An Analysis of Manoeuvring in Dense Crowds

S.A. Stüvel

M.F. de Goeij

A.F. van der Stappen

A. Egges

Technical Report UU-CS-2015-006
April 2015

Department of Information and Computing Sciences
Utrecht University, Utrecht, The Netherlands
www.cs.uu.nl

ISSN: 0924-3275

Department of Information and Computing Sciences
Utrecht University
P.O. Box 80.089
3508 TB Utrecht
The Netherlands

An Analysis of Manoeuvring in Dense Crowds

S.A. Stüvel, M.F. de Goeij, A.F. van der Stappen, A. Egges *

April 23, 2015

Abstract

While walking through a crowd, a person balances several desires, such as reaching some goal position, avoiding collisions with others, and conserving energy. Crowd models generally try to mimic this behaviour by planning short paths that avoid collisions. However, when the crowd density increases, choosing a collision-free path becomes more difficult. In such high-density crowds, one can observe torso twists; people rotate their upper body to decrease their width perpendicular to the motion path, in order to squeeze through narrow spaces between other crowd members. In this paper we investigate this behaviour, by recording and analysing dense crowds. We show that the paths chosen by the participants can be predicted by generalized Voronoi diagrams, and identify relations between instantaneous speed and look-ahead distance, and between the participants' torso orientations and goal positions.

1 Introduction

In real high-density crowds, people twist their torso to decrease their width perpendicular to their motion path, so they can squeeze through narrow spaces between other people in the crowd. Rather than *walking*, such motions are better described as *manoeuvring* through the crowd. In this paper, we describe our experiment, consisting of the recording and subsequent analysis of such crowd manoeuvring.

In our experiment, the participants form a crowd of such density that it is sparse enough to manoeuvre through, but of dense enough to require torso rotations in order to do so. In each trial, the person in the centre of the crowd must manoeuvre towards a predefined goal position, while the other participants remain more or less stationary. To reduce ambiguity in the analysis of the data, in each trial only a single participant will manoeuvre towards their goal position.

Main contribution We investigate properties of dense crowd manoeuvring. We show a relation between the orientation of the torso with respect to the goal position, and the lack of relation between angular and linear velocity. Furthermore, we show that in dense crowds people follow generalized Voronoi diagrams based on a line-segment representation of the agents.

Organization The rest of the paper is organized as follows. Section 2 discusses related work. The details and execution of the experiment are described in Section 3, with analysis of the results in Section 4. The implications and possible future work are discussed in Section 5, and conclude the article.

2 Related work

The study of crowd behaviour spans a large research area, ranging from computer vision techniques to assess human behaviour [ZL14] to the simulation of evacuation scenarios [ZZL09] and application in (serious) games. In this section we focus on the works relevant to the study of dense crowds.

Two possible ways of simultaneously capturing the motions of multiple individuals in a crowd are video recordings or motion capture systems. *Video recordings* are most often used. As these require no markers to be attached to the participants, the data are easier to obtain, and the resulting techniques can be used in a wider range of situations. A common approach is use feature point tracking, in order to determine movement of people in a crowd and classify behaviour [SBTM08, RSLA11]. Lee et al. use a similar tracking method to train a crowd simulation system, such that it exhibits behaviours imitating real human crowds [LCHL07]. The

*VHTLab, Universiteit Utrecht

opposite is also possible, by using a crowd simulation model to improve the result of feature tracking in videos [MOS09, BM14]. All these works use the video data to determine the position of the recorded people; under the assumption of nonholonomic motion, the orientation is determined from the velocity vector. Hence, such techniques are not suitable for studying the torso twisting behaviour in crowds. Furthermore, Jacques et al. [JJRMJ10] state that "in a crowded situation, it is difficult to segment and track accurately each individual, mostly due to severe occlusions. In fact, when high-density video sequences are employed, the accuracy of traditional methods for object tracking tends to decrease as the density of people increases." For these reasons, we have chosen to employ a *motion capture system* for our experiment. Such systems have been used by Wolinski et al. [WJGO+14] to optimize parameters for various crowd simulation systems to increase similarity between the simulated and recorded crowd behaviour. Lemercier et al. [LJK+12] used motion capture to study following behaviour in crowds, by recording people following each other in a circular fashion. In both works, the participants are represented as moving discs, even though a motion capture system might have been able to capture more detailed motions. Hence, as with the video-driven techniques, the resulting data are unsuitable for the study of torso twisting behaviour. Full-body motion capture data has also been used as the basis for crowd animation techniques. Lee et al. [LCL06] introduced "motion patches", later extended by Yersin et al. [YMPT09] and Kim et al. [KHHL12]. These approaches use precomputed human motion, often obtained from motion capture, to animate and stitch together cyclic and collision-free behaviour. A small number of people (possibly just one) is recorded simultaneously, and multiple recordings are stitched together to form a crowd. As such, these techniques are ill suited to study torso twisting behaviour.

