



11-2012

Pedestrian Anomaly Detection Using Context-Sensitive Crowd Simulation


Cory D Boatright
University of Pennsylvania

Mubbasir Kapadia
University of Pennsylvania

Jennie M. Shapira
University of Pennsylvania

Norman I. Badler
University of Pennsylvania, badler@seas.upenn.edu

Follow this and additional works at: <http://repository.upenn.edu/hms>

 Part of the [Engineering Commons](#), and the [Graphics and Human Computer Interfaces Commons](#)

Recommended Citation

Boatright, C., Kapadia, M., Shapira, J. M., & Badler, N. I. (2012). Pedestrian Anomaly Detection Using Context-Sensitive Crowd Simulation. *First International Workshop on Pattern Recognition and Crowd Analysis*, Retrieved from <http://repository.upenn.edu/hms/167>

Pedestrian Anomaly Detection Using Context-Sensitive Crowd Simulation

Abstract

Detecting anomalies in crowd movement is an area of considerable interest for surveillance and security applications. The question we address is: What constitutes an anomalous steering choice for an individual in the group? Deviation from “normal” behavior may be defined as a subject making a steering decision the observer would not, provided the same circumstances. Since the number of possible spatial and movement configurations is huge and human steering behavior is adaptive in nature, we adopt a context-sensitive approach to assess individuals rather than assume population-wide homogeneity. When presented with spatial trajectories from processed surveillance data, our system creates a shadow simulation. The simulation then establishes the current, local context for each agent and computes a predicted steering behavior against which the person’s actual motion can be statistically compared. We demonstrate the efficacy of our technique with preliminary results using real-world tracking data from the Edinburgh Pedestrian Dataset.

Disciplines

Computer Sciences | Engineering | Graphics and Human Computer Interfaces

Pedestrian Anomaly Detection Using Context-Sensitive Crowd Simulation

Cory D. Boatright

Mubbasir Kapadia

Jennie M. Shapira

Norman I. Badler

University of Pennsylvania

{coryb, mubbasir, jshapira, badler}@seas.upenn.edu

Abstract

1 *Detecting anomalies in crowd movement is an area of*
2 *considerable interest for surveillance and security appli-*
3 *cations. The question we address is: What constitutes an*
4 *anomalous steering choice for an individual in the group?*
5 *Deviation from “normal” behavior may be defined as a*
6 *subject making a steering decision the observer would*
7 *not, provided the same circumstances. Since the number*
8 *of possible spatial and movement configurations is huge*
9 *and human steering behavior is adaptive in nature, we*
10 *adopt a context-sensitive approach to assess individu-*
11 *als rather than assume population-wide homogeneity.*
12 *When presented with spatial trajectories from processed*
13 *surveillance data, our system creates a shadow simula-*
14 *tion. The simulation then establishes the current, local*
15 *context for each agent and computes a predicted steering*
16 *behavior against which the person’s actual motion can*
17 *be statistically compared. We demonstrate the efficacy of*
18 *our technique with preliminary results using real-world*
19 *tracking data from the Edinburgh Pedestrian Dataset.*

20 1. Introduction

21 Anomaly detection is increasingly important in mod-
22 ern security operations, which must observe increasing
23 numbers of people for suspicious behavior. By automat-
24 ing the detection of such behavior, we can lift the bur-
25 den on personnel and help focus their limited resources.
26 Anomaly detection remains an open research problem
27 because of the challenge in finding a model to serve as
28 the basis of normality while accommodating the diverse
29 range of human behavior. Previous efforts have used
30 such techniques as Gaussian Mixture Models and Hid-
31 den Markov Models to define how an average person may
32 act in a particular location with outliers being declared
33 anomalous. A more robust model of “normal” that prop-
34 erly reflects the qualitatively different situations a person
35 may experience is still needed.

36 Modeling human behavior is precisely the aim of
37 crowd simulation, making these two research endeavors
38 complementary. Data-driven approaches to simulation
39 in particular try to generalize the relationship between
40 environmental stimuli and a corresponding action, mak-
41 ing them a strong fit to this application. Training such
42 models on real-world data has presented problems with
43 the unpredictability of what will be observed, and subse-
44 quent disagreement of model and human is blamed on the
45 steering algorithm. However, with a high-quality model
46 it is reasonable to question which is truly abnormal. For
47 instance, an intoxicated person’s behavior would show
48 that the simulation model is not always at fault. With
49 an adequate simulation, we can analyze the behavior of
50 real people without artificially restricting expectations to
51 averages and other statistical figures.

