Pattern Recognition Letters 44 (2014) 16–29

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Towards an integrated approach to crowd analysis and crowd synthesis: A case study and first results

Stefania Bandini [a], Andrea Gorrini [b], Giuseppe Vizzari [a,*]

[a] Complex Systems and Artificial Intelligence Research Centre, Università degli Studi di Milano–Bicocca, Milano, Italy
[b] Information Society Ph.D. Program, Department of Sociology and Social Research, Università degli Studi di Milano–Bicocca, Milano, Italy

ARTICLE INFO

Article history:
Available online 15 October 2013

Communicated by Rita Cucchiara

Keywords:
Crowd analysis
Crowd synthesis
Agent-based modelling and simulation
Groups

ABSTRACT

Studies related to *crowds* of pedestrians, both those of theoretical nature and application oriented ones, have generally focused on either the *analysis* or the *synthesis* of the phenomena related to the interplay between individual pedestrians, each characterised by goals, preferences and potentially relevant relationships with others, and the environment in which they are situated. The cases in which these activities have been systematically integrated for a mutual benefit are still very few compared to the corpus of crowd related literature. This paper presents a case study of an integrated approach to the definition of an innovative model for pedestrian and crowd simulation (on the side of synthesis) that was actually motivated and supported by the analyses of empirical data acquired from both experimental settings and observations in real world scenarios. In particular, we will introduce a model for the adaptive behaviour of pedestrians that are also members of groups, that strive to maintain their cohesion even in difficult (e.g., high density) situations. The paper will show how the synthesis phase also provided inputs to the analysis of empirical data, in a virtuous circle.

1. Introduction

The modelling and simulation of pedestrians and crowds is a consolidated and successful application of research results in the more general area of computer simulation of complex systems. It is an intrinsically interdisciplinary effort, with relevant contributions from disciplines ranging from physics and applied mathematics to computer science, often influenced by (and sometimes in collaboration with) anthropological, psychological, sociological studies and the humanities in general. The level of maturity of these approaches was sufficient to lead to the design and development of commercial software packages, offering useful and advanced functionalities to the end user (e.g., CAD integration, CAD-like functionalities, advanced visualisation and analysis tools) in addition to a simulation engine.[1] Nonetheless, as testified by a recent survey of the field by Schadschneider et al. (2009) and by a report commissioned by the Cabinet Office by Challenger et al. (2009), there is still much room for innovations in models improving their performances both in terms of *effectiveness* in modelling pedestrians and crowd phenomena, in terms of *expressiveness* of the models (i.e. simplifying the modelling activity or introducing the possibility of representing phenomena that were still not considered by existing approaches), and in terms of *efficiency* of the simulation tools. In addition to the above directions, we want to emphasise the fact that one of the sometimes overlooked aspects of a proper simulation project is related to the *calibration* and *validation* of the results of tools related to the *synthesis* of the pedestrians and crowd behaviour in the considered scenario. These phases are essentially related to the availability of proper empirical data about or, at least, relevant to, the considered scenario ranging from the pedestrian demand (a description of the points of origin and destination of pedestrians, the so-called "origin–destination matrix", plus an indication of the pedestrian arrival rate for all origins, and so on), preferences among different alternative movement choices (e.g., percentage of persons employing stairs, escalators and elevators in a multiple-level scenario), but also the average waiting times at service points (i.e. queues), the average time required to cover certain paths, the spatial distribution of pedestrians in specific environmental conditions that is required to evaluate the so-called "level of service" associated to portions of the environment as defined by Fruin (1971). These data are results of activities of *analysis*, some of which can be fruitfully automated given, on one hand, the wide diffusion of cameras employed for video surveillance of public areas and, on the other, considering the level of maturity of video processing and analysis techniques.

* Corresponding author. Tel.: +39 0264487865.
E-mail addresses: bandini@disco.unimib.it (S. Bandini), a.gorrini@campus.unimib.it (A. Gorrini), vizzari@disco.unimib.it (G. Vizzari).

[1] see http://www.evacmod.net/?q=node/5 – last checked September, 2013 for a large although not necessarily complete list of pedestrian simulation models and tools. The list comprises more than 60 models, of commercial and academic nature, general purpose or specifically designed for certain situations and scenarios, maintained or discontinued.