In our experiment, we compare the motions of people in dense crowds to the motions predicted by a generalized Voronoi diagram. It has been shown that such diagrams can be used to steer medium-density crowds [SAC+08]. However, it is unknown whether dense crowds also follow a Voronoi diagram. To our knowledge, we are the first to investigate this.

3 Experiment

This section describes the experiment design, execution, and results. In short, the goal of the experiment is to obtain information on how people manoeuvre through dense crowds.

3.1 Experiment design

The experiment consists of a repetition of trials. In each trial, the person in the centre of the crowd, the designated *walker*, manoeuvres towards a predefined goal position, while the other participants remain more or less stationary. The experiment is aimed at providing ground truth data, which can be analysed to further understand crowd behaviour, as well as provide empirical data to improve crowd simulation methods. We are particularly interested in the following aspects:

- What are typical values for linear and angular velocity, and is there any correlation between them?
- What are typical values for the angle between the orientation of the torso and the direction of a goal position?
- A mathematical model that predicts chosen path through the crowd.

When recording people using an optical motion capture system, marker occlusions are very common. The chance of occlusions occurring increases when the number of people grows, and more so when those people stand closer together. This makes it impossible to capture the full-body motion of each member of a large, dense group. Consequently, we chose to reduce the marker set, and place them on the body in areas that are least likely to be occluded from the overhead cameras. Participants are adorned with three motion capture markers: one on the right shoulder, and two on the left. The asymmetry in the marker layout allows us to distinguish between the participants' left and right side. Each participant started at a predetermined position, simplifying the identification and labelling when post-processing the recorded data.

A circle was drawn in the middle of the motion capture studio. All participants except the walker were prohibited from leaving this circle during each trial, to ensure a consistent crowd density. The location and size of the circle depended on the limitations of the motion capture system as well as the intended crowd density. Around the circle, the letters A-H, printed on A4 paper, were hung 2 meters high at 45-degree intervals. The motion capture system consisted of fourteen cameras recording at 100 Hz.

The experiment consisted of multiple sets of trials. For each set, a single participant was chosen for the role of *walker*. For each walker, we recorded a set of seven trials. The walker received a randomized set of task cards: a random subset of six of the eight letters, and a question mark. He/she was instructed to keep these cards hidden for the other participants. As two letters were excluded for each walker, the other participants would not be able to predict the tasks.

3.2 Execution

Participants were invited on the notion that the intended goal was to test the limits of our motion capture lab: to see how much people it can hold, and how many markers it can see at once. This way, the behaviour of the participants would not be not influenced by any knowledge about the actual goal of the experiment. Furthermore, the participants were asked to treat the situation as a densely packed bar and act naturally, and to crowd together to such a degree that it would be non-trivial but still possible to manoeuvre through. As for dealing with the walker, we asked participants to let him/her through as they would have in similar situations in real life and not anticipate their movement too much.

Each participant was measured for their chest width, chest thickness, and distance between the left and right shoulder markers. The first two measurements define the properties of our abstract representation (see Section 3.3). The last measurement served as extra reference for identification purposes. Furthermore we recorded their name, age and sex, and assigned starting positions for each trial.

Each *walker* was hand-picked from the participants on the basis of height and gender. We chose participants of average height, on the assumption that being too small or too large could influence the behaviour – we leave the influence of height on the behaviour in crowd to future work. We also alternated the gender of the walker to eliminate gender bias.

