
Simulation of Crowd Problems for Computer Vision

Ernesto L. Andrade, Robert B. Fisher

School of Informatics

University of Edinburgh, JCMB, The King's Buildings, Mayfield Road, Edinburgh, EH9 3JZ – UK

Telephone: +44 131 651 3441, Fax: +44 131 651 3435

E-mail: {eaneto,rbf}@inf.ed.ac.uk

Abstract

This work presents an approach for generating video evidence of dangerous situations in crowded scenes. The scenarios of interest are those with high safety risk such as blocked exit, collapse of a person in the crowd, and escape panic. Real visual evidence for these scenarios is rare or unsafe to reproduce in a controllable way. Thus there is a need for simulation to allow training and validation of computer vision systems applied to crowd monitoring. The results shown here demonstrate how to simulate the most important aspects of crowds for performance analysis of computer based video surveillance systems.

Keywords

Surveillance, simulation, computer vision, behaviour, crowd problems, social force model

1. Introduction

The problem of understanding and defining what a human crowd is and how it behaves can be analysed at different levels. For instance, crowds are represented as an aggregate of individuals having a set of motivations and basic rules [1-5]. This representation permits a flexible analysis of a large spectrum of crowd densities and complicated behaviours. Continuous crowds (high densities) are also represented as fluids and simulations from fluid mechanics are applied to represent the crowd flow. Another approach for crowd representation is that of a “thinking fluid” [6], working also for continuous crowds allowing an analytical model for the flow. The numerical simulation of crowds as a fluid for complex geometries or as discrete pedestrians to represent complicated behaviours is well established. Most of the work up to date addresses applications in building safety planning [2], crowd rendering for virtual world animation [1][21] or motion picture special effects [18]. However, in recent years computer vision systems are trying to model and interpret different sorts of human activities [7] including the behaviour of crowds [8]. Most recent computer vision systems for behaviour understanding rely on a statistical model. In order to build such models usually there is the need for large amounts of visual data with examples of the activities to be learnt. In other application domains such as gesture recognition [11], cargo bay activity [11-12] and office activity recognition [13] the reproducibility and complexity of the scene elements are simpler, whereas gathering interesting crowd behaviours in a surveillance

context is much more challenging. In the crowd surveillance domain there is the additional issue of legal restrictions for gathering statistically significant CCTV data from public surveillance systems. Given the practical complexities of gathering crowd video evidence and the need for repeatable and predictable ways to generate video evidence of crowd behaviour our work applies recent techniques of crowd simulation as discrete pedestrians to visual evidence synthesis. The visual evidence generation simulates CCTV video for computer vision modules performing image understanding. One of the applications of such video analysis is the detection of unusual events in the crowd. As an illustration one uses the simulated video output to collect long sequences of “normal” crowd behaviour for the training of statistical models and short sequences of “abnormalities” are generated to test whether or not the statistical model is able to correctly understand the test sequence as an unusual event in the crowd.

This paper is structured as follows. The second section discusses the modeling of events for the surveillance scenario and the context where the crowd simulations are applied. The third section presents the details of the simulation model for visual evidence generation. The fourth section shows results of applying the model for different dangerous crowd scenarios. The fifth section draws the final comments and conclusions.

2. Modelling Behaviour for Surveillance

Previous work on surveillance systems for crowd monitoring has been applied mainly to indoor environments such as underground stations [15]. They concentrate on detecting events in crowds such as overcrowding and congestion, detection of unusual and forbidden directions of motion, intrusion and detection of stationary people in forbidden areas. These systems use prior knowledge about the scene to i) set the limits for event detection and ii) model abnormal situations. Recently there has been a series of attempts to use computer vision to automatically learn and recognize human activities [16-17]. These systems attempt to understand visual data to derive a set of behaviours and activities, which can be used for recognition and classification. However, there has been no attempt to date of applying such techniques to model the activities and behaviours of a large group of individuals. Using such techniques for large groups of people, such as crowds, assumes that the normal behaviour of the crowd can be learnt from long sequences of observations. This behavioural model then can be used to evaluate the degree of abnormality for unusual or emergency events. In order to achieve this goal, reliable and reproducible data of people’s movements is necessary to test the modelling assumptions.

