Entertainment Computing 27 (2018) 179–187

Contents lists available at ScienceDirect

Entertainment Computing

journal homepage: www.elsevier.com/locate/entcom

Generating background NPCs motion and grouping behavior based on real video sequences

Paulo Knob, Marlon Alcântara, Estêvão Testa, Rodolfo Favaretto, Gabriel Lima, Leandro Dihl, Soraia Raupp Musse*

Graduate Course in Computer Science, School of Technology, Pontifical Catholic University of Rio Grande do Sul, Brazil

A R T I C L E I N F O

Keywords:
Background population
Human trajectories
Computer vision
Group behavior

A B S T R A C T

Nowadays, there are many different types of NPCs (Non-Playable Character) in games, which can be endowed with different abilities, levels of intelligence and more or less realistic motion. How much the NPCs abilities are sophisticated, more complex are techniques to model and control them. Concerning the realism of NPCs motion, the ideal result is to provide characters that act and move in a way that seems natural to people. However, it can take a long time and a lot of work for developers, since designers should define characteristics for each different background character. This work aims to introduce a trajectory learning system derived from real video sequences, which can be used to generate NPCs motion for games. In addition to NPCs trajectories, we propose to detect in video sequences the formed groups and its characteristics to be used to simulate more realistic NPCs. We performed some evaluations using a pre-defined set of real video sequences and compared them with generated simulations. The obtained results indicate that our motion learning and groups detection techniques can be a valuable resource in order to generate NPCs and groups of NPCs trajectories, in an automatic way and similar to the real-life.

1. Introduction

Despite the breakthrough in gaming technologies, the game industry still seeks to reflect the real world in a better way. One relevant subject, in particular, is the role of the Non-Playable Characters (NPCs), where most of them have their behavior and motion manually defined by game designers, normally using decision trees in a limited way. As Warpefelt [29] comments in his work, "*Many of the problems with NPCs stem from the fact that they do not achieve a sufficient level of believability, particularly in the social arena. This is primarily related to the fact that the NPCs do not behave in ways that align with the expectations of the player. This can lead to the player misunderstanding the role and purpose of the NPCs, which damages the believability of the game. By extension, this lessens the enjoyment the player can derive from the game*". Since it restricts the possibilities, it can generate unrealistic motion or behavior, which in turn can make the players bored or uneasy. Also, game designers spend a significant amount of time and resources on this task, which could be used to work in more important matters.

The main goal of this paper is to propose a way to generate NPCs and groups of NPCs behaviors based on learning trajectories from real video sequences. While the process to generate trajectories use tracked motion from video sequences at each frame, groups are also identified in the videos, as well some of their characteristics. Examples of computed features are speeds, orientations and distances among people. Then, captured and computed data are used to simulate crowds based on video sequences, having both individuals and groups. It is not the first time computer vision and games are related. As presented by Freeman et al. [9], "*the appeal of computer games may be enhanced by vision-based user inputs*".

Fink et al. [8] also explore the use of computer vision in games. In their work, authors make use of computer vision techniques to extract information from a played game. In other words, they use the graphical output of a given game (during the game play) to learn the behavior of NPCs and other entities, using the method presented.

Although some works were found in the area, the use of computer vision to create NPCs motion is still incipient. Section 2 discusses some related work. The main contribution of this paper is to use captured data from video sequences, using computer vision, to generate individual and grouped trajectories of NPCs for games, in an automatic way. Computer vision methods are used to detect and track individuals and groups of people. Indeed, we propose two different ways of control: the Continuous Based Control (*CC*) which aims to continue a specific video sequence keeping NPCs and groups behaving exactly like in the video, and the User Based Control (*UC*), which allows the user to inform

* Corresponding author.
E-mail address: soraia.musse@pucrs.br (S.R. Musse).

https://doi.org/10.1016/j.entcom.2018.06.003
Received 1 March 2018; Received in revised form 10 May 2018; Accepted 17 June 2018
Available online 25 June 2018
1875-9521/

the number of groups and agents to be created in the simulation/application.

The remaining of this paper is organized as follows: Section 3 presents the method we adopt to detect and track people in video sequences and describes our proposal to learn and estimate new trajectories for NPCs, as well our proposal to identify groups of people and compute groups parameters. Then, we developed an Unity application that allows to read a short video and provides simulation using the two possible controls: *CC* or *UC*, based on the learning trajectories. This application allows to visualize the whole film, partially from real life, partially computer-generated. Such aspects are discussed in Section 4. Lastly, Section 5 presents our final considerations and future work.

2. Related work

This section presents some related work where authors discuss the importance of NPCs believability in the games success. Moon et al. [21] affirms that the majority of the NPCs have a limited range of actions contained in the game parameters. Therefore, this experience can be boring for the player and he/she may lose interest in the game in a rapid pace. From this, it is possible to say that it is of paramount importance to have methods able to generate and control NPCs behaviors in an automatic and effective way.