Each trial consisted of the following steps:

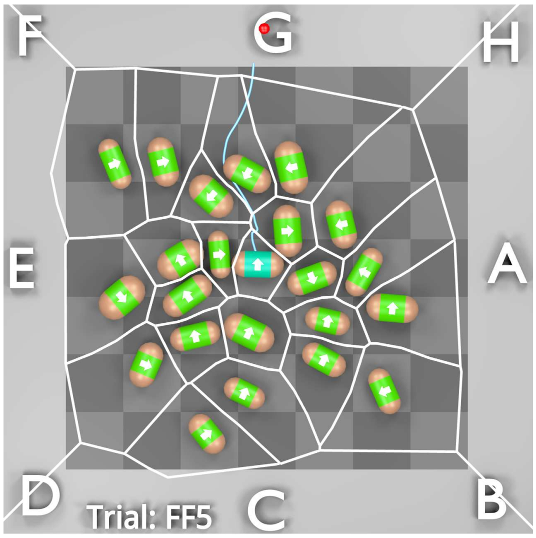


Figure 2: A situation of a task, showing the walker in the starting configuration and letter G as the goal.

1. Each participant moves to their predefined starting position. Recording starts.

2. The *walker* moves to the centre of the circle and rotates to face letter G, while the other participants walk around in a more or less random fashion. This ensures the crowd is different in each trial, and gives the walker ample time to covertly inspect his/her set of cards to determine the task to perform.

3. When the crowd is sufficiently randomized and evenly distributed around the walker, a verbal signal is given. The participants move to fill possible gaps in their vicinity, and then stop walking.

4. The walker manoeuvres through the crowd. Depending on the task, he/she tries to reach the goal letter, or, when the task card indicates a question mark, tries to exit the crowd in the easiest direction of his/her own choice.

5. When the task is complete, recording stops, and the next trial begins.

23 people (16 male, 7 female) participated in the experiment, with an average age of 24 years ($\sigma = 8.4$). Their average chest width was 0.44 metre ($\sigma = 0.03$), and chest thickness 0.23 metre ($\sigma = 0.03$). 7 participants took the role of "walker", and a total of 47 usable trials were

Figure 1: Photos of the predefined starting positions (left) and randomization step (right).

recorded; two recordings were rejected due to respectively a technical issue and a participant not adhering to the task.

3.3 Representation

After the experiment, all motion capture data was post-processed, manually labelled, and mapped to an abstract representation. This representation consists of an oriented line segment of length w with a radius r, and is shown in Figure 3. For each participant, the measured chest thickness T determines radius $r = T/2$, and the measured chest width W determines the line segment length $w = W - 2r$.

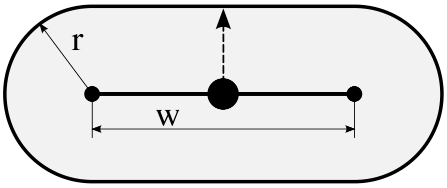


Figure 3: Our abstract crowd agent representation, consisting of an oriented line segment with a radius. The dashed arrow indicates the *forward vector* of the agent.

We use the centroid of the two right shoulder markers as a support point for the agent. The line defined by the ground projections of that support point and the left shoulder marker determines the orientation of the abstract agent. The centroid of those two points determines the position of the agent. The *forward vector* of the agent is defined as the forward-facing normal of the line segment. Hence, it is always perpendicular to the line segment, regardless of the linear velocity. This process is repeated for each agent at each frame of the motion capture data, and results in the set of agents translating and rotating in the ground plane used in our analysis.

4 Results

In this section we describe our experimental results, based on an analysis of the abstract representation of the motion capture data as described in the previous section.

The linear and angular velocities observed in our recordings are shown in Figure 4. These were calculated by numerical differentiation and applying a smoothing filter to improve numerical stability. The average linear velocity is 0.41 m/s ($\sigma = 0.23$), and the average angular velocity is 39 deg/s ($\sigma = 31$), with maxima at respectively 1.36 m/s and 176 deg/s. Figure 4 shows a large spread of the velocities in a near-Gaussian distribution. We have found no correlation between angular and linear velocities; fitting a linear correlation results in $R^2 = 0.01$.