3. Simulation Model

The assumptions for the simulation are that pedestrians in the crowd react to their environment, other individuals, entrances/exits and obstacles, using a set of basic rules and local observations. This simplification is appropriate for simulating dangerous situations. If they were too smart and had knowledge of the whole world they would not put themselves in danger. One model of particular interest, which allows for these modelling assumptions, is the social force model [2-3]. It assumes that the interaction between pedestrians and their world is modelled by a force corresponding to the motivation of a pedestrian to move in a given direction at a certain speed. Forces representing the need to deviate from other pedestrians and obstacles counterbalance the motivational force. This simulation framework easily provides a compact representation for two of the main behavioural aspects of a pedestrian in a crowd:

following a given path and performing collision avoidance with obstacles. The motion behaviour derived from these two key aspects represents the crowd accurately enough in a CCTV surveillance domain.

After simulating a realistic pattern of motion for the crowd the next task is to render an image representation. For this the dimensions of the simulated world correspond to measurements of a real location. To position the world objects in the simulation the world is discretized to a grid of 1 cm^2 . To represent objects on the simulated world back into the real world, the ground plane of the simulated world is mapped on the real world image ground plane using a homography [10] transformation. The homography is computed with measurements from the real scene [19]. Overlaying image sprites representing pedestrians on the real scene image performs the final image rendering. The sprite area is scaled as a function of their relative depth in the image to represent their relative distance from the camera. The next subsections detail the simulation model for the pedestrian motion.

3.1. Pedestrian Path Model

The desired pedestrian path is modeled as a series of control points shown in figure 1. The beginning point defines the dimensions of the entrance area (E_{area}) for pedestrians. The intermediary way points for each pedestrian are obtained by adding uniform noise ($[U_{\text{path}_x}, U_{\text{path}_y}]$) to the original user specified way point coordinates resulting in a diverse set of way points per pedestrian. Each pedestrian attempts to follow its unique variant of the path. The ending point is located over a pre-defined exit representing the final goal for a pedestrian leaving the scene.

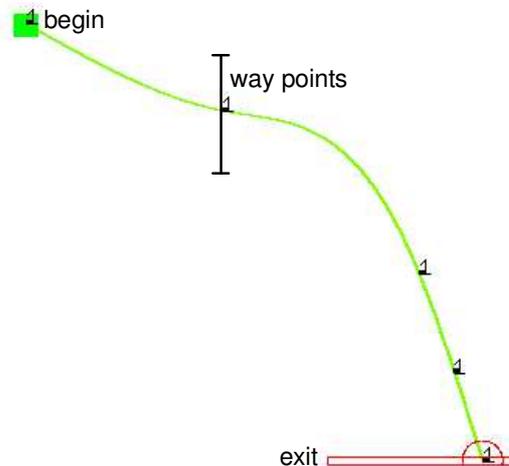


Figure 1: Model for the pedestrian path.

3.2. Pedestrian Body Model

The body model for the pedestrian on the ground plane is defined as a circle of area 900 cm^2 [5], which agrees with statistics of population size distribution and permits the simulation of interesting pedestrian densities, i.e. approaching 10 people per square meter in situations of panic and overcrowding. If the pedestrian were made too large there would be a self-imposed limit in the simulation allowing for an inbuilt safety factor in the pedestrian densities. The pedestrian body is considered uncompressible and the simulation does not take into account crushing or deformation effects in the crowd.