52 We propose an anomaly detection system which uses
53 a simulation of “shadow agents” to represent real pedes-
54 trians. The system maintains a score for each person
55 according to deviations from their shadow agent’s nav-
56 igation. Our simulation uses a data-driven, compound
57 model of steering which dynamically adjusts each agent’s
58 decisions as its environment evolves from its own per-
59 spective. The idea of contexts for a crowd are not new,
60 but we extend this idea by allowing each individual to
61 determine its own context rather than setting a crowd-
62 wide context. This model of anomaly detection has sev-
63 eral advantages over other techniques. First, the system
64 permits a variety of appropriate behaviors co-existing to-
65 gether rather than assuming the agents are homogeneous.
66 Second, the system guards against the problem where
67 a small, early difference has unnecessarily large influ-
68 ence on the anomaly score by accumulating short-term
69 deviations. This metric depends on the validity of the
70 steering model used, be it our context-sensitive model
71 or any other algorithm. This framework simultaneously
72 checks both the population and the model’s accuracy,
73 as an overabundance of anomaly detections are strong
74 evidence of an inaccurate steering algorithm.

75 This paper makes the following contributions:

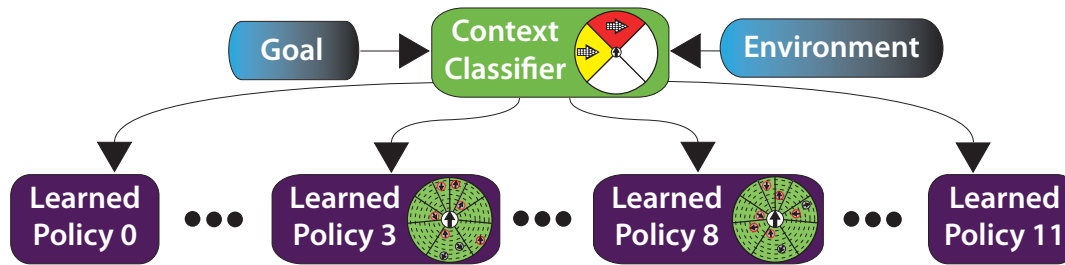


Figure 1. Our compound steering model dynamically chooses between classifiers based on the agent’s environment.

- A framework for detecting anomalous pedestrian trajectories in real-time which uses crowd simulation as the basis for comparison and is sensitive to the context each individual is experiencing rather than enforcing a group norm.
- A real-time, cumulative scoring model which is robust against late-starting anomalous behavior, does not artificially weight early decisions higher than those occurring later, and reveals inaccurate models when used on real data.

In Section 2 we frame this paper in the past work found in the literature. Section 3 gives more detail of our simulation, with the anomaly detection discussed in Section 4. Last, we give preliminary results of the technique in Section 5 with conclusions in Section 6.

2. Related Work

This paper proposes to bridge the gap between two areas of research: crowd simulation and anomaly detection in pedestrian movement. While we provide a review of the most applicable crowd literature, those interested in a more thorough survey of the field are directed to [17, 24]. Similarly we give a brief look at some of the common anomaly detection techniques, with further surveys of such work being [6, 4].

Crowd Simulation and Evaluation. Early crowd simulation [19] focused on agent throughput: getting many agents to move on screen and look like a group. In the quarter-century since that seminal work, the field has expanded and moved towards representing more complex dynamics. Emulation of the cognition behind human decision-making [26, 21, 1] has been an active area of research, and provides support for individual roles in the simulation.

In contrast to cognitive approaches, data-driven techniques [15, 11, 13] use machine-learning to map agent stimuli to actions. These techniques seek to fit a single model to the full spectrum of scenarios an agent may encounter through best-match databases. Other works [10, 16, 25] use clustering of their databases to account for the possible encounters which lead to different actions given the same stimuli.