The angle τ between the forward vector of the walker and the vector towards the target position tells us something about how often people keep their body oriented towards their target, and which angles are generally preferred. A histogram of these angles is shown in Figure 5. Fitting a Gaussian curve using a minimum-squared-error-approach results in $\hat{\tau} = 45^o$ ($\sigma = 35^o$). The right-hand graph in the same figure shows the percentage of time in which this angle is within a certain range. The relation between the angle limit and the time spent within that limit is more or less linear below 72 degrees, and then gradually flattens out until reaching its maximum of 100% at $\tau = 120^o$. Angle $\tau \leq 50^o$ in 50% of the time, and $\tau \leq 84^o$ in 90% of the time.

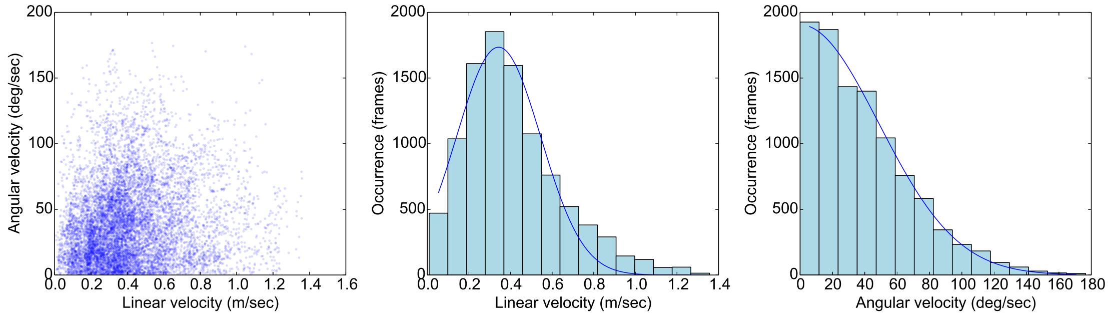


Figure 4: Linear and angular velocities. The scatter plot shows that there is no significant correlation between the two velocities. The histograms show the densities of the samples and a Gaussian fit.

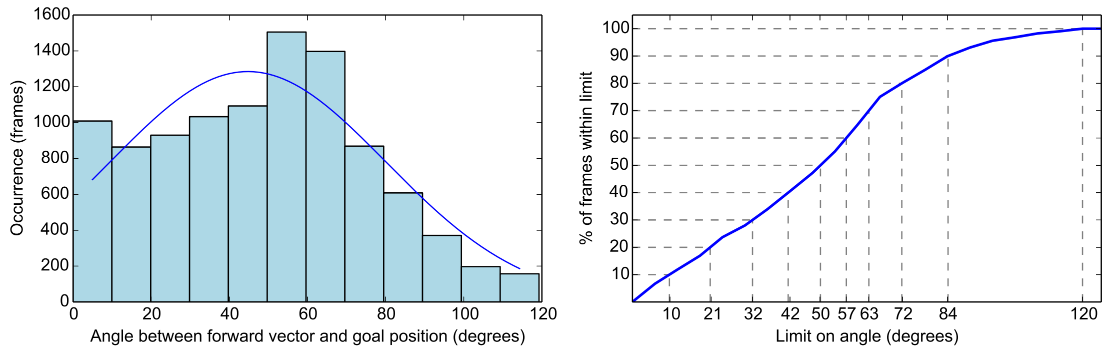


Figure 5: Angle between the agents' forward vector and to-goal vector. The right-hand graph shows the percentage of recorded frames for which that angle is within a certain limit.

For the "question mark" tasks the letter most often chosen was E, with 4 out of 7 walkers choosing this letter. Second was the letter G with 2 out of 7. The last letter picked was C. The letter E is to the left of the walkers, letter G directly in front, and letter C directly behind the walkers. It is interesting to see that someone took the effort to turn 180^o to find an easy way out of the crowd. However, the statistical significance is probably low due to the sample size. We intend to use the recordings of these as-fast-as-possible crowd escape tasks in future work, to optimize crowd simulation performance for such situations.

It is a known fact that people minimize their perceived energy use when walking [Zip49, GCC+10]. We predict that this holds true for dense crowds as well, and hypothesize that in such a crowd people:

1. no longer attempt avoiding single individuals, but choose an opening between pairs of individuals to pass through;
2. move towards the *midpoint* of the opening, as this provides a path with least chance of collision;
3. move towards the *area of largest clearance* behind that opening when that midpoint is closer to the agent than the distance from that agent to the other individuals.