3.3. Pedestrian Motion Model

The simulation model for pedestrian motion is based on a modification of the social force model described in [2] and [3]. In the modified version the interaction forces are simplified to not consider friction effects and same group attraction, and to use hard potentials for the repulsive forces. The model has been augmented with the path model instead of using a single goal attraction potential, allowing a richer range of behaviors for the pedestrians that compose the simulation. The simulation is updated for all pedestrians at a desired frame rate. The change of velocities between frames is represented by the following acceleration equation:

$$m(i) \frac{d\mathbf{v}(i)}{dt} = \mathbf{F}_{\text{int}} + \mathbf{F}_{\text{CA}} \quad (1)$$

where,

- $m(i)$ mass of i -th pedestrian
- $\mathbf{v}(i)$ actual velocity of i -th pedestrian
- \mathbf{F}_{int} internal force on the i -th pedestrian
- \mathbf{F}_{CA} repulsion force among pedestrians for collision avoidance

The desired motion direction \mathbf{e}^d is given by,

$$\mathbf{e}^d(i) = \frac{\mathbf{W}_p(i) - \mathbf{P}(i)}{|\mathbf{W}_p(i) - \mathbf{P}(i)|} = \angle \theta \quad (2)$$

where,

- $\mathbf{W}_p(i)$ p -th point on the path that leads to the final destination of the i -th pedestrian
- $\mathbf{P}(i)$ current position of the i -th pedestrian
- θ is the motion direction angle towards the goal

The actual velocity $\mathbf{v}(i)$ is defined based on the current force interactions upon the i -th pedestrian. These forces are divided into i) internal forces, which measure the degree of motivation of a pedestrian to reach its goal and ii) repulsive forces acting among the i -th pedestrian the other pedestrians and obstacles. To simplify the evaluation of these forces in the pedestrian motion all pedestrian are assumed to have equal mass (1 Kg).

Internal forces:

$$\mathbf{F}_{\text{int}}(i) = F_{\text{mot}}(i) \mathbf{e}^d(i) \quad (3)$$

where,

- F_{mot} models the attraction for the current destination of the i -th pedestrian $F_{\text{mot}}(i) = m(i)v^d(i)/\tau(i)$, where $v^d(i)$ is the desired velocity and $\tau(i)$ is the adaptation time for the i -th pedestrian and $m(i)$ its mass.

Repulsive Forces:

$$\mathbf{F}_{\text{CA}}(i) = \sum_{j=1, j \neq i}^n -F_{\text{rep}}(\mathbf{P}(j), \mathbf{P}(i)) \frac{\mathbf{P}(j) - \mathbf{P}(i)}{|\mathbf{P}(j) - \mathbf{P}(i)|} + \mathbf{Avoid}(|\mathbf{P}(j) - \mathbf{P}(i)|) \quad (4)$$

$$F_{rep}(\mathbf{P}(j), \mathbf{P}(i)) = \begin{cases} |\mathbf{P}(j) - \mathbf{P}(i)|^2 \geq d_0^2 & F_{rep} = 0 \\ d_1^2 < |\mathbf{P}(j) - \mathbf{P}(i)|^2 \leq d_0^2 & F_{rep} = r_0 \\ d_2^2 < |\mathbf{P}(j) - \mathbf{P}(i)|^2 \leq d_1^2 & F_{rep} = r_1 \\ |\mathbf{P}(j) - \mathbf{P}(i)|^2 \leq d_2^2 & F_{rep} = r_2 \end{cases} \quad (5)$$

$$\mathbf{Avoid}(D) = \begin{cases} D > d_2 & \mathbf{Avoid} = [0,0]^T \\ D \leq d_2 & \mathbf{Avoid} = r_{av} \left[\frac{rnd_x}{\sqrt{rnd_x^2 + rnd_y^2}}, \frac{rnd_y}{\sqrt{rnd_x^2 + rnd_y^2}} \right]^T \end{cases} \quad (6)$$

where,

- n number of pedestrians plus obstacles in the collision avoidance radius
- $F_{rep}(\mathbf{P}(j), \mathbf{P}(i))$ the repulsion force between pedestrians j and i
- d_0, d_1, d_2 reference distances for the repulsion force intensity
- r_0, r_1, r_2 repulsion force intensities
- r_{av} repulsion force in case of collision
- rnd_x and rnd_y uniform random coordinates biased in the direction of the goal θ .

