Classifying Pedestrian Behaviour using Random Forests

Panayiotis Charalambous*and Yiorgos Chrysanthou†
Department of Computer Science, University of Cyprus

1 Introduction

Crowds of pedestrians are an important feature of both virtual and real worlds. Over the past few years, various scientific fields have been concerned with studying crowds. Some, such as Computer Vision, study them in attempt to understand the intents of real people captured by sensors such as cameras. Others, such as Computer Graphics, in attempt to convey the desired intend in the simulation of virtual people. Being able to extract higher level understanding of crowds from their trajectories can have a wide array of applications such as crowd synthesis, crowd evaluation and outlier detection. Having a method for automatic labeling of crowd trajectories (or parts of them) with high level descriptions such as walking, standing still, talking or as being part of a group, real and virtual crowds can be better understood. As far as crowd synthesis is concerned, trajectories from large databases could be extracted so that they exhibit a particular behaviour pattern and then used to train a data driven crowd simulator resulting in more believable crowds. Typical methods to evaluate a virtual crowd consist of quantitative approaches such as gathering statistics from the simulated trajectories and getting scores for various tasks such as time to reaching goals, minimum time to collision, average speeds etc. or qualitative approaches that employ typically user evaluations. By comparing higher level characteristics and behaviours of a simulated crowd to a real world one, better evaluation algorithms can be developed.

We propose a set of features that help in the classification of trajectories in high level behaviours based on local and global statistical measures. This selection of local and global features is capable of capturing behaviours that only one the two set of features would fail to do due to both local interactions between pedestrians and their global state.

2 Methodology

Preprocessing In the proposed classification scheme, users are requested to tag tracked people from video sources into six categories: *walk, wander, walk fast, wait, meet and talk and group*. Even though some of the classes are very similar having subtle differences (such as for example the walk and wander classes), a person has the intrinsic ability to differentiate between them. By allowing a person to tag these trajectories, a classifier can be trained to learn what are the decision boundaries used by people subconsciously to differentiate between different behaviours. In a sense, the quantitative statistical measures over the trajectories are correlated to the qualitative measures that real people gain from experience of being part of and observing crowds in their everyday lives.

After the tagging phase, each trajectory is split into smaller fixed time overlapping segments. For each one of these segments, a feature vector consisting of 12 statistical measures were calculated (Table 1). These measures are split into two categories: history and local. History statistics represent the entire history of the pedestrian up to the end of the segment and in essence encode past behaviour whereas the local statistics deal with local interactions. This scheme was chosen since the behaviour of a person is consistent on one hand with his past behaviour and on the other hand on the the local interactions/actions he takes at any given time. By taking into account both local and history statistics we have found that the classification accuracy increased significantly. The generated feature vectors then act as examples of a particular behaviour which are used to train a Random Forest classifier.

	Statistic	$E(.)$	$\sigma(.)$	Description
History	$U(t)$	✓	✓	Speed
	$w(t)$	✓	✓	Rotational Speed
	$nn_d(t)$	✓	✓	Distance to nearest neighbor
Segment	$U(t)$	✓	✓	Speed
	$w(t)$	✓	✓	Rotational Speed
	$nn_{dist}(t)$	✓	✓	Distance to nearest neighbor

Table 1: *The statistics used in this work consist of 12 distinct scalar values for speed, rotational speed and distance to nearest neighbor.*

Experiments To test and evaluate our methodology, pedestrians from a crowd on a busy street were tracked and each trajectory was tagged by a user to belong in one of the six possible behaviour classes. These trajectories were then segmented into smaller overlapping segments of 4 seconds length which were then used to train the classifier. Out method was used to tag trajectory segments from both video datasets ([Lee et al. 2007; Lerner et al. 2007; Pellegrini et al. 2010]) and simulation data ([Reynolds 1999]) with promising results. We are in the process of manually tagging the data by a number of users to verify classification accuracy.

3 Discussion

We see a lot of potential applications of the proposed method such as *Motion Mining* for data driven crowd synthesis and scenario editing. *Evaluation* for crowd simulations can benefit from the higher level understanding of the simulated data such as a distribution of groups or social interactions and not just numerical data or user evaluations. Additionally *video tracking* and *surveillance* could also benefit. As feature work we are considering on looking into additional statistical measures and multiple tags to allow for behaviour combinations.

References

BREIMAN, L. 2001. Random forests. *Machine Learning 45*, 1, 5–32.

LEE, K., CHOI, M., HONG, Q., AND LEE, J. 2007. Group behavior from video: a data-driven approach to crowd simulation. *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*, 109–118.

LERNER, A., CHRYSANTHOU, Y., AND LISCHINSKI, D. 2007. Crowds by Example. *Computer Graphics Forum 26*, 3, 655–664.

PELLEGRINI, S., ESS, A., AND VAN GOOL, L. 2010. Improving data association by joint modeling of pedestrian trajectories and groupings. *Computer VisionECCV 2010*, 452–465.

REYNOLDS, C. 1999. Steering behaviors for autonomous characters. *Game Developers Conference 1999*.

*e-mail:totis@cs.ucy.ac.cy
†e-mail:yiorgos@cs.ucy.ac.cy