A *midpoint* is defined as the middle of the opening between two agents. In our model, we use the middle of the line segment that represents the shortest distance between the line segments representing the two agents. By definition, this point lies on an edge of a generalized Voronoi diagram (GVD), or on an extension thereof. The centre of the *area of largest clearance* behind the opening corresponds to a vertex of the GVD.

To verify our expectations, we have analysed the difference between the actual direction of movement and the direction predicted by the expectations described above. In our analysis, we use three different methods to test our expectations, first partially and then completely:

MIDPOINT Only *midpoints* are considered, regardless of the distance to the agent. This corresponds to only testing the first and second hypothesis mentioned above.

VERTEX Only *vertices* of the Voronoi cell defined by the agent are considered. This corresponds to only testing the first and third hypothesis mentioned above, regardless of the distances.

LIMITED MIDPOINT *Midpoints* are considered, but limited to the closest Voronoi *vertex* on their corresponding edges. This corresponds to all hypotheses mentioned above.

We aim to compare the directions predicted by the methods described above with the overall direction of movement of the participant. However, this direction cannot directly be determined by the instantaneous velocity vector, since this vector varies too rapidly. Such rapid changes are especially noticeable at low velocities, where merely shifting weight can rotate the velocity vector by 180^o. To filter out these variations, we consider the recorded positions of the agent, and obtain the vector from the current position to the recorded position at a given Euclidean distance of $D \in [0.05, 1.5]$ metres. This position is uniquely defined due to the nature of our recordings, as the participants moved more or less monotonically towards their target. The distance is named the *look-ahead distance*, as it could indicate how far people look ahead in order to plan their direction. An example is shown in Figure 6. To our best knowledge, the relation between the agent's instantaneous speed and the preferred look-ahead distance has not been studied in the context of dense crowds.

Our analysis will be limited to the period of *dense manoeuvring*. Every recorded trial shows three phases:

Startup, where the walker turns towards the goal position, investigates the situation, and shifts balance in order to start manoeuvring.

Dense manoeuvring, where the walker manoeuvres through the crowd. This is the longest of the three phases, and subject of our analysis.

Crowd exit, where the walker is no longer part of the dense crowd, and reaches the goal position.

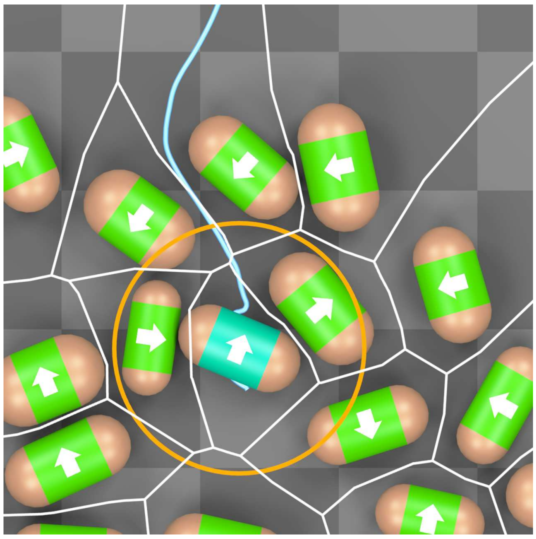


Figure 6: Recorded movement is shown as a light blue trajectory. The yellow circle indicates a look-ahead distance of 0.4 metre.

The phase transitions depend on what is considered as "dense manoeuvring". The transition from the first to the second phase is subjectively defined by the start of actual manoeuvring, as our experimental design ensures we start in a dense situation. The transition from the second to the third phase is defined by properties of the walker's Voronoi cell. The generalized Voronoi diagram (GVD) is a pair $\{V, E\}$ of vertices V and edges E that partition the ground plane. Each cell of this partition is defined by an agent, and consists of all points that are closer to that agent than to any other. We continually measure the cell's surface area and neighbouring cells. When the surface area is larger than 0.32 m^2, or when the cell is incident to a Voronoi cell defined by the bounding box, we no longer consider the agent to be in a dense crowd. There are other possible definitions of "dense crowd manoeuvring"; this one worked well in practice for our recordings. An important aspect is that our definition does not consider the aggregate crowd density, but categorizes the situation of individual agents. In most of our analyses we simplify the GVD by using the agents' line segments, rather than using their exact shape. We will also compare this GVD to a Voronoi diagram defined by the agents' centre points.