The biasing procedure is given by:

$$[rnd_x, rnd_y] = [\cos(\alpha), \sin(\alpha)] = \begin{cases} U[0,1] > \gamma & \alpha \in U[\theta - \pi/8, \theta + \pi/8] \\ U[0,1] < \gamma & \alpha \text{ is uniformly distributed over} \\ & \text{the other possible directions} \end{cases} \quad (7)$$

where,

- $U[x,y]$ is a uniform random distribution between x and y
- γ is the biasing factor (typical value $\gamma = .2$)
- α is the avoidance angle

In (7) the angle θ is the same as in (2) and the random deviations for collision avoidance are obtained relative to θ . The change in position $\mathbf{P}(i)$ is given by the velocity $\mathbf{v}(i) = d\mathbf{P}(i)/dt$. The desired velocities for the pedestrians are given by a Gaussian distribution of typical mean $v_{mean}^d = 1.5$ m/s and typical standard deviation $v_{\sigma}^d = 0.5$ m/s. The desired velocity distribution is truncated at one standard deviation. The maximum velocity is set to 3 m/s. The typical values for other simulation parameters are listed in Table 1.

Table 1: Simulation parameters.

Parameter	Additional Description	Typical Values
W_w	world ground plane width	14.80 m.
W_h	world ground plane height	7.50m
E_{area}	pedestrian entrance area	0.36 m ²
FR	frame rate	25 frames/second
PGR	pedestrian generation rate	0.1-0.001 ped./frame
PBR	pedestrian body radius	0.01692 m
d_0	repulsion outer distance	1.80 m
d_1	repulsion middle distance	1.00 m
d_2	repulsion inner distance	0.60 m
r_0	outer repulsion force	1.5 Kg*m/s ²
r_1	middle repulsion force	3 Kg*m/s ²
r_2	inner repulsion force	4.5 Kg*m/s ²
r_{av}	collision repulsion force	9 Kg*m/s ²

4. Simulation Results

The results for the crowd simulation are used to render a frame representing visual evidence of the crowd motion, see Figure 2. This figure shows a simulated scenario where two distinct groups enter the scene and leave through the same exit. The visual information from the simulation rendering is used to compute the crowd's optical flow. The computer vision modules use this information to infer the crowd's actions and behaviour.

