Virtual Vision: Visual Sensor Networks in Virtual Reality

Faisal Z. Qureshi*
Department of Computer Science
University of Toronto, Toronto, ON, Canada

Demetri Terzopoulos†
Computer Science Department
University of California, Los Angeles, CA, USA

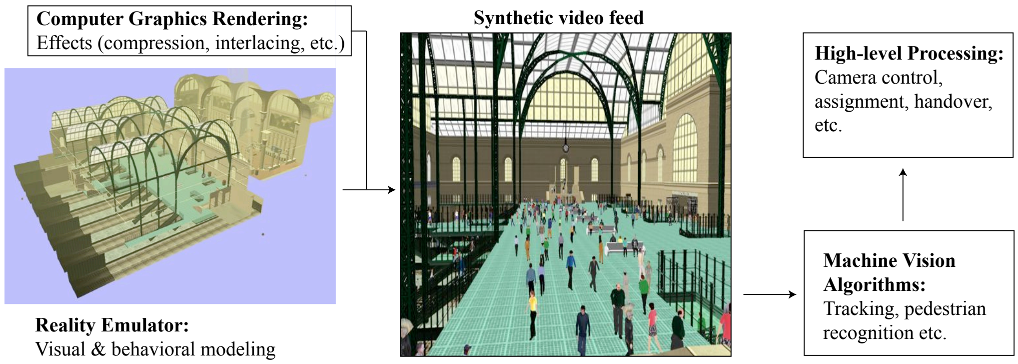

Figure 1: *The virtual vision paradigm.*

Abstract

The virtual vision paradigm features a unique synergy of computer graphics, artificial life, and computer vision technologies. Virtual vision prescribes visually and behaviorally realistic virtual environments as a simulation tool in support of research on large-scale visual sensor networks. Virtual vision has facilitated our research into developing multi-camera control and scheduling algorithms for next-generation smart video surveillance systems.

CR Categories: I.3.7 [Three-Dimensional Graphics and Realism]: Animation—Virtual Reality

Keywords: Virtual vision, reality emulator, smart cameras

1 Introduction

A computer simulated world that approaches the complexity and realism of the real world, inhabited by virtual humans that look, and move, and behave like real humans, could be used in revolutionary ways with profound impact across multiple scientific disciplines. In his VRST 2003 invited paper, Terzopoulos [2003] proposed the idea of employing such visually and behaviorally realistic "reality emulators" in designing machine vision systems, an approach called *Virtual Vision*. Shao and Terzopoulos [2005] implemented a prototype reality emulator, comprising a reconstructed model of the original Pennsylvania Train Station in New York City, populated with virtual pedestrians. The latter are autonomous agents with functional bodies and brains combining perceptual, behavioral, and cognitive components. The emulator can efficiently simulate well over a thousand self-animating pedestrians performing a rich variety of activities in the large-scale indoor urban environment.

Further developing the virtual vision paradigm over the past two years, we have employed this reality emulator in our research on camera sensor networks, advocating the design of simulated visual sensor networks and the meaningful experimentation with such simulated systems (Fig. 1). Camera networks are becoming increasingly important to next generation applications in visual surveillance, in environment and disaster monitoring, and in the military. In contrast to current video surveillance systems, camera sensor networks are characterized by smart cameras, large network sizes, and ad hoc deployment. To develop such systems, issues in machine vision and sensor networks must be addressed simultaneously. Virtual vision has enabled us to develop novel control strategies for smart camera systems capable of carrying out persistent visual surveillance tasks automatically or with minimal human intervention.

2 Motivation for Virtual Vision

Deploying a large-scale visual sensor network in the real world is a major undertaking whose cost can easily be prohibitive for most researchers interested in designing and experimenting with sensor networks. Moreover, privacy laws generally restrict the monitoring of people in public spaces for experimental purposes.

Legal impediments and cost considerations aside, the use of realistic virtual environments in sensor network research offer significantly greater flexibility during the design and evaluation cycle, thus expediting the engineering process. The virtual world provides readily accessible ground-truth data for the purposes of visual sensor network algorithm validation. Experiments are perfectly repeatable in the virtual world, so we can readily modify algorithms and parameters and immediately determine their effect. The hard real-time constraints of the real world can easily be relaxed in the simulated world. Finally, despite its sophistication, our simulator

*e-mail: faisal@cs.toronto.edu
†e-mail: dt@cs.ucla.edu

VRST 2007, Newport Beach, California, November 5–7, 2007.
© 2007 ACM 978-1-59593-863-3/07/0011 $5.00