Now that we can compute a direction of movement and find the period of dense manoeuvring, we can test our hypothesis. Since there is a correlation between crowd density and walking speed [Daa04], we have included the speed of the agent as a variable in our analysis. Figure 7 shows the graphs for the various testing methods. For each frame F in our recordings, we determine the instantaneous speed of the walker $s(F)$. This determines the bin column in which the data for that frame is plotted. We then vary look-ahead distance D to determine direction of movement vectors $\vec{d}$, determining the bin row in the graph, and calculate error $e(\vec{d}, P)$ with respect to a prediction P, as described in the next paragraph. The error is stored at the appropriate bin; the graph shows the average (left) and standard deviation (middle) of the binned errors, and the number of data points in each bin (right).

The error $e(\vec{d}, P)$ depends on prediction P, which is the set of points considered by the different prediction methods described earlier[1]. The error is defined as the minimal angle between the direction of movement $\vec{d}$ and the prediction points in P:

$$e(\vec{d}, P) = \min_{\vec{p} \in P} \left[\text{acos} \left(\vec{d} \cdot \frac{\vec{p} - \vec{x}}{||\vec{p} - \vec{x}||} \right) \right]$$

For each binned speed we find the look-ahead distance bin that results in the smallest predicted error. These bins are marked with a white dot in Figure 7. The white line indicates the best linear regression through those bins, weighted by the number of data points.

5 Discussion & Conclusion

Figure 7 shows, for each speed, white dots at the bins with the minimal error. These points identify the most likely look-ahead distance L for speed v. Fitting a line shows a slightly positive relation $L = 0.2v + 0.3$, indicating that people tend to plan over longer distances when their speed increases. We suspect this is due to the relation between speed and density; when the density decreases, distance to others increases as well as the speed.

Our most prominent result is the correlation between dense crowd manoeuvring patterns with generalized Voronoi diagrams. We generated this Voronoi diagram by modelling the participants as line segments, and analysed the motions of the agents. Our results show that with an average error of less than 7^o ($\sigma = 5^o$) our LIMITED MIDPOINT method successfully matches the paths of our participants. In comparison, representing agents as points instead of line segments, results in a slightly higher average error, but also higher fluctuations in the standard deviation. We can conclude that, for people in dense crowds, line segments are a better fitting model than points.

We have measured the participants in the directions that are relevant to an analysis performed on the ground projection of the data. It would be interesting to investigate the influence of the participant's relative height to the other participants on their behaviour, as this may both influence the participant as well as the others in their proximity.

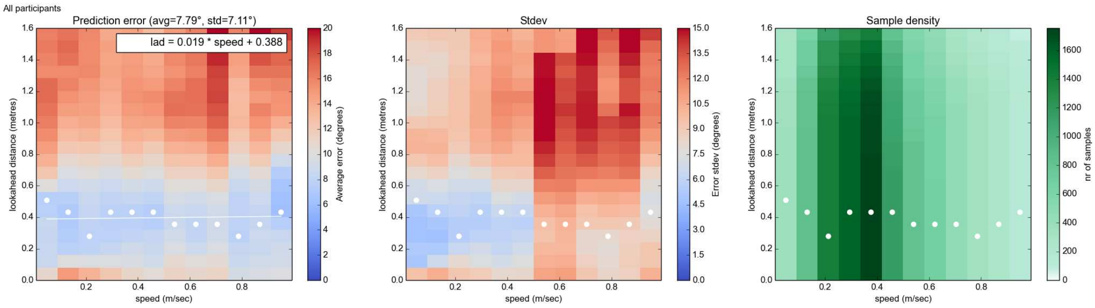
As we focus on torso orientations, the orientation of the head was not recorded during our experiment. A future experiment, employing a cap with motion capture markers, could record head movement. Such recordings could produce more natural results when applying full-body animation, as well as provide interesting data on the impact of vision on manoeuvring behaviour.

During the experiment, the participants let us know that they felt uncomfortably close to each other. We suspect that this is not only caused by the higher density; after all, in daily life there are denser crowds where even manoeuvring is sheer impossible. However, in such cases there is often something that draws away the attention of the crowd, such as a performing artist, whereas in our case there was little to focus on except each other. It would be interesting to investigate the effect of such a distractor on the natural density of a crowd.