Evaluation of crowds has often been by subjective observation, but statistical techniques have been proposed [7, 22, 8, 12, 9]. We leverage the concept of quantitative crowd metrics for our own anomaly detection system.

Anomaly Detection. In the interest of automated surveillance, computer vision has been interested in a variety of techniques and applications of anomaly detection. The most common technique is to use observations of a real population to fit a model of normal behavior. By focusing on the general flow of the crowds [5], these statistical models can then be used to detect high-level anomalous behavior such as an emergency evacuation [3]. Other works have focused on specific behavior of an individual, but not steering within a crowd [27, 20].

Comparison to the Literature. Both fields have acknowledged the problem of acquiring sufficient real-world data for training models and the potential for synthetic data in developing and training these systems [3, 18, 2]. This work is the realization of such suggestions, as we use an active crowd simulation as the model for normal behavior.

Furthermore, the model itself is egocentric, with each agent in the simulation capable of experiencing a different steering context from its neighbors. This is an extension to [12], where an entire crowd must be considered under the same context. Through the use of steering contexts and a hierarchical data-driven model, we avoid the single-model problem of defining a univer-

149 sally normal behavior for qualitatively different dynamic
150 environments.

151 3. Hierarchical Steering Model

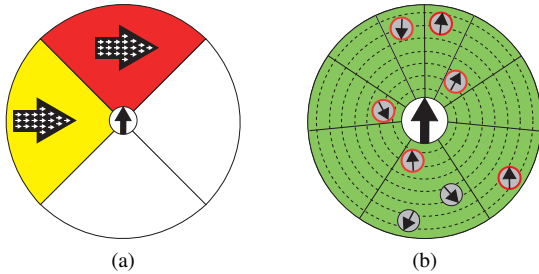


Figure 2. The environment is classified using long-horizon density and average trajectory tracked in each region, seen left. A shorter-range, more precise feature set seen right is used by the selected specialized model to decide the agent’s next action.

152 We use a compound machine-learned model for agent
153 steering, outlined in Figure (1). This model is constructed
154 by first identifying qualitatively different steering scenarios
155 an agent may encounter during a simulation, which we
156 call steering contexts. These contexts represent variation
157 such as cross traffic, oncoming traffic, and varying pop-
158 ulation densities. Each context has a specialized model
159 trained for it, and a top-level classifier is fit to take an
160 agent’s environment and decides which context model
161 should be used.

162 The action space for our model is discretized foot-
163 steps [23] and we use synthetic training data from a
164 short-horizon, space-time planner as a steering oracle
165 algorithm. Scenarios representing each context are
166 stochastically generated and the oracle’s decisions are
167 recorded. We then use the GPL C5.0 decision tree library
168 (www.rulequest.com) to train a model for each foot in
169 each context.

170 The features used in classifying a context focus on
171 general regional information, particularly each region’s
172 population and the average velocity of the agents present.
173 A second feature set is used for more precise measure-
174 ments of nearby agents. The area around the subject is
175 divided into slices with a higher resolution to the front to
176 simulate human vision. Each slice records the discretized
177 distance to the nearest agent as well as the agent’s rela-
178 tive velocity to the subject. Both sets are visualized in
179 Figure (2).

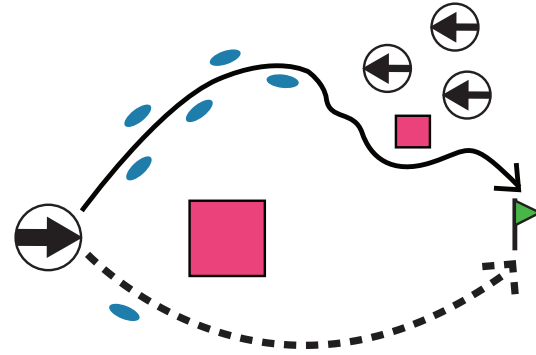


Figure 3. Shadow agents are forced to take the route of the person. After the first step above, the agent and person agree on the subsequent steering choices, reducing the likelihood of an anomaly.

180 4. Technique

181 Our system first creates a “shadow” agent in the sim-
182 ulation for each tracked person in the real world. Then
183 we calculate when the divergence between the two is
184 sufficient to merit flagging the behavior as anomalous.
185 Section 4.1 explains how our data-driven model for steer-
186 ing is converted into an observational tool applicable to
187 real humans. The calculation details are given in Sec-
188 tion 4.2.