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.10.003

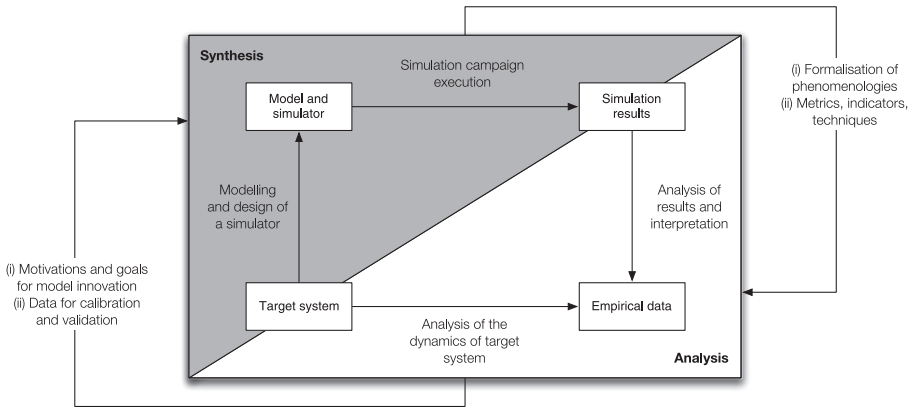

Fig. 1. A schema putting together activities of modelling and execution of simulation campaigns (synthesis) and activities related to the interpretation of simulation results, analysis of empirical data observed in reality and their comparison for sake of validation (analysis).

An *integrated approach* to pedestrians and crowd studies encompasses both the application of analysis and synthesis techniques that, in a virtuous circle, can mutually benefit one from the other, to effectively (i) identify, (ii) face and (iii) provide innovative solutions to challenges towards the improvement of the understanding of crowding phenomena (see Fig. 1). It must also be stressed that this virtuous circle is not merely a disciplinar enterprise, but it rather requires the fruitful interaction among various disciplinary efforts also from social sciences about human behaviour in crowded environments, such as the analysis of spatial social interaction among people. In particular, *proxemics* has a prominent role both in the modelling (see, e.g., Was, 2010 and in Manzoni et al. (2011)) and automated analysis Junior et al., 2010 of pedestrians and crowd behaviour. The term was first introduced by Hall (1966) with respect to the study of a type of nonverbal behaviour related to the dynamic regulation of interpersonal distance between people as they interact.

Within this context the aim of this paper is to provide a comprehensive framework comprising both the synthesis and analysis of pedestrians and crowd behaviour: in this schema we suggest both ways in which the results of the analysis can provide fruitful inputs to modellers and, on the other hand, how results of the modelling and simulation activities can contribute to the (automated) interpretation of raw empirical data. The following section will present relevant existing efforts that represent useful contributions covering part of this cycle, while Section 3 will describe and discuss the proposed framework. An example of unfolding of the conceptual and experimental pathways suggested by the framework will be described thanks to the introduction of an adaptive model for group cohesion (Section 4) that was motivated and that will be effectively calibrated and validated by means of analyses on observed crowd behaviour (Section 5). A discussion on how the two experiences of analysis and synthesis actually had a mutual benefit and they also produced an advancement beyond the current state of the art will follow. Conclusions and future works will end the paper.

2. Related work

The first type of collaboration between crowd analysis and synthesis exploits results of the former activity for supporting or improving modelling and simulation: first of all, empirical evidences are actually necessary for the identification of values or at least intervals for model parameters, but the observation of the crowded situations can also support the identification of phenomenologies not currently represented and managed by a model. On the other hand, the potential outcomes of synthesis activities that can have an influence on analysis are related to two categories of contributions: first of all, the definition of mechanisms for the generation of a given phenomenon leads to its formalisation, that can be extremely useful for the creation of mechanisms for its automated detection; moreover, the modeller and developer of a simulation system actually defines and develops metrics, indicators and techniques to evaluate the outcomes of his work, even before reaching the validation phase. These by-products can also represent a starting point for the actual development of automated analysis approaches. In the following, we will briefly present relevant works that fit the above schema and that are worth considering in the construction of an integrated framework for crowd analysis and synthesis.

2.1. Analysis as a support to synthesis

A relevant work in this category is the one described by Patil et al. (2011): the authors show how computational fields guiding simulated pedestrians movement in a simulated environment can be automatically derived by video footages of actual people moving in the same space. From the broader perspective of offering a useful service to crowd managers, works like the one by Georgoudas et al. (2011) and Wagoum et al., 2013 describe overall systems integrating computer vision techniques as input blocks to pedestrian simulation models able to suggest crowd management solutions (e.g., guidance signals) to avoid congestion situations in evacuation processes.

The so called *data driven* approach to crowd simulation can also be set in this category: Lee et al., 2007 present a first attempt of actually *learning* a simulation model from observations, a trend continued more recently by Junges and Klügl (2012). This kind of effort has provided useful suggestions especially in an attempt to give a more systematic characterisation of crowd dynamics but it is still relatively far from having a definitive impact of the (semi-) automatic definition of models, since results are not completely validated against empirical data in a sufficiently wide set of scenarios. This approach has highlighted the possibility of defining formal and computable metrics of similarity between different crowd dynamics, either generated by analyses or synthesis activities, for sake of their comparison and evaluation, as suggested by Lerner et al. (2009) and Guy et al., 2012. This possibility represents a useful validation element that can provide useful indications in selected and limited spatial/temporal sections of reference scenarios, for which abundant and high quality empirical data are available.

As previously mentioned, efforts in the analysis of pedestrians and crowd dynamics can also provide the motivations and goals for the innovation and extension of models. An example of