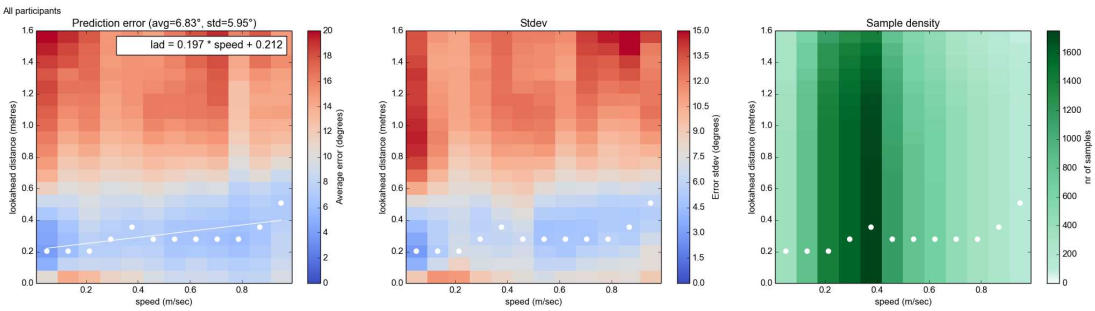
Acknowledgements

We want to thank everybody who took the time to participate in our experiment. This research is supported by the Dutch nationally funded project COMMIT.

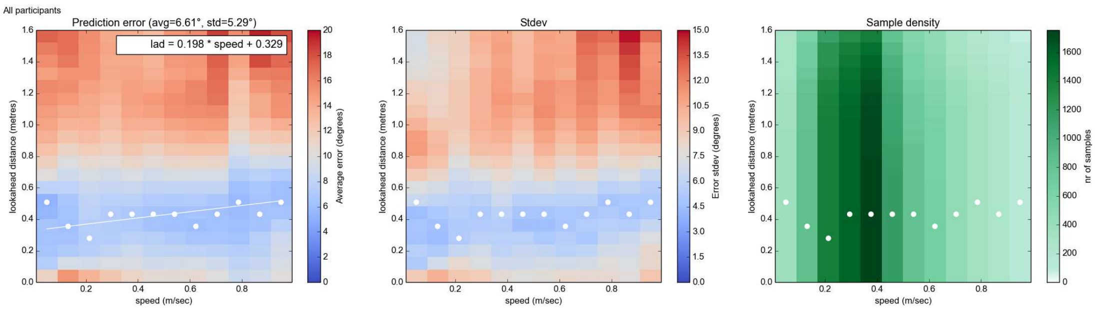
[1] We refer to the methods MIDPOINT, VERTEX and LIMITED MIDPOINT



(a) MIDPOINT method



(b) VERTEX method



(c) LIMITED MIDPOINT method

Figure 7: Results of our analysis. The agent's instantaneous speed is shown on the x-axis; the chosen look-ahead distance D is shown on the y-axis. The colour indicates the average error over all recorded trials (left), standard deviation (middle) and number of data points (right).

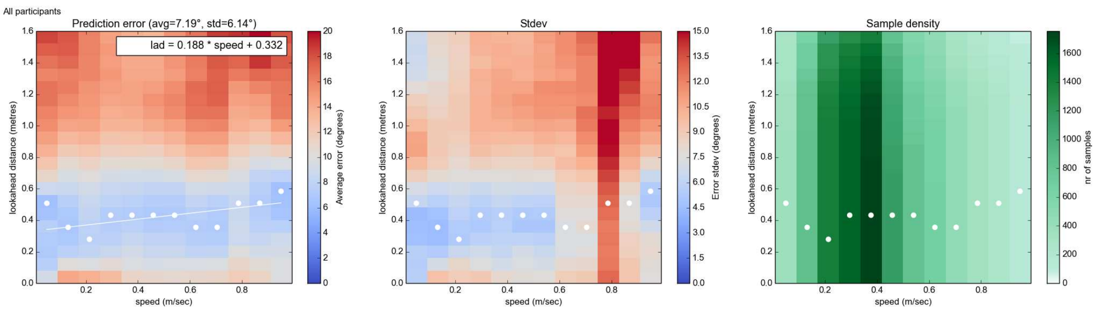


Figure 8: LIMITED MIDPOINT method, but using points instead of line segments to represent the agents.

