Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 2 (2014) 228 – 236

Transportation Research Procedia

www.elsevier.com/locate/procedia

The Conference on Pedestrian and Evacuation Dynamics 2014 (PED2014)

Optimization-based pedestrian model calibration for evaluation

David Wolinski [a], Stephen J. Guy [b], Anne-Hélène Olivier [a], Ming C. Lin [c], Dinesh Manocha [c], Julien Pettré [a,*]

[a] Inria, Campus de Beaulieu, 263 Avenue Général Leclerc, 35042 Rennes, France
[b] University of Minnesota, 200 Union Street SE, Minneapolis MN 55455, USA
[c] University of North Carolina, Campus Box 3175, Brooks Computer Science Building, Chapel Hill N.C. 27599-3175, USA

Abstract

The evaluation and comparison of crowd simulation algorithms (complex, high-dimensional, multi-scale systems) is an important question. "Realism" being dependent on target applications, comparisons with real measurements are not easy. Promising solutions have been suggested for such evaluations (Guy et al. (2012)). Here, we address estimating simulation parameters before evaluating: what do evaluation results mean if the assessed model is not performing at its best? We propose an optimization-based approach encompassing: reference data, metrics, simulation algorithms and optimization techniques. We demonstrate finding good parameter values setting simulation results as close as possible to reference data, enabling fair and meaningful comparisons.

Peer-review under responsibility of Department of Transport & Planning Faculty of Civil Engineering and Geosciences
Delft University of Technology

Keywords: crowd simulation; evaluation; parameter estimation; optimization; ground-truth data

1. Introduction

Simulating crowds has recently received considerable attention with the introduction of a certain number of simulation algorithms. These include multi-agent algorithms that can be based on very different principles and are widely used in computer games, virtual reality, animation as well as pedestrian security.

With this increasing number of simulation algorithms, however, it becomes increasingly important to be able to evaluate them objectively and rigorously. One obvious way of doing that is comparing against real-world data. One more issue then becomes apparent, which is the case of the algorithms' parameters, which can often be tuned in order to obtain very different results. In such a case, one would want to compare the algorithms on the real-world datasets while being assured that the algorithms being compared perform to the best of their ability. This implies the use of the optimal set of parameter values leading to the best match between the simulation and data.

Methods to acquire the observed, ground-truth data are also becoming increasingly ubiquitous, and the data is becoming more readily available. With these increasing numbers of simulation algorithms and datasets, being able to automatically and rigorously compare the algorithms with respect to this data becomes also apparently important.

* Corresponding author. Tel.: +33 (0) 2 99 84 22 36; fax: +33 (0) 2 99 84 71 71.
E-mail address: julien.pettre@inria.fr

Peer-review under responsibility of Department of Transport & Planning Faculty of Civil Engineering and Geosciences Delft University of Technology
doi:10.1016/j.trpro.2014.09.041

Here, we present a framework that can be used to automatically evaluate simulation algorithms against various datasets. Of course, this evaluation includes using metrics to compare the algorithms to the data but also a broadly applicable scheme to estimate these algorithms' parameter values to best match the data. We formulate this process as an optimization problem where we find the optimal parameter values that minimize the error between a simulation and the data as given by the metrics. Then, the rest is merely a comparison of this error, as obtained with the different simulation algorithms. Our framework is general in this regard and many algorithms and metrics can be used or integrated.

We illustrate the usage and benefits of our framework on datasets of different natures such as recorded trajectories of individual pedestrians or macroscopic quantities such as fundamental diagrams. Our framework is open-source so that others can use it to evaluate crowd simulation algorithms with reference data. We demonstrate its performance on various multi-agent algorithms across various data with different numbers of agents. In our benchmarks, we observe that algorithms which use the information on the agents' instantaneous velocities (e.g. RVO2 (van den Berg et al. (2008)), Tangent (Pettré et al. (2009)), etc...) result in lower errors than techniques such as Boids (Reynolds (1987)) or Social-Force (Helbing and Molnar (1995)).

The rest of the paper is organized as follows. In Section 2, we give an overview on related work in crowd simulation, parameter calibration, and algorithm evaluation. Section 3 describes our parameter estimation framework and its key components: algorithms, metrics, reference data, and optimization techniques. Finally, concrete examples and applications are presented in section 4 to demonstrate the benefits of our solution.

2. Related work

With their applications to graphics, robotics or fire safety, numerous crowd simulation algorithms have emerged. Many of these choose to represent each individual agent in a crowd (e.g. as particles) and model the interactions between them. A very representative example of microscopic approaches is Reynolds' seminal Boids model (Reynolds (1987)). It aims to match an agent's speed to that of its neighbors and as a result, multiple behaviors emerge at the crowd level. This work was later extended to include more interactions such as herding and path following (Reynolds (1999)). Physically-based models formulate interactions as repulsive forces among agents (Helbing and Molnar (1995)) and can be combined with various rules to improve behaviors (Lamarche and Donikian (2004); Pelechano et al. (2007)). Lately, velocity-based algorithms have been introduced which aim to determine velocities for agents that are collision-free over certain future time-windows (van den Berg et al. (2008); Pettré et al. (2009); Karamouzas et al. (2009)). Other recent models include cognitive ones (Yu and Terzopoulos (2007)), affordance (Kapadia et al. (2009)), short-term planning with a discrete approach (Singh et al. (2011)), as well as a synthetic vision-based model using perceptual variables derived from the optic flow (Ondřej et al. (2010)).

Many other algorithms exist as well, such as algorithms operating at the macroscopic level (Treuille et al. (2006)). A major problem is then developing metrics to be able to compare these algorithms, which rely on very different principles and thus qualitatively very different (for instance velocity-based anticipation of future trajectories versus repulsive forces based on position alone).

A solution is to use experimental data for the evaluation; but in this case, parameters need to be taken into account. For instance, Pettré et al. (2009) set the parameter values of a collision-avoidance model based on data by using Maximum Likelihood Estimation. Lerner et al. (2009) compared local decisions taken from local context between simulations and real-world data. Lemercier et al. (2012) considered more complex situations where experimental data capturing both microscopic and macroscopic features of pedestrian motion was used to calibrate and compare various approaches.

Learning methods were also used to calibrate models from vision data (Pellegrini et al. (2009)), learn model parameters from real-world data (Charalambous and Chrysanthou (2010); Ju et al. (2010)), or learn agent motion from real-world motion (Lee et al. (2007); Lerner et al. (2007); Kim et al. (2012)). These approaches, while intrinsically realistic (based on real data), are also rather limited to specific scenarios. Furthermore, realistic portions of trajectories do not necessarily guarantee the realism of the compositions of the trajectories. In comparison, we present a technique to perform parameter estimation in a general, model-independent manner, considering various types of datasets.

In terms of evaluation, Kapadia and colleagues proposed a framework that evaluates the steering algorithms on different scenarios with respect to path smoothness or the number of collisions (Singh et al. (2009); Kapadia et al. (2011)). Other works also compared simulated trajectories with real-world data. The entropy-based metric (Guy et al.