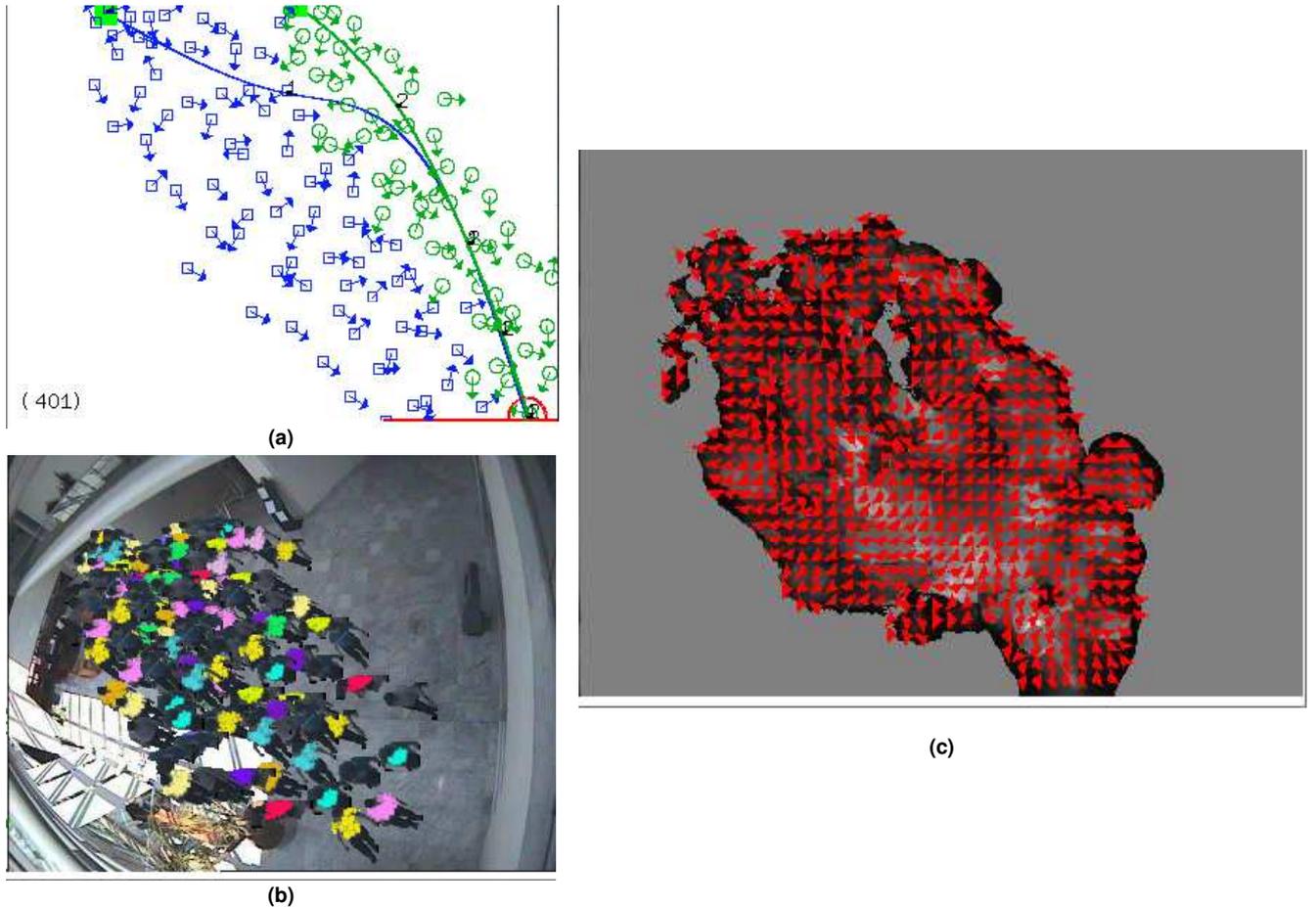


Figure 2: Simulation graphical output. a) Simulated world ground plane with two pedestrian groups following distinct paths. The mean paths are lines. Pedestrians on the given path have different shapes.

b) Rendering of the simulation on the image plane. c) Optical flow field computed for the rendered image frames, arrows indicate flow direction and lighter colours more intense flow.

Using the crowd simulations we are able to represent different emergency situations. To demonstrate this we show quantitative examples for two emergency scenarios: i) blocked exit at the lower right and ii) person collapse. In the blocked exit scenario the exit in figure 2 is blocked after 2000 simulated frames. Using the assumption that the people in the back of the crowd are unable to react to this situation and keeps traveling in their normal routes we observe the build up of the pedestrian density after the blockage. This situation is shown in Figure 3 and the visual data is ready for analysis by the computer vision modules.