References

[BM14] A. Bera and D. Manocha. Realtime Multilevel Crowd Tracking using Reciprocal Velocity Obstacles. *IEEE International Conference on Pattern Recognition*, 2014.

[Daa04] W. Daamen. *Modelling passenger flows in public transport facilities.* Delft University Press, 9 September 2004.

[GCC+10] S.J. Guy, J. Chhugani, S. Curtis, P. Dubey, M. Lin, and D. Manocha. Pledestrians: a least-effort approach to crowd simulation. In *Proceedings of the 2010 ACM Siggraph/Eurographics Symposium on Computer Animation*, SCA '10, pages 119–128, Aire-la-Ville, Switzerland, Switzerland, 2010. Eurographics Association.

[JJRMJ10] J.C.S. Jacques Jr., S. Raupp Musse, and C.R. Jung. Crowd analysis using computer vision techniques. *Signal Processing Magazine, IEEE*, 27(5):66–77, Sept 2010.

[KHHL12] M. Kim, Y. Hwang, K. Hyun, and J. Lee. Tiling motion patches. In *Proceedings of the 11th ACM SIGGRAPH / Eurographics Conference on Computer Animation*, EUROSCA'12, pages 117–126, Aire-la-Ville, Switzerland, Switzerland, 2012. Eurographics Association.

[LCHL07] K.H. Lee, M.G. Choi, Q. Hong, and J. Lee. Group behavior from video: A data-driven approach to crowd simulation. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '07, pages 109–118, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.

[LCL06] K.H. Lee, M.G. Choi, and J. Lee. Motion patches: Building blocks for virtual environments annotated with motion data. *ACM Trans. Graph.*, 25(3):898–906, July 2006.

[LJK+12] S. Lemercier, A. Jelic, R. Kulpa, J. Hua, J. Fehrenbach, P. Degond, C. Appert-Rolland, S. Donikian, and J. Pettré. Realistic following behaviors for crowd simulation. *Computer Graphics Forum*, 31(2pt2):489–498, 2012.

[MOS09] R. Mehran, A. Oyama, and M. Shah. Abnormal crowd behavior detection using social force model. In *IEEE conference on Computer vision and pattern recognition*, pages 935–942, June 2009.

[RSLA11] M. Rodriguez, J. Sivic, I. Laptev, and J.-Y. Audibert. Data-driven crowd analysis in videos. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 1235–1242, Nov 2011.

[SAC+08] A. Sud, E. Andersen, S. Curtis, M. Lin, and D. Manocha. Real-time path planning for virtual agents in dynamic environments. In *ACM SIGGRAPH 2008 Classes*, SIGGRAPH '08, pages 55:1–55:9, New York, NY, USA, 2008. ACM.

[SBTM08] S. Saxena, F. Brémond, M. Thonnat, and R. Ma. Crowd behavior recognition for video surveillance. In Jacques Blanc-Talon, Salah Bourennane, Wilfried Philips, Dan Popescu, and Paul Scheunders, editors, *Advanced Concepts for Intelligent Vision Systems*, volume 5259 of *Lecture Notes in Computer Science*, pages 970–981. Springer Berlin Heidelberg, 2008.

[WJGO+14] D. Wolinski, S. J. Guy, A.-H. Olivier, M. Lin, D. Manocha, and J. Pettré. Parameter estimation and comparative evaluation of crowd simulations. *Computer Graphics Forum*, 33(2):303–312, 2014.

[YMPT09] B. Yersin, J. Maïm, J. Pettré, and D. Thalmann. Crowd patches: Populating large-scale virtual environments for real-time applications. In *Proceedings of the 2009 Symposium on Interactive 3D Graphics and Games*, I3D '09, pages 207–214, New York, NY, USA, 2009. ACM.

[Zip49] G.K. Zipf. *Human behavior and the principle of least effort*. Addison-Wesley Press, 1949.

[ZL14] F. Zhao and J. Li. Pedestrian motion tracking and crowd abnormal behavior detection based on intelligent video surveillance. *Journal of Networks*, 9(10), 2014.

[ZZL09] X. Zheng, T. Zhong, and M. Liu. Modeling crowd evacuation of a building based on seven methodological approaches. *Building and Environment*, 44(3):437 – 445, 2009.