SPS Chapter Research Award Proposal

Quantitative evaluation of pedestrian movement models:
A *real* many-body problem

School:	Purdue University
Chapter:	5781
Requested:	$1089.00

Abstract

A number of papers have been published since 2000 which attempt to model pedestrian crowd flow dynamics using basic equations of motion. Here we propose a study which aims to critically evaluate a handful of these models in terms of their predictive accuracies, then categorize them according to their strengths and weaknesses. This will be accomplished by taking aerial footage of crowds on campus using an unmodified quadcopter drone.

1 Proposal Statement

1.1 Overview of Proposed Project

In this project, we will attempt to answer 1. Whether a visible-range camera mounted to an unmodified quadcopter drone can be used to effectively collect quantitative crowd flow data, and 2. Whether the models we will discuss in the background section accurately model crowd flow in reality.

This project is important for two main reasons. First, a given visible-range camera is significantly cheaper than an infrared camera of the same resolution. So, if we are able to successfully recover pedestrian position data in crowded settings using only a visible range camera, future studies can be performed with lower upfront costs.

Second, in contemporary science, there is an appalling dearth of replication studies due to severe lack of funding for such research. In a field such as quantitative sociology, theoretical models are almost exclusively semi-empirical, thus placing a heightened importance on experimental data. In our study, we hope to provide a critical, quantitative evaluation of each theory presented in such a way that tangible, robust conclusions about its methods or assumptions may be drawn.

In this project, we will use a quadcopter drone equipped with a built-in camera to record the movement of crowds from some fixed location above them. Using simple machine vision methods, we will extract pedestrian positions and velocities as a function of time from video data. Finally, we will compare model predictions to this observed data. A specific example of the type of error analysis we will employ is given in the "Description of Proposed Research" section.

In this project, our members will be exposed to applied computing challenges such as modeling, machine learning, and error analysis. They will also be challenged to connect ideas between frameworks that are expressed in different or unfamiliar notations. We intend to educate every member of this project to an extent that they could independently implement one of the models discussed in a scripting language of their choosing. At a national level, we believe that if our project is successful and if our results are properly disseminated, other chapters might be inspired to attempt research of their own.

1.2 Background for Proposed Project

Quantitative models of pedestrian movement have been present in technical literature since at least the mid-1900's. An early example is given by Henderson [6], in which a crowd is modeled as a sort of gas whose dynamics are determined by the Boltzmann transport equation. This analytical but highly idealized approach demonstrated that shock waves could occur in crowds, and that numerically predictable collective action could arise among a set of independent, free-willed agents. Though this has massive implications for sociology, philosophy, and civil engineering, a majority of the work leading up to the turn of the century was done by physicists and engineers who were intrigued by the parallels between collective human behavior and fluid dynamics [5].

In 1995, Dirk Helbing et al. published a seminal work in the field of crowd flow dynamics titled Social force model for pedestrian dynamics. In this work, he modeled humans as psychologically driven automata subject to reasonably defined social forces. These forces, which include a motivational force (i.e. a desire to reach a destination), a general attractive force, and an interpersonal repulsive force, act on the agent as though it were a particle obeying Newtonian mechanics. In simulation, this simple model yielded rich, recognizable phenomena such as lane formation in passing crowds and entryway obstruction due to group egress. Five years after his introduction of this model, Helbing examined the phenomenon of escape panic using a slightly more nuanced framework. This investigation, which was published in Nature, demonstrated how the social force model could be augmented to include true (i.e. external, non-psychological) external forces and thus give a quantitatively accurate depiction of personal injury risk as a function of escape route geometry and person density.

We plan to examine a number of predictive models which were published in the 2000's that expand upon these foundational works. A few [8, 12, 13] focus on panicked crowd dynamics and flow through pinch points, while others [8, 9, 3, 10] attempt to advance the social force model. We use the term predictive to describe models which compute crowd state as a function of time given some initial conditions and free input parameters. Since these models are chaotic, even the most comprehensive among them will produce solutions that diverge after a long enough time. Accuracy will therefore be judged over small time domains. Here, each of the models is described in terms of its inputs, outputs, basic assumptions, and scope.