189 4.1. The Shadow Simulation

190 Our system takes in tracked data of pedestrians and
191 extracts the necessary information for running a shadow
192 simulation. A shadow agent is created for each person,
193 with the person’s first tracked position and last tracked
194 position becoming the agent’s spawn and goal points,
195 respectively. The tracking data is also used to force the
196 shadow agent to follow the person’s path. Figure (3) il-
197 lustrates a person’s choice to turn left rather than right
198 having large consequences in the total trajectory as more
199 obstacles and people must be avoided to reach the goal.
200 Forcing the agent along the real path instead of simply
201 simulating the scene and comparing the resulting trajec-
202 tories nullifies inconsequential path diversity. With lim-
203 ited knowledge of each pedestrian’s internal state, such
204 singular differences are not sole indicators of anomalies.

205 At the beginning of each simulated footstep, the agent
206 uses the compound model from Section 3 to project its fu-
207 ture expected position. It also compares its current posi-
208 tion, which is the end of the previous footstep, against the
209 person’s real position. These measurements are used in

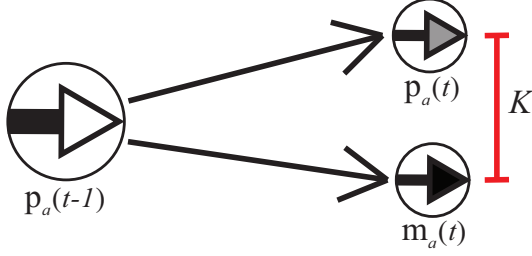


Figure 4. Regular comparison is made between the position of a person and that of its corresponding shadow agent in the virtual world.

210 Equation (1) to initiate an update of the agent’s anomaly
211 score.

212 Once indicated by a sufficiently high cumulative score,
213 the person is flagged as anomalous by the simulation.
214 This anomaly flag can optionally be removed with enough
215 subsequent expected behavior.

216 4.2. Flagging Anomalies

217 At each measurement time t , every agent a has two
218 positions, the real-world position $\mathbf{p}_a(t)$ and the position
219 indicated by the simulation $\mathbf{m}_a(t)$. We use the indicator
220 function in Equation (1) to decide whether or not the
221 deviation from one step to another is significant based
222 on difference kernel K . The tunable parameter d adjusts
223 the sensitivity of the system’s detection to allow for such
224 things as measurement error in the tracking data.

$$\mathbf{1}(a, t) = \begin{cases} 1 & \text{if } K(\mathbf{p}_a(t), \mathbf{m}_a(t)) \geq d \\ 0 & \text{else} \end{cases} \quad (1)$$

225 We let the variable $s_a(t)$ be the score for agent a at
226 time t . The value of $s_a(t)$ is defined according to Equa-
227 tion (2) where ω is a constant decay amount subtracted
228 from the score when normal behavior is observed, χ
229 is the confidence value of the shadow agent’s decision
230 from the compound model, and γ is set to reflect the
231 expected accuracy of the specialized classifier used for
232 this particular step. We constrain the value of $s_a(t)$ to
233 be nonnegative.

$$s_a(t) = \sum_{i=0}^t \chi(i) \gamma(i) \mathbf{1}(a, i) - \omega(1 - \mathbf{1}(a, i)) \quad (2)$$

234 Tuning ω adjusts the time window over which too
235 many deviations result in higher scores, with larger values

236 creating a more forgiving system. The benefit of this
237 decay-based accumulation function is that an anomaly
238 can start at any time and the score maintained as the
239 shadow agent moves through various contexts. This is
240 an improvement over using a finite time window, where
241 enough early normal behavior can dilute the ability to
242 detect late anomalies through an average score.

243 We define τ_{anom} to be the score threshold which indi-
244 cates anomalous behavior in a pedestrian. Additionally,
245 let $\tau_{\text{norm}} \leq \tau_{\text{anom}}$ be a score threshold which indicates
246 a return to normality. The latter is chosen to introduce
247 hysteresis in the detection system to prevent rapid tog-
248 gling of the anomaly flag. τ parameters can be chosen
249 together with ω to set a desired cooldown time.