The person collapse scenario is depicted in Figure 4, where a person collapses in the middle of the crowd, at frame 4000, and the other pedestrians start to strongly avoid running over that person. Figure 5 shows the mean delay for a pedestrian to cross the scene. The person collapse increases this mean delay due to the avoidance action taken by the remaining pedestrians. Figure 6 shows a qualitative assessment of the simulation results where we observe the formation of lanes given two opposite pedestrian flows. This is one of the self-organizing aspects of the crowd behaviour that emerge from social force models. We acknowledge that crowd simulations are hard to prove or validate. However like in [2] and [3] the experiments show that the quantitative effects are as one intuitively expects. For more qualitative results see the URLs: a) for the normal case <http://homepages.inf.ed.ac.uk/eaneto/normal.avi>; b) for the exit blocking scenario <http://homepages.inf.ed.ac.uk/eaneto/blocked.avi>; c) for the person collapse scenario <http://homepages.inf.ed.ac.uk/eaneto/fall.avi>; d) for the lane formation in crossed flows <http://homepages.inf.ed.ac.uk/eaneto/crossedflows.avi>.

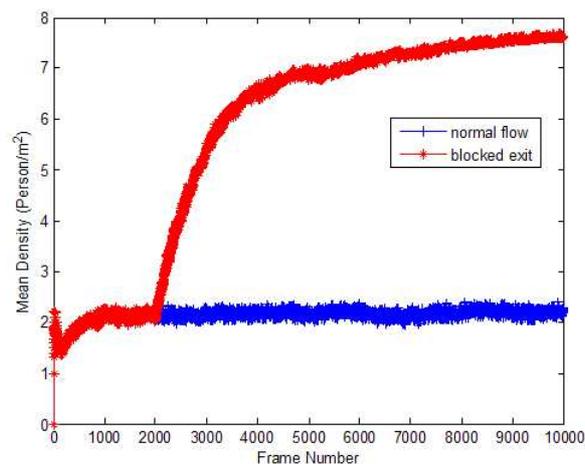


Figure 3: Person density for the blocked exit scenario. The exit is blocked at frame 2000.

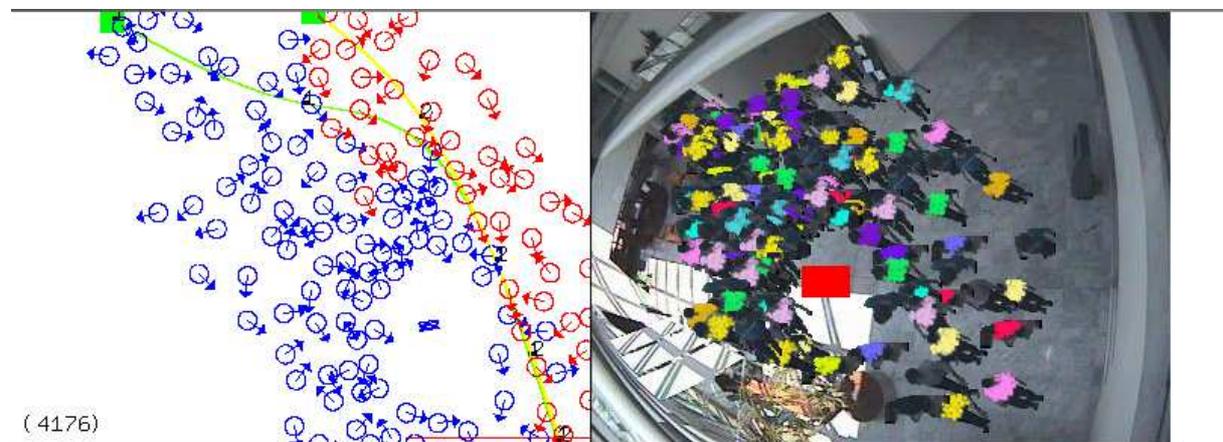


Figure 4: Illustration of person collapse scenario. Left, notice the hole in crowd caused by avoiding the collapsed person. Right, respective rendered image.