Langston et al. [8] provide an extension to [4]. In order to better understand flow dynamics in tightly confined situations, each pedestrian is modeled as the union of three overlapping physical circles which represent a torso and shoulders, and a circumscribing psychological personal space circle. Each pedestrian is given an aim point and a set of semi-empirical parameters that determine their biometrics and behavior. Initial positions and velocities determine the state of the system at all future times according to Newtonian dynamics and numerical integration in time. To evaluate this model, one must provide obstacle information, pedestrian position and velocity, aim points (which can be determined ad hoc by final pedestrian position) and estimates for biometric parameters. This model is appropriate for 10s of people in medium to high density configurations.

Song et al. [12] provide new conclusions for the system proposed in [4] by combining the social force model with a lattice gas model. This augmentation provides insight into panicked crowd phase transitions and the scaling of egress rate with respect to exit size. These points are not addressed by [4]. In this model, each actor shares the same aim point, which is the exit. Pedestrian location is quantized to a grid, and movement during each timestep is determined probabilistically according to social force considerations. As such, it may only be possible to evaluate the accuracy of this model in crowded areas with few exits. Such situations occur regularly before and after large lectures, though panic is not necessarily present.

Cristiani et al. [3] model crowd flow on both microscopic and macroscopic scales simultaneously. In the microscopic picture, individuals are discrete elements with movement defined by ODEs, while in the macroscopic picture, the crowd is a continuous density distribution that evolves according to PDEs. Using a rigorous measure-theoretic framework, these representations are mixed in proportion to a dimensionless free parameter $\theta \in [0, 1]$. At $\theta = 0$, the model is fully discrete, while $\theta = 1$ gives the exclusively continuous description. For all other values, actor position, velocity, density, and current density are interdependent. In the paper, simulations are performed with crowds of 10 to 100 people. Larger crowds are ostensibly included in the theoretical scope, but in examining them computational load may become prohibitive.

Narain et al. [10] consider the problem of scalably simulating large (i.e. 10^5 individuals), dense crowds. Their model, like that of cristiani et al., proposes a mixing of continuous and discrete representations. Unlike the cristiani model, no scaling parameter similar to θ exists, and the continuous representation is only used as an internal state to help update the discrete representation. Agent positions and velocities first update the continuous distribution, which is then transformed according to the Unilateral Incompressibility Constraint (UIC). Finally, the continuous distribution is transformed back into the discrete representation and motives are recalculated. The UIC prevents crowd density from exceeding a chosen maximum value, and accounts for the fact that human crowds are not purely compressible or incompressible. Furthermore, the UIC reduces computational complexity of collision avoidance massively, thus allowing for simulation of massive crowds. This model is unlikely to provide good performance in dilute crowds, though in low-panic high-density situations, it may perform better than panic-focused models such as [Song]. No attempt is made by the authors to fit model performance to experimental data.

Mousad and Helbing et al. [9] propose a model which emphasizes cognitive science principles and steps away from the pairwise interaction scheme. Each actor behaves according to two heuristics. First, a person walks in the direction of the most direct path towards their destination point. Second, a person will maintain a distance from obstacles in the chosen path that allows for a minimum stopping time τ. In order to implement this model, we will have to provide values for several empirical parameters, though the paper provides good estimates. Furthermore, it will be necessary to detect each pedestrians orientation. Detection of this parameter is not required for any other model we are investigating. This model is most well suited to 10s to 100s of people in dilute crowds, where each person has a well-defined objective.

In order to visualize flow segmentation, data will be characterized according to a method proposed by Ali and Shah [1]. Their method uses optical flow to estimate the Lyapunov exponent at each point in the image, then segments the spatial domain according to separatrices (defined by locally large Lyapunov exponent). In doing so, the crowd can be viewed in terms of regions of similar fate. The accuracy of this model cannot be tested in the same way as the others, for it does not make quantitative predictions about individual movement. Instead, it will be used to examine movement structures that may have otherwise gone unnoticed. Ali and Shah present this method in the context of very large (10^3 to 10^4 people) crowds in wide open areas, but it will be applied here on much smaller crowds in more confined areas.

1.3 Expected Results

The first question which we will attempt to answer is whether or not an off-the-shelf drone and our homemade software can provide accurate, automated pedestrian movement tracking. Even if every predictive model we examine does an awful job of predicting pedestrian movement, it will still be worthwhile as a proof of concept to develop the proposed data acquisition method.