250 Each agent then has a Boolean flag f_a which at time
251 t has the value set by Equation (3).

$$f_a(t) = \begin{cases} 1 & \text{if } s_a(t) \geq \tau_{\text{anom}} \\ f_a(t-1) & \text{if } \tau_{\text{norm}} < s_a(t) < \tau_{\text{anom}} \\ 0 & \text{if } s_a(t) \leq \tau_{\text{norm}} \end{cases} \quad (3)$$

252 5. Results

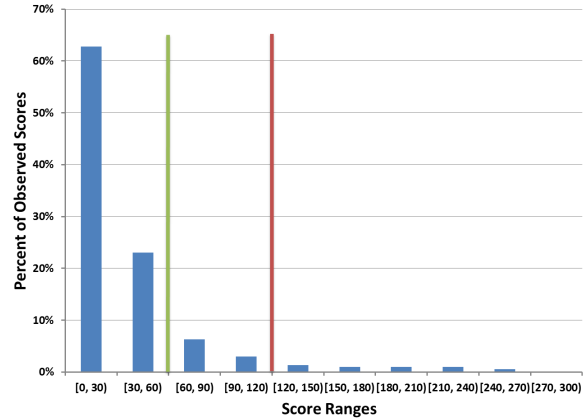


Figure 5. Histogram of score values from running the system used to find values for the anomaly and normality thresholds. The red and green lines are anomaly and normal thresholds, respectively.

253 To test our system, we used the Edinburgh Informatics
254 Forum Pedestrian Database [14]. Our compound steering
255 model consists of 12 contexts, with each context using
256 5000 sample scenarios to generate training data. An
257 additional 1000 sample scenarios were withheld for each
258 context as a validation set. The models were evaluated for
259 accuracy using this set to calculate our values for γ , seen

Context Number	0	1	2	3	4	5	6	7	8	9	10	11
γ	.79	.79	.80	.81	.80	.80	.80	.80	.81	.80	.79	.80

Table 1. The accuracy across the specialized classifiers is highly uniform, making no particular context a strength or weakness for the anomaly detection scores.

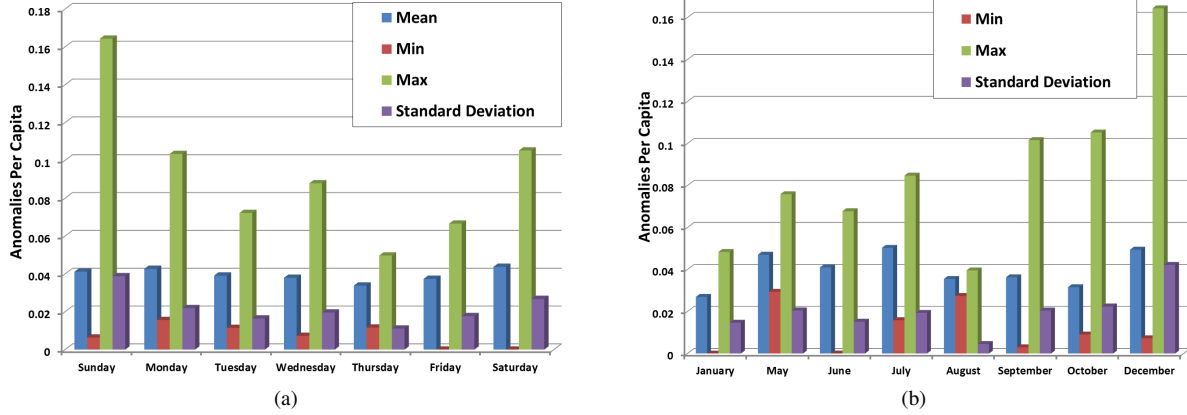


Figure 6. Statistical analyses of anomalies per capita for days of the week and months of the year.

260 in Table (1). A shadow simulation was created for each
 261 day of the database, and a histogram of anomaly scores
 262 generated using an ω value of 0.1. The distribution of
 263 scores can be seen in Figure (5) and strongly suggest the
 264 choice of 120 for τ_{anom} and 60 for τ_{norm} , owing to the
 265 small value for ω .

