elements, these elements have a lot of relationship with others elements of the group, and with others elements of the other groups. So the number of the elements and relationships increase exponentially. This is very important, because some of the techniques use to activity recognition could be not feasible in time.

This new field still has not many researches compared to its predecessor, and it is the same with the number of specific datasets. INEF12Basketball dataset aims

to contribute to this field of research, providing a set of low-level data such as location of the system's elements and other higher-level data related to the individuals and groups behavior.

So we try to create a useful dataset that could be used to improve and prove the novel techniques developed for group behavior recognition issue.

5.1 Domain

All the information of the dataset was gathered in a basketball training session, carried out by four players (two versus two). This session was recorded in 225 seconds of video composed by twenty seven sequences. Each sequence went to one basketball move.

All the moves played in the scene are composed by one pick and roll attack and some of these ways of defense: to fight over, to go below, to help and recover, to show and go, to show and recover, to switch and trap.

Group behavior information is related with the type of defense played, so all the moves have the same attack and the behavior information is focused on the defense. In a pick and roll move, the attacking team try to set a screen (pick) for a teammate handling the ball and then slips behind the defender (rolls) to accept a pass. It could be defended by several types of defenses; here we can read a short description of each of them.

Fight over: In this case, the blocked player trays to skip the screen fighting over them.

Go below: In this case, the blocked player trays to skip the screen going below, this take some extra time.

Help and recover: The player whose defense to the player that plays the screen tries to help to his partner with a short lateral movement.

Show and go: In this case the second defender goes with the ball for some time, and he gives some help to his partner.

Show and recover: This case is similar to previous one, but in this case the help is longer in time.

Switch: Both defenders change their marks.

Trap: Both defenders go to mark to the owner of the ball, in a 2vs1 situation.

So all the information is gathered in twenty seven consecutively sequences of basketball training session, with one type of attack and seven different types of defenses.

5.2 Description

INEF12Basketball dataset is composed by a video data recorded by four cameras in fixed positions on a basketball court.
There are 27 plays recorded, in three minutes and 45 seconds. At 25 frames per second the video has 5.627 frames.
Video was recorded in MPEG4 DivX avi container and no audio associated. The size of each frame is 320 pixels wide and 240 pixels high. Apart from the video, CVBASE12Basketball dataset provides useful information about the scene; like player location, group action, owner of the ball, etc. Player's location is annotated

around the whole video. This information is presented in one plane text file, in two coordinated axes: screen and court. This information was captured using a modified version of ViPER program. The tracking was performed by a tracker algorithm with expert supervision, all this information is saved in the plane text file.

As well as the position, dataset contains information about which player has the ball in each frame. (There are frames in which the ball is not had by any player).

This information was annotated by an expert. The file contains the frame in which the player has the ball, and the frame in which the player louse the ball. Moreover, in order to serve as a useful research fields related to the group behavior recognition, segmentation or classification; group activity information has been incorporated.

There are information about witch type of defense is played in each turn, in the section there are the explanation of each type of defense.

Coach called Ignacio Refoyo was the expert that provides all the technical data, like defense played by the players.

All this information and much more could be found in (13).

5.3 Using the dataset

All the information of INEF12Basketball dataset described above is stored in files (.avi and .txt) but could be managing by a few set of Matlab functions described below. These functions allow to the users manage all the information without knowing the format and the structure of the txt files. And provide a quick and easy way to start working on the algorithms designed, taking from the outset all the information available in the Matlab workbench.

There are seven Matlab functions:

1. CVBASE12BasketballInit
2. CVBASE12BasketballGetPos
3. CVBASE12BasketballShowTrajectories
4. CVBASE12BasketballGetFrames
5. CVBASE12BasketballShowFrames
6. CVBASE12BasketballGetTeamActivity
7. CVBASE12BasketballHasBall

The first one is used to initialize all the system, loading the information on the workbench. Functions 2, 3; 6 and 7 allow the users manage all the information about the location of the players, the defense played and the owner of the ball respectively.