Assuming we succeed in developing the data acquisition method to a useful state, the next goal will be to determine which of the discussed methods is superior. Since it is unlikely that one model performs better than all others in every way, our results on this front will be multi-faceted. Consequently, numerous error quantifiers will be proposed and tested in an attempt to find the specific circumstances under which a given model fails or succeeds.

1.4 Description of Proposed Research - Methods, Design, and Procedures

The first challenge to address in this study is the extraction of individual pedestrian movement data from video captured on a quadcopter drone. Video will be taken from a stationary overhead vantage point, thus each pedestrian will have approximately the same profile. This configuration provides a high level of feature regularity and should be suitable for characterization under most statistical classification methods. In particular, we intend to try detection via background subtraction [7] and support vector machines (SVM) [11]. In both of these relatively simple approaches, no external library of training data will be necessary, although there are more complicated classification routines requiring such inputs that could be employed [2].

In addition to determining the feasibility monitoring crowd flow with an ordinary drone, the goal of this study is to quantitatively evaluate the predictive powers of each model discussed in the background. Many of these models represent pedestrians discretely, i.e. as a set of positions $\{\mathbf{x}_i\}_{i=1}^N$ and velocities $\{\mathbf{v}_i\}_{i=1}^N$. In the case of [12], positions are quantized to a grid, and velocity is fixed. In [10] and [3], continuous representations are used, but only in [3] is it an external (i.e. predicted observable) variable of the model. When such special considerations are not needed, defining a function which measures error locally in time is not difficult. Let $\mathbf{x}_i^n, \mathbf{v}_i^n \in \mathbb{R}^2$ give the true position and velocity of actor i at frame n, and let the timestep in the model be given by the reciprocal of video framerate. This establishes a direct correspondence between predicted values and observed values. Borrowing notation from [1], let $\phi_{t_0}^t$ denote the predicted flow map associated with position and define $\psi_{t_0}^t$ similarly for velocity. Then, for s time steps, $t = s\Delta t + t_0$. Accordingly, we choose to rewrite $\phi_{t_0}^t$ as ϕ^s. Given the true position $\mathbf{x}_i^n$ of pedestrian i at timestep n, the model predicts that in s timesteps, that actor should be at location $\phi^s\mathbf{x}_i^n$. Since their true position is $\mathbf{x}_i^{n+s}$, the s-step position error at time t_n can be defined:

$$\epsilon_{ni}^{(s)} = ||\phi^s\mathbf{x}_i^n - \mathbf{x}_i^{n+s}|| \tag{1}$$

A similar quantity can be defined for velocity:

$$\eta_{ni}^{(s)} = ||\psi^s\mathbf{v}_i^n - \mathbf{v}_i^{n+s}|| \tag{2}$$

Time-dependent error will be computed according to these expressions for each model. Locally large error can then be correlated to system state and used to elucidate weaknesses in model assumptions. Furthermore, by examining error in terms of step number s, solution stability can be addressed in terms of its rate of divergence from accurate behavior.

To demonstrate this process concretely, example calculations are performed on a simulation of Helbing's original social force model [5]. For brevity, the mathematical details of the model are omitted here. In this demonstration, 10 pedestrians are moving in a hallway of width 7 m. The initial state of the system is given in Figure 1. Velocities are given by black arrows, force of attraction to objective points is blue, interpersonal repulsions are red, and repulsions from the walls on either side are green. One pedestrian is singled out for error analysis. Their initial position is circled in red, and their trajectory is given by a dashed line. In Figure 2, the path of each other pedestrian is given.