266 Figure (6) shows statistical analyses for the number
 267 of anomalies our system detected per capita for each
 268 day of the week and month of the year from the dataset.
 269 The population count varied greatly for each of the data
 270 points, ranging from 5 to 2804. However, the average
 271 anomalies per capita across the days and months re-
 272 mained consistent under our system, providing a val-
 273 idation of its robustness. We also note the weekend has
 274 a particularly high standard deviation for anomalies de-
 275 tected, indicative of the less uniform crowd flow dur-
 276 ing those days. Not all months were present in the dataset,
 277 and May consisted of only 3 days of tracking information.

278 Manual inspection of the simulation provided an inter-
 279 esting observation where we noticed anomalous agents
 280 under seemingly normal circumstances. On review of
 281 the dataset, we found that the floor can reflect the person,
 282 causing two agents to be spawned in the same location.
 283 In this case the agents continuously try to separate from
 284 each other but cannot, causing the high anomaly score.

285 6. Conclusions and Future Work

286 This paper presented an initial exploration into the
 287 use of a data-driven, context-sensitive crowd simulator
 288 for pedestrian anomaly detection. We used our prototype
 289 framework to examine the Edinburgh Dataset by report-
 290 ing the computed anomalies for the tracked pedestrian
 291 trajectories over 115 days.

292 We are actively exploring several avenues of future
 293 work. Our framework is fast, operating on a day of
 294 tracked data in minutes, suggesting potential for use in
 295 live surveillance. Our system is currently constrained to
 296 pedestrian movement, but we would also like to expand
 297 the contexts we use to include such things as small groups
 298 walking together to increase the quality of our algorithm
 299 and the breadth of its impact. Correlation-based metrics
 300 are another set of scoring techniques we could explore.
 301 An important validation of our technique will be to com-
 302 pare it against existing anomaly detection frameworks,
 303 such as the model provided with the dataset [14].

304 7. Acknowledgements

305 The research reported in this document/presentation
 306 was performed in connection with Contract Number
 307 W911NF-10-2-0016 with the U.S. Army Research Labo-
 308 ratory. The views and conclusions contained in this doc-
 309 ument/presentation are those of the authors and should
 310 not be interpreted as presenting the official policies or

311 position, either expressed or implied, of the U.S. Army 366
312 Research Laboratory, or the U.S. Government unless so 367
313 designated by other authorized documents. Citation of 368
314 manufacturer's or trade names does not constitute an offi- 369
315 cial endorsement or approval of the use thereof. The U.S. 370
316 Government is authorized to reproduce and distribute 371
317 reprints for Government purposes notwithstanding any 372
318 copyright notation heron. 373