Because of this need to make NPCs actions more natural (or human-like), in order to increase games believability, the NPCs complexity is enhancing with time. Methods which apply Artificial Intelligence (AI) in the field of behavior control of the NPCs can be seen in Zhou et al. [31], Hussain and Vidaver [14] and Akbar et al. [1]. In order to control one or more NPCs in a dynamic way, Hussain and Vidaver [14] proposed a genetic algorithm based on robotic controllers. Akbar et al. [1] chose to make use of a Fuzzy technique to model team coordination. Therefore, one smart agent acts like the team leader in an attacking scenario. To make sure that each NPC has a more natural and unpredictable behavior, a variation is introduced with a Gaussian distribution.

In cases where NPCs are not a key component of the game (for example, a crowd which acts as a background population), Khoo et al. [16] says that is not necessary to use any AI method and scripted approaches are enough. However, such scripted behaviors can generate non-natural behaviors and can be seen as awkward for players. Some commercial games make use of some level of NPC's behavior control, like Planet Coaster [3][1] and Age of Empires [6].[2]

Therefore, the contribution of this work is to use computer vision techniques in order to extract people trajectories and use this data to control NPCs motion and their grouping structure, automatically. Our following-up approach also aims to identify groups of people in the video sequences, as well to be able to calculate some of their characteristics (speed, distances between people, etc.). In fact, Favaretto et al. [7] does a similar work.

Based on short video sequences, they calculate the mean speeds, distances between people and orientations, in order to find groups of people differences and relate that to cultural aspects. In the case of this work, we are more interested about mimic a specific video sequence with respect to (w.r.t.) the individuals behaviors in order to simulate and animate NPCs. Other works in the area try to mimic group behaviors as well. Using the ORCA method, He et al. [12] present an algorithm to simulate group behavior, similar with the observed behavior in real life. Such groups are dynamic, meaning each one can have any format and number of agents inside it. The concept used for ORCA of Velocity Obstacles (VO) is used in this work for collision avoidance purposes between groups. Even so, it is possible for a large group to divide itself when passing by another group, just to join again later. Our work differs from He et al. [12] specially in the group formation. While He et al. [12] uses a spatial clustering algorithm to assign a group for each agent (based on the positions and velocities of all agents), our work create groups based on the analysis of video sequences.

Krontiris et al. [18] also propose to add more realism to the crowd. More specifically, their work aims to develop an activity-centric framework for authoring functional, heterogeneous virtual crowds in semantically meaningful environments. The simulated environment is classified with attractors (for example: food, coffee or vendor) and each agent computes the so-called "influence maps", which is the influence of its surroundings on its behavior. Besides locations, it is possible for an agent to be influenced by other agents, which allows the framework to work with group behavior as well.

Courty and Corpetti [4] present a way to animate a crowd based on velocity fields estimated from a real-life video. Their work focus on the optical flow of the macroscopic crowd, defining the structure of the crowd motion and being more fit to be used with dense crowd of people. Also, their work is able to generate continuous crowd motion even with a short video sequence as input. Based on a criteria of similarity, they can create some transitions between given instants of the vector field, which generates a movement vector similar with the observed in the video. Using real-data video sequences, they run tests with two different scenarios: Strike Sequence (all people walking in the same direction) and Entrance Sequence (people entering a stadium). The results achieved show that their method was able to reproduce the real scenes with success. Nevertheless, some limitations were found. The quality of the simulation is linked with the initial density of the crowd and agents are not able to avoid collision.

Kim et al. [17] combine online tracking algorithms, non-linear pedestrian motion models and machine learning techniques to create an interactive approach for analyzing crowd videos and generate content for multimedia applications. The authors claim that their method can be used for three different ends. First, it can be used to segment the movement of the crowd, gathering similar behaviors in different clusters. Second, it can be used to detect anomaly behaviors in the crowd, finding a few of them that are rarely observed throughout the video sequence. Third, it can be used for crowd replication, using the observed behaviors as input to generate virtual agents with similar trajectory behaviors as people in a given real-life video. Although the tests show that the method works as intended, some limitations are pointed. The chosen tracking algorithm is vital for the performance and accuracy of their method, since it can generate noises or lose track of people. Also, their online learning algorithm can only find local data of pedestrians, while an offline learning method could be able to compute many global characteristics of them.

Solmaz et al. [26] aim to identify five different crowd behaviors in crowd scenes, namely bottlenecks, fountainheads, lanes, arches, and blocking. Using the optical flow definition, they make use of linear dynamical systems and Jacobian matrix to analyze the crowd flow in a given video sequence. Then, local peaks are clustered and regions of interest (ROI) can be found and checked for any abnormal behavior. Results achieved show the flexibility and capability of the method for different scenarios. Although, the authors list some limitations, like the problems faced when the video sequences do not have a consistent flow.

Our main contribution is to propose a method where NPCs are controlled in terms of their trajectories as well as their grouping behavior based on video sequences. In fact, the use of video sequences to help the crowd simulation is not new, and many works were already proposed in this area Lee et al. [19], Lerner et al. [20], Paravisi et al. [23], Hu et al. [13], Pettré et al. [24], Junior et al. [15]. When trajectories of individuals are learned from a video sequence and used to determine simulated agents behavior, the detection of the crowd characteristics is usually called as data driven crowd simulation. The difference in this paper is the way to control such behaviors and the group structure learning and simulation, based on video sequence.

[1] Available at https://www.planetcoaster.com/en-gb, visited at 05-07-2018.
[2] Available at https://www.ageofempires.com/, visited at 05-07-20018.