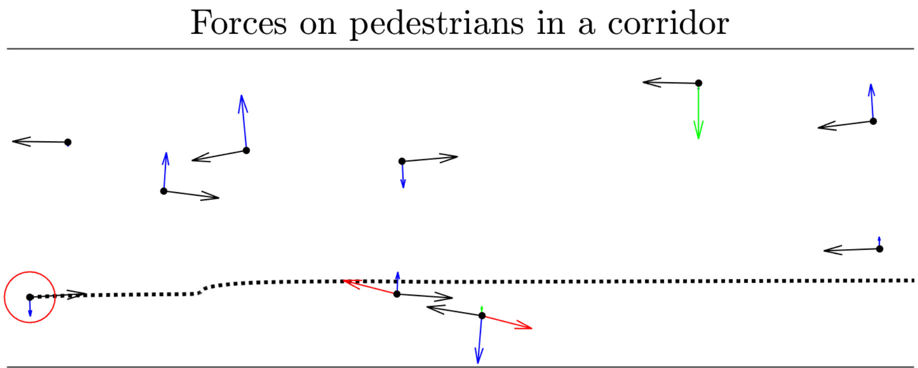

Figure 1: Initial positions, velocities, and forces for 10 pedestrians in a corridor. Model parameters are $\tau_\alpha = 0.5$ s, $v_\alpha^0 = 1.34$ m/s, $V_{\alpha\beta}^0 = 2.1$ m^2/s^2, $U_{\alpha B}^0 = 10$m^2/s^2, $v^{max} = 1.3 v_\alpha^0$, $\sigma = 0.3$ m, $R = 0.2$ m, $\phi = 100°$, $c = 0.5$, and $\Delta t = 2$ s.

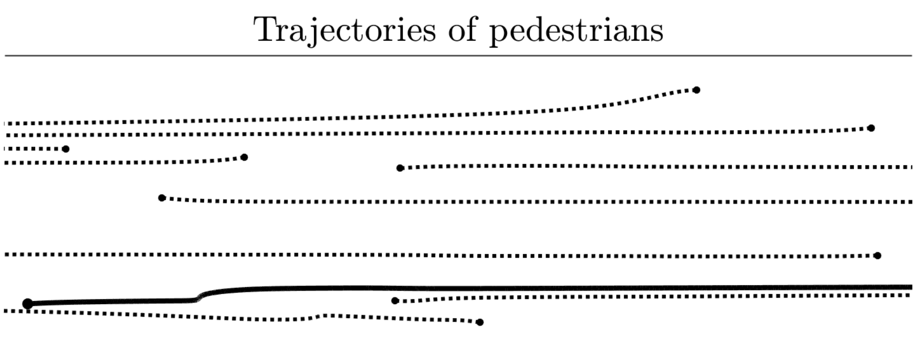

Figure 2: Trajectories of all pedestrians. Solid circles denote starting locations. Integration was computed using Euler's method with $dt = 0.005$ s.

In order to introduce an error which can be analyzed, the model is again applied but with one input parameter adjusted slightly. In particular, the relaxation time τ_α is changed from 0.5 s to 0.25 s. Equation (1) is then computed at every time t_n with $s = 1$ and ϕ representing one finite integration step with the new relaxation time influence. In Figure 3, $\epsilon_i^{(1)}(t)$ is given on a $\log_{10}$ scale for each trajectory. Because τ_α affects the time it takes a pedestrian to reorient towards their goal point, the error is largest when many reorientations are occurring, i.e. in the transient regime. Furthermore, some peaks in error can be seen to align in time and magnitude. When pedestrians are approaching each other, their interpersonal potentials are made nearly identical, and therefore they both experience reorientation forces of similar magnitude. It is thus shown that physical meaning can be extracted from the error plot, even in this highly simplified example.

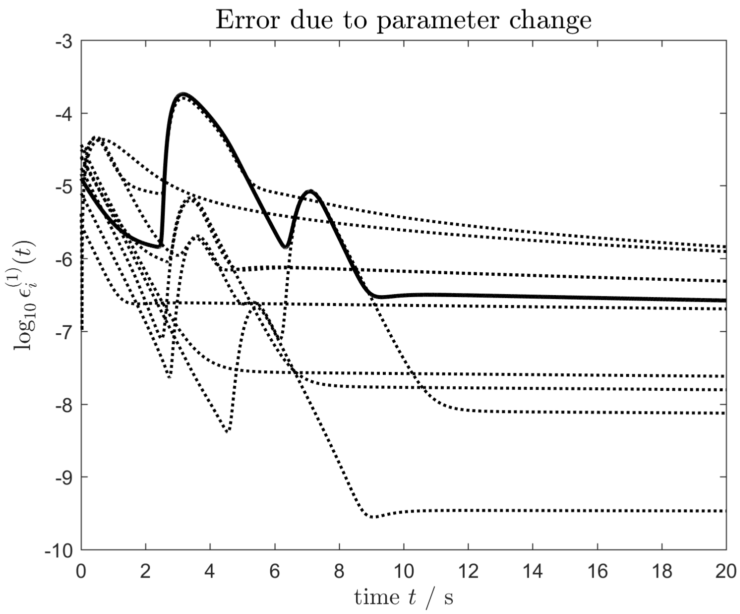

Figure 3: Local error induced by smaller relaxation time for each trajectory.