The following image shows the information presented to the user with the function 3, that print the trajectories of the players in the couch.
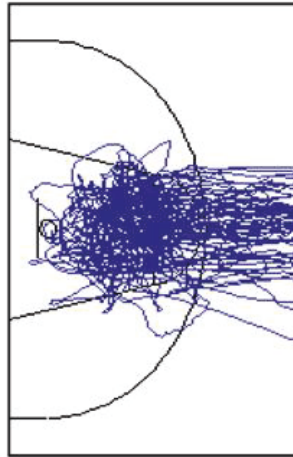
Fig. 3: Show trajectories

At last, functions 4 and 5 allow the users manage the raw video information, 4 obtains the chosen frame information provided by different cameras and the time of the frame in several variables, and 5 show the video information as we could see below.

Fig. 4: ShowFrames

6 Conclusions

Group behavior recognition is a field of research that emerges as an extension of the activity recognition, when the one-element restriction is eliminated. The elements and the features in group behavior recognition have several differences and it must be manage with different techniques and approaches.

So we have a novel field of research, which is still a long way to go, and it needs a lot of work to have a good approaches. One of the current difficulties on this field is that there are not many dataset to create the approaches and improve and prove it with experiments. So we need more dataset specializing in group behavior recognition issue.

This paper describes a new dataset that aims to solve this problem, using the information gather in a basketball training session.

Because in the most of the group behavior domains the location of the elements is a important feature to recognize the group behavior, INEF12Basketball dataset includes this data (the location of each player in each frame), and some other information like the current owner of the ball, and, of course, the group activity carried out by the group.

In addition, the dataset includes a set of Matlab functions that allow to the users handle the data in a quickly and easy way. So researchers only need to know some about Matlab language and could have all the gathered information in their hands.

References

1. Liu, J, Luo, J and Shah, M. *Recognizing realistic actions from videos in the wild.* 2009.

2. Schindler, K and van Gool, L. *Action snippets: How many frames does human action recognition require?* 2008.

3. Ruonan, Li, Rama, Chellappa and Shaohua, Kevin Zhou. *Learning Multimodal Densities on Discriminative Temporal Interaction Manifold for Group Activity Recognition.* New York : s.n., 2009.

4. Ramos, Fernando and Ayanegui, Huberto. Tracking behaviours of cooperative robots within multi-agent domains. Tlaxcala : s.n., 2010.

5. Schüldt, Cristian, Laptev, Ivan and Caputo, Barbara. *Recognizing human actions: a local SVM approach.* 2004. Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. pp. 32 - 36 Vol.3.

6. Gorelick, Lena, Blank, Moshe and Shechtman, Eli. A*ctions as Space-Time Shapes.* 2005. Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. pp. 1395 -1402 Vol. 2.

7. Jhuang, H, et al. s.l. *A biologically inspired system for action recognition.* : Ieee, 2007. Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on. pp. 1-8.

8. Niebles, J C and Fei-Fei, L. s.l. *A hierarchical model of shape and appearance for human action classification.* : IEEE, 2007. Vols. Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, pp. 1-8.

9. Pers, Janez. *CVBASE 06 Dataset: A Dataset for Development and Testing of Computer Vision Based Methods in Sport Environments.* Ljubljana : s.n., 2005.

10. Asada, M and Kitano, H. *RoboCup...: robot soccer World Cup...* s.l. : Springer Verlag, 1999.

11. Li, Ruonan and Chellappa, Rama. *Group Motion Segmentation Using a Spatio-Temporal Driving Force Model.* Maryland : s.n., 2010.

12. Siddiquie, Behjat, Yacoob, Yaser and Davis, Larry S. *Recognizing Plays in American Football Videos.*

13. Refoyo, Ignacio, et al. *INEF12Basketball Dataset.* Madrid : s.n., 2011.

14. Schindler, K and Van Gool, L. s.l. *Action snippets: How many frames does human action recognition require?* : IEEE, 2008. Vols. Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, pp. 1-8.