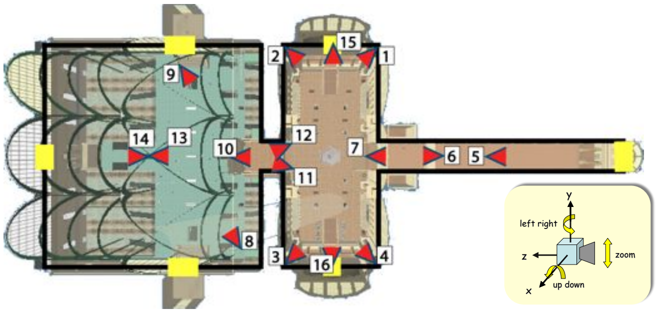

Figure 2: *Plan view of the virtual Penn Station environment with the roof not rendered, revealing the concourses and train tracks (left), the main waiting room (center), and the long shopping arcade (right). (The yellow rectangles indicate station pedestrian portals.) An example visual sensor network comprising 16 simulated active (pan-tilt-zoom) video surveillance cameras (lower right) is illustrated.*

runs on high-end commodity PCs, thus obviating the need to come to terms with special-purpose hardware and software.

3 Visual Sensor Network Research

Little attention has been paid to the problem of controlling/scheduling active cameras to provide intelligent visual coverage of a large public area, such as a train station or an airport. Our virtual vision approach has conveniently enabled us to carry out research that addresses this problem [Qureshi and Terzopoulos 2006]. Specifically, we have demonstrated visual sensor networks comprising multiple static and active simulated video surveillance cameras that provide perceptive coverage of a large virtual public space; in our case, a train station populated by autonomously self-animating virtual pedestrians (Fig. 1). The virtual cameras generate synthetic video feeds that emulate those generated by real surveillance cameras monitoring public spaces. We have developed a novel camera network control strategy that does not require camera calibration, a detailed world model, or a central controller. Our scheme is robust to node (camera) failures and communication errors. It enables a network of smart cameras to provide perceptive scene coverage and perform persistent surveillance with minimal human intervention.

We have tested our algorithms by deploying virtual camera networks comprising up to 16 uncalibrated active pan-tilt-zoom and passive cameras within a visually and behaviorally realistic train station simulator (Fig. 2). The simulation environment is populated with up to 100 self-animating pedestrians exhibiting realistic commuter behaviors. Fig. 3 demonstrates our distributed surveillance system following a pedestrian as she makes her way through the train station. Without any human intervention, cameras automatically collaborate to keep the pedestrian persistently in view despite occasional tracking failures. We have also demonstrated the ability of our system to successfully resolve camera assignment conflicts when multiple observation tasks are active.

The details of our virtual vision work are given in [Qureshi 2007].

4 Conclusion

We have developed a prototype intelligent surveillance system in a virtual train station environment populated by autonomous, lifelike pedestrians. This simulator facilitates our ability to design large-scale sensor networks in virtual reality and experiment with them on commodity personal computers. The future of such advanced simulation-based approaches appears promising for the purposes of low-cost design and experimentation.

Figure 3: *A pedestrian is persistently observed by a network of active surveillance cameras (see Fig. 2) as she makes her way through the train station to the concourse. (a-d) Cameras in rest configuration observing the station. (e) Operator selects a pedestrian in video feed 7. (f) Camera 7 has zoomed in on the pedestrian, (g) Camera 6, which is recruited by Camera 7, acquires the pedestrian. (h) Camera 6 zooms in on the pedestrian. (i) Camera 7 reverts to its default mode after losing track of the pedestrian—it is now ready for another task (j) Camera 6 has lost track of the pedestrian. (k) Camera 2. (l) Camera 2, which is recruited by Camera 6, acquires the pedestrian. (m) Camera 2 tracking the pedestrian. (n) Camera 3 is recruited by Camera 6; Camera 3 has acquired the pedestrian. (o) Camera 3 zooming in on the pedestrian. (p) Pedestrian is at the vending machine. (q) Pedestrian walking towards the concourse. (r) Camera 10 is recruited by Camera 3; Camera 10 tracks the pedestrian. (s) Camera 11 is recruited by Camera 10. (t) Camera 9 is recruited by Camera 10.*

Acknowledgements

The research reported herein was supported in part by a grant from the Defense Advanced Research Projects Agency (DARPA) of the Department of Defense. We thank Tom Strat, formerly of DARPA, for his generous support and encouragement. We also thank Wei Shao and Mauricio Plaza-Villegas for implementing the Penn Station simulator.

References

QURESHI, F.Z. 2007. *Intelligent Perception in Virtual Camera Networks and Space Robotics*. PhD thesis, Department of Computer Science, University of Toronto.

QURESHI, F.Z. AND TERZOPOULOS, D. 2006. Surveillance camera scheduling: A virtual vision approach. *ACM Multimedia Systems*, 12:269–283. Special Issue on ”Multimedia Surveillance Systems”.

SHAO, W., AND TERZOPOULOS, D. 2005. Autonomous pedestrians. In *Proc. ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, 19–28.

TERZOPOULOS, D. 2003. Perceptive agents and systems in virtual reality. In *Proc. 10th ACM Symposium on Virtual Reality Software and Technology*, 1–3.