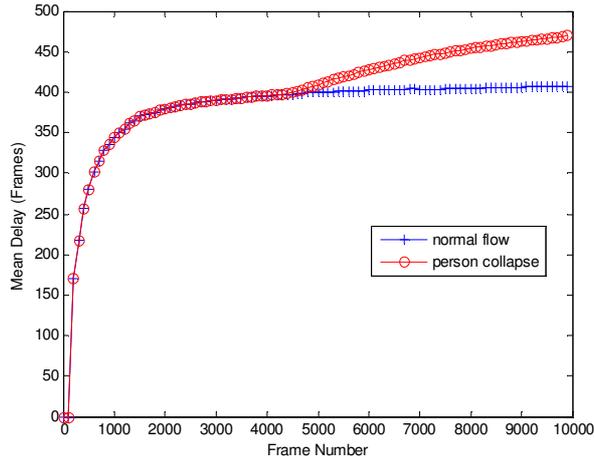


Figure 5: Mean delay to leave the scene for the person collapse scenario. The person collapses at frame 4000.

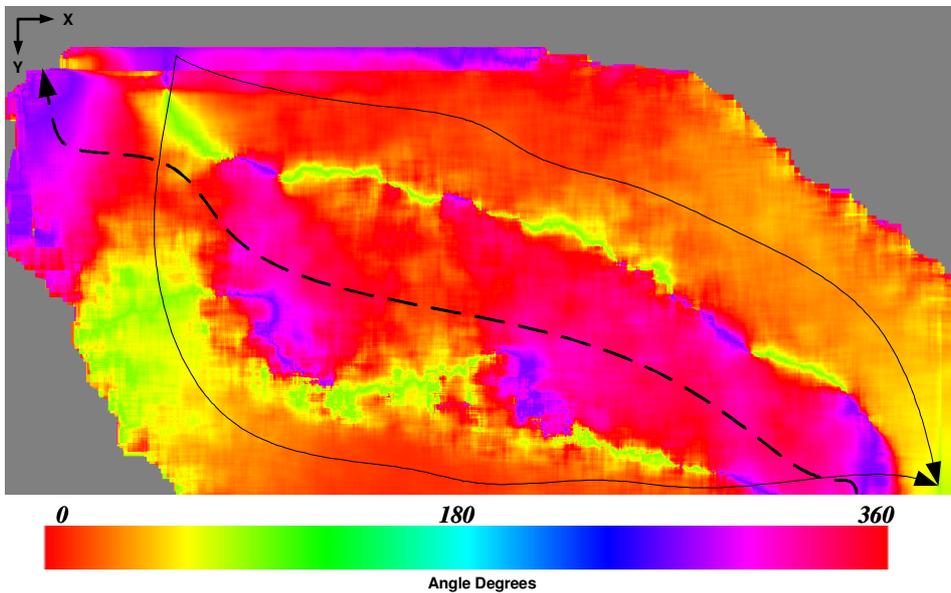


Figure 6: Example of lane formation given by the simplified social force model for two opposite flows. The image shows the mean values for the motion direction on the simulated ground plane after 9000 frames. The arrows indicate the main paths for the observed lanes. The colour (shade) encodes the average simulated motion direction at each point. Note the clearly distinct path regions and some turbulence where the paths are adjunct and at the crossing points.

5. Conclusions

This work presented a crowd simulation approach to generate visual evidence for training and testing of image understanding systems in computer vision. The main contribution is the generation of visual evidence for crowd emergency situations where such evidence is rare or hard to obtain. The results of the visual simulation are ready to be applied to computer vision systems that infer crowd behaviour from visual evidence, e.g. crowd density or optical flow. Although the simulation rendering has enough quality to train and test behaviour recognition in computer vision systems a better rendering strategy would allow for the representation of more subtle behavioral variations in the crowd. For instance visually representing with more accuracy the transitory states from a small group of people to a larger crowd, where details of individuals are more noticeable and not well represented by simple sprites. Such details would be better rendered by the projection of a pedestrian 3-D model at the cost of extra

computational expenses [9]. The simulation model can also be extended to incorporate more detailed behavioral model of the individuals in the crowd, for instance by applying the cost models for pedestrian motion in [20] to a dense crowd.