319 We also thank Corey Novich for her assistance with 374
320 our illustrations. 375
376
377

321 References

322 [1] J. M. Allbeck. CAROSA: a tool for authoring NPCs. In 381
323 *MIG 2010*, pages 182–193. Springer, 2010. 382
324 [2] J. M. Allbeck and N. I. Badler. Distributed Video Sensor 383
325 Networks. In B. Bhanu, C. V. Ravishankar, A. K. 384
326 Roy-Chowdhury, H. Aghajan, and D. Terzopoulos, editors, 385
327 *Distributed Video Sensor Networks*, pages 193–205. 386
328 Springer London, London, 2011. 387
329 [3] E. L. Andrade, S. Blunsden, and R. B. Fisher. Modelling 388
330 Crowd Scenes for Event Detection. *ICPR '06*, pages 389
331 175–178, 2006. 390
332 [4] H. M. Dee and S. A. Velastin. How close are we to solving 391
333 the problem of automated visual surveillance? *Machine* 392
334 *Vision and Applications*, 19:329–343, May 2007. 393
335 [5] L. F. Henderson. The statistics of crowd fluids. *Nature*, 394
336 229:381–383, 1971. 395
337 [6] W. Hu, T. Tan, L. Wang, and S. Maybank. A survey on 396
338 visual surveillance of object motion and behaviors. *IEEE* 397
339 *Transactions on Systems, Man, and Cybernetics—Part* 398
340 *C: Applications and Reviews*, 34(3):334–352, 2004. 399
341 [7] M. Kapadia, S. Singh, B. Allen, G. Reinman, and 400
342 P. Faloutsos. Steerbug: An Interactive Framework for 401
343 Specifying and Detecting Steering Behaviors. In *SCA* 402
344 *2009*, volume 1, pages 209–216, 2009. 403
345 [8] M. Kapadia, S. Singh, and W. Hewlett. Egocentric afford- 404
346 ance fields in pedestrian steering. *ISD '09*, 1(212):215– 405
347 224, 2009. 406
348 [9] M. Kapadia, M. Wang, S. Singh, G. Reinman, and 407
349 P. Faloutsos. Scenario space: Characterizing coverage, 408
350 quality, and failure of steering algorithms. In *SCA 2011*, 409
351 pages 53–62. ACM, 2011. 410
352 [10] K. H. Lee, M. G. Choi, Q. Hong, and J. Lee. Group 411
353 behavior from video: a data-driven approach to crowd 412
354 simulation. In *SCA 2007*, volume 1, pages 109–118. 413
355 Eurographics Association, 2007. 414
356 [11] A. Lerner, Y. Chrysanthou, and D. Lischinski. Crowds 415
357 by Example. *Computer Graphics Forum*, 26(3):655–664, 416
358 Sept. 2007. 417
359 [12] A. Lerner, Y. Chrysanthou, A. Shamir, and D. Cohen- 418
360 Or. Context-Dependent Crowd Evaluation. *Computer* 419
361 *Graphics Forum*, 29(7):2197–2206, 2010. 420
362 [13] A. Lerner, E. Fitusi, Y. Chrysanthou, and D. Cohen-Or. 421
363 Fitting behaviors to pedestrian simulations. In *SCA 2009*, 422
364 volume 1, pages 199–208, New York, New York, USA, 423
365 2009. ACM Press. 424

[14] B. Majecka. *Statistical models of pedestrian behaviour in the Forum Master of Science Artificial Intelligence School of Informatics University of Edinburgh*. Masters, University of Edinburgh, 2009.
[15] R. A. Metoyer and J. K. Hodgins. Reactive pedestrian path following from examples. In *16th International Conference on Computer Animation and Social Agents*, volume 20, pages 149–156, Nov. 2003.
[16] S. R. Musse, C. R. Jung, J. C. S. Jacques Jr., and A. Braun. Using computer vision to simulate the motion of virtual agents. *Computer Animation and Virtual Worlds*, 18:83–93, 2007.
[17] N. Pelechano, J. Allbeck, and N. Badler. *Virtual Crowds: Methods, Simulation, and Control*. Morgan & Claypool, 2008.
[18] F. Z. Qureshi and D. Terzopoulos. Surveillance camera scheduling: A virtual vision approach. *Multimedia systems*, 12:269–283, 2006.
[19] C. W. Reynolds. Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, 21(4):25–34, Aug. 1987.
[20] P. Scovanner and M. F. Tappen. Learning pedestrian dynamics from the real world. In *2009 IEEE 12th International Conference on Computer Vision*, pages 381–388. Ieee, Sept. 2009.
[21] W. Shao and D. Terzopoulos. Autonomous pedestrians. *Graphical Models*, 69(5-6):246–274, Sept. 2007.
[22] S. Singh, M. Kapadia, P. Faloutsos, and G. Reinman. SteerBench: a benchmark suite for evaluating steering behaviors. *Computer Animation and Virtual Worlds*, 20:533–548, 2009.
[23] S. Singh, M. Kapadia, G. Reinman, and P. Faloutsos. Footstep navigation for dynamic crowds. *Computer Animation and Virtual Worlds*, 22:151–158, 2011.
[24] D. Thalmann and S. R. Musse. *Crowd Simulation*. Number May. Springer, 2007.
[25] P. Torrens, X. Li, and W. A. Griffin. Building Agent-Based Walking Models by Machine-Learning on Diverse Databases of Space-Time Trajectory Samples. *Transactions in GIS*, 15:67–94, July 2011.
[26] Q. Yu and D. Terzopoulos. A decision network framework for the behavioral animation of virtual humans. In *SCA 2007*, pages 119–128. Eurographics Association, 2007.
[27] H. Zhong, J. Shi, and M. Visontai. Detecting unusual activity in video. *CVPR 2004*, 2:819–826, 2004.