In practice, the "original" paths against which models are compared will be given by experimental data. It is expected that errors will be much larger than those observed here. By looking at $\epsilon^{(s)}$ and $\eta^{(s)}$ for larger values of s, it may also be possible to determine if a given model can provide robust qualitative predictions despite poor immediate accuracy.

1.5 Plan for Carrying Out Proposed Project

We will first build programs in Python or MATLAB (perhaps both) for each model discussed in the background section. Before any data is taken, we can test each model to verify that it reproduces the results given in its corresponding publication. When this is achieved, we can be certain that our programming is sound and will be suitable for use in error analysis. This work can occur before the drone is received, but will likely take quite a while due to the challenge of programming higher complexity models.

Once the drone is obtained, we will gather preliminary training data. During these trial runs, we will experiment with hovering height, lighting (i.e. time of day), and weather. Additionally, we will attempt to determine whether the presence of a drone is sufficiently distracting to contaminate our data. As was noted in the 'Methods' section, we expect a simple method such as SVM or background subtraction to be sufficient for extracting person position in every frame. It is important that only one video characterization method is employed for all models, though its possible that several will be developed in this stage.

Once we have a functioning, automated method of extracting position and velocity data from overhead video, we will begin recording data. Though all analysis is performed after data has been taken in full, waiting until the method is optimized will ensure that fewer repeat runs are necessary. Assuming that the data extraction method works well, it should not be difficult to obtain hundreds of data sets over the course of a few months.

After a sufficient volume of data has been collected, we will analyze the accuracy of each model according to the same metrics. For example, we may choose to measure ϵ and η in each model for several values of s. This procedure is discussed more thoroughly in the previous section. Using the unearthed discrepancies (or the lack thereof) we will characterize the regimes in which each model thrives or falls short. At this stage, the specific goals will be more malleable, and more models or error classifiers may be introduced to the scope of our study if time permits.

At the head of this project is Adam Kline, the current President of SPS Purdue. He is a senior with three years of research experience in computational modeling, machine learning, and gross looking math. He will provide direction for the project and technical instruction wherever it is necessary. Charles Li is a senior in electrical and computer engineering and holds the "secretary" officer position. He is the most experienced programmer on the team, has the technical know-how to fix and fabricate electrical components, and will help provide support to younger members. Dawith Lim is a junior and holds the "Stockmaster" officer position within this chapter. He is familiar with modern research tools and techniques in statistical and stochastic physics, and will help assimilate and understand interesting literature. David Lamey is a senior in math and physics and holds the position of Vice President. For very mathematically rigorous papers, his input will be invaluable. The remaining officers are Emily Firehammer, Nathan Shrum, Nathan Glotzbach, and Connor Mohs. Each one is supremely talented and will no doubt be of great service to the team.

Hiram Diaz, Andrew Gustafson, Braden Buck, Porter Hunley, Akshat Jha, Wyatt Montgomery, Henry Dawson, and Hao Wan are a few of the underclassmen who will be participating in the project. They helped to put together this proposal and have displayed enormous eagerness to learn more about the topic of this study and physics more generally.

1.6 Project Timeline

- January 8th February 19th

  Building and verifying programs for each model

  Training of dedicated drone operator(s)

  Obtaining test run videos to find optimal operating geometry (i.e. flight height, locations, etc.)

- February 19th March 8th

  Start obtaining training sets for feature definition

  Develop fully automated pedestrian data extraction method

- March 8th - April 6th

  Obtain as much pedestrian data as possible

- April 6 - May 1st

  First round of model analyses

  Prepare interim report

- September 1st - October 6th

  Review results from interim report and decide on new directions

  Consider new models if necessary

  Update error definitions if necessary

- October 6th - December 1

  Polish data, results, and code

  Begin writing final report

  Prepare poster if appropriate

2 Budget Justification

Our SPS chapter has selected the DJI Mavic Pro Drone Quadcopter as our primary research apparatus to collect human movement information. Given the features and specifications of this particular drone, we expect it to provide workable data with greater expediency than others. To demonstrate this, we will be comparing the Mavic drone with a lower tier drone, the DJI Phantom 3 Standard Quadcopter. In particular, we will show that the workarounds necessary to produce high quality results with the lower tier drone are not compensated for by its lower cost. The primary issues to address are image data stabilization, pilot training, necessary flight time, and image resolution.