Acknowledgements

This research is supported by EPSRC BEHAVE project under grant GR/S98146.

References

1. Loscos, C., Marchal, D., Meyer, A., **Intuitive Behaviour in Dense Urban Environments using Local Laws**, IEEE Theory and Practice of Computer Graphics, p.p. 122-129, 2003.
2. Helbing, D., Farkas, I., Vicsek, T., **Simulating Dynamical Features of Escape Panic**, Nature, p.p. 487-490, 2000.
3. Helbing, D., Molnar, P., **Social Force Model for Pedestrian Dynamics**, Physical Review, p.p. 4282-4286, 1995.
4. Feurty, F., **Simulating the Collision Avoidance Behaviour of Pedestrians**, Msc. Thesis, Department of Electronic Engineering, University of Tokyo, 2000.
5. Still, K., **Crowd Dynamics**, PhD. Thesis, Mathematics Department University of Warwick, 2000.
6. Hughes, R. L., **The Flow of Human Crowds**, Annu. Rev. Fluid. Mech., p.p. 169-182, 2003.
7. Hu, W., Tan, T., Wang, L., Maybank, S., **A Survey on Visual Surveillance of Object Motion and Behaviours**, IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews, p.p. 334-352, August, 2004.
8. Boghossian, B. A., Velastin, S. A., **Motion Based Machine Vision Techniques for the Management of Large Crowds**, IEEE Int. Conf. on Electronics, Circuits and Systems, p.p. 961-964, 1999.
9. Ren, L., Patrick, A., Efros, A., Hodgins, J., Rehg, J., **A Data-Driven Approach to Quantifying Natural Human Motion**, ACM Transactions on Graphics (SIGGRAPH 2005), p.p. 1090-1097, 2005.
10. Hartley, R., Zisserman, A., **Multiple View Geometry in Computer Vision**, Cambridge University Press, 2000.
11. Galata, A., Johnson, N., Hogg, D. C., **Learning Variable Length Markov Models of Behaviours**, Computer Vision and Image Understanding, p.p. 398-413, 2001.
12. Hamid, R., Johnson, A., Batta, S., Bobick, A., Isbell, C., Coleman, G., **Detection and Explanation of Anomalous Activities: Representing Activities as Bags of Event n-Grams**, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p.p. 1031- 1038, 2005.
13. Xiang, T., Gong, S., **Discovering Bayesian Causality among Visual Events in a Complex Outdoor Scene**, IEEE International Conference on Advanced Video- and Signal-based Surveillance, p. p. 177-182, 2003.
14. Brand, M., Kettner, V., **Discovery and Segmentation of Activities in Video**, IEEE Transactions on Pattern Analysis and Machine Intelligence, p.p. 855-851, August, 2000.
15. Velastin S., Boghossian, B., Lo, B., Sun, J., Vicencio-Silva, M., **"PRISMATICA: Toward Ambient Intelligence in Public Transport Environments"**, IEEE

Transactions on Systems, Man, and Cybernetics - Part A, p. p. 164-182, January, 2005.

16. Ivanov, Y., Bobick, A., **Recognition of Visual Activities and Interactions by Stochastic Parsing**, IEEE Transactions on Pattern Analysis and Machine Intelligence, p.p. 852-872, August, 2000.
17. Oliver, N., Rosario, B., Pentland, A., **A Bayesian Computer Vision System for Modeling Human Interactions**, IEEE Transactions on Pattern Analysis and Machine Intelligence, p.p. 831-843, August, 2000.
18. <http://ai-implant.com/> (15-09-2005)
19. <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/> (15-09-2005)
20. Antonini G., Bierlaire M., Weber M., **Simulation of pedestrian behavior using a discrete choice model calibrated on actual motion data**, Swiss Transport Research Conference, 2004.
21. Tang W., Wan T., Patel, S., **Real-time crowd movement on large scale terrains**, Theory and Practice of Computer Graphics, p.p. 146-153, 2003.