In order to accurately collect pedestrian movement without adding a bias velocity, we will need to completely stabilize the drone in midair. This issue either needs to be resolved in software with drift correction, or simply having a camera with negligible drift. While software-based drift correction is a well-developed field programming such measures manually would prove time consuming and arduous. The Mavic provides absolute stability GPS capabilities, essentially negating any need for drift correction software. In comparison, the Phantom has a lower tier GPS system and is expected to drift 10-15 feet. Its more advanced stability features like Point of Interest and Wayfinding are unavailable and may remain so indefinitely.

If we had an individual in our group with drone flying experience, we would be more comfortable selecting a lower end drone, as he or she would be able to compensate for drift with skill. As it stands, no one in our group has such experience, and we will need to train a pilot to fill this role. On this front, the Mavic is again the better choice not only for its ease of use, but also for its built-in obstacle-avoidance feature, something the Phantom does not have. With this, we can more comfortably have a novice pilot fly our equipment without fear of catastrophic failure on impact with the environment.

Another relevant specification is total flight time. We will be collecting data over a 15 minute interval, as our window of interest is the 10 minute period of peak activity between classes at Purdue. We will want to be in the air 1-2 minutes before this period in order to collect classes that tend to be released earlier, and track the group dynamics of a less crowded system. We would similarly like to capture data directly after the passing period. These constraints do not yet account for the time required for setup and navigation to recording position. All latencies considered, we will need a drone with an in-air lifetime that is comfortably above the minimum 10 minute mark. The Mavic has a life expectancy of 27 minutes, while the Phantom is only expected to last 10 minutes. In order to mutually ensure a sufficiently long flight time and the safety of pedestrians, it necessary to choose the Mavic.

Finally, the Mavic has twice the resolution of the Phantom, 1080p vs 720p. In many scenarios, the higher pixel count would be irrelevant, or even a hinderance if computation time were an issue. This would be especially true in a case where real time analysis was necessary. However, image processing will be done using a trained computer vision model in the post processing phase, thus a higher resolution can only be better. A 1080p image has twice the pixel count of a 720p one, which on average doubles the number of features our detection algorithms can work with.

References

[1] Saad Ali and Mubarak Shah. A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–6. IEEE, 2007.

[2] Md Zahangir Alom and Tarek M Taha. Robust multi-view pedestrian tracking using neural networks. *arXiv preprint arXiv:1704.06370*, 2017.

[3] Emiliano Cristiani, Benedetto Piccoli, and Andrea Tosin. Multiscale modeling of granular flows with application to crowd dynamics. *Multiscale Modeling & Simulation*, 9(1):155–182, 2011.

[4] Dirk Helbing, Illés Farkas, and Tamas Vicsek. Simulating dynamical features of escape panic. *Nature*, 407(6803):487–490, 2000.

[5] Dirk Helbing and Peter Molnar. Social force model for pedestrian dynamics. *Physical review E*, 51(5):4282, 1995.

[6] Leroy F Henderson. On the fluid mechanics of human crowd motion. *Transportation research*, 8(6):509–515, 1974.

[7] ZuWhan Kim. Real time object tracking based on dynamic feature grouping with background subtraction. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[8] Paul A Langston, Robert Masling, and Basel N Asmar. Crowd dynamics discrete element multi-circle model. *Safety Science*, 44(5):395–417, 2006.

[9] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences*, 108(17):6884–6888, 2011.

[10] Rahul Narain, Abhinav Golas, Sean Curtis, and Ming C Lin. Aggregate dynamics for dense crowd simulation. In *ACM Transactions on Graphics (TOG)*, volume 28, page 122. ACM, 2009.

[11] Constantine Papageorgiou and Tomaso Poggio. Trainable pedestrian detection. In *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, volume 4, pages 35–39. IEEE, 1999.

[12] Weiguo Song, Xuan Xu, Bing-Hong Wang, and Shunjiang Ni. Simulation of evacuation processes using a multi-grid model for pedestrian dynamics. *Physica A: Statistical Mechanics and its Applications*, 363(2):492–500, 2006.

[13] Yusuke Tajima and Takashi Nagatani. Scaling behavior of crowd flow outside a hall. *Physica A: Statistical Mechanics and its Applications*, 292(1):545–554, 2001.