

Rapid Prototyping of Parallel Robot Vision Systems Using Virtual Reality and Systems Simulation

Progress Report and Continuation Proposal, NSF CDA-9401142

Thomas LeBlanc, Dana Ballard, Christopher Brown, Randal Nelson, and Michael Scott
Computer Science Department, University of Rochester

June 1998

1 Continued Funding and Form 1328 Certification

Continued Funding is requested, as detailed in Section 5.2.

I certify that to the best of my knowledge (1) the statements herein (excluding scientific hypotheses and scientific opinions) are true and complete, and (2) the text and graphics in this report as well as any accompanying publications or other documents, unless otherwise indicated, are the original work of the signatories or individuals working under their supervision. I understand that the willful provision of false information or concealing a material fact in this report or any other communication submitted to NSF is a criminal offense (U.S. Code, Title 18, Section 1001.)

PI Signature:

2 Introduction

2.1 Institution

The University of Rochester is a small, private University established in 1850. During the early 20th century, the University grew significantly, in part due to the efforts of George Eastman, the founder of Eastman Kodak. During this period, the Medical School, the Institute of Optics, and the Eastman School of Music, all currently nationally known, were established. Today, the University is home to 4600 undergraduates and 2500 graduate students, and operates with a philosophy of providing the academic opportunities of a renowned research institution in an environment scaled to the individual.

The Department of Computer Science at the University of Rochester offers an intense, research-oriented program leading to the degree of Doctor of Philosophy, with particular emphasis on the areas of computer vision and robotics, knowledge representation and natural language understanding, systems software for parallel computing, and the theory of computation. A recently-inaugurated undergraduate major aims to capitalize on these research strengths. The focused research interests reflect our desire to achieve excellence in a core of important issues, rather than to try to cover all areas.

2.2 Project

The main focus of our research is a laboratory that combines sensory interaction with simulated physical environments (otherwise known as virtual reality), physical and sensory interaction with mixed physical and simulated environments (otherwise known as augmented reality), and the reconstruction of physical environments from sensory inputs (otherwise known as computer vision), with execution-driven simulation

of complex parallel systems in the design and control of visually-controlled robotic systems. We also develop systems tools to manage large parallel computing applications and development environments. This RI-supported research leverages Rochester's unique combination of expertise in active vision systems, behavioral robotics, virtual reality, and parallel programming environments and systems. (see the URL <http://www.cs.rochester.edu:80/research/iip/>).

Our laboratory has two parts: one is for building working systems in the *real world*; the other part is for prototyping and experimentation in the *virtual world*. The components of the real-world laboratory are the effectors and sensors for interacting in the real world, and the computational machinery required to run the control algorithms. The virtual-world laboratory includes the hardware and software for creating the virtual world, (models of sensors, effectors, and their interaction with an environment). A key innovation of the work is the ability to intermix real and virtual components, allowing us, for example, to use the same robotic control system in either a real or virtual environment. System support includes the large parallel computers used for computation-intensive applications and simulations, and research software systems for running them.

3 Year's Activities

3.1 Goals, Objectives and Targeted Activities

This year we wanted to integrate components into working demonstration systems as well as solve some difficult basic research issues. We wanted to extend one of our premiere demonstrations of simulation and real-world interaction by increasing the repertoire in the library of task-specific behaviors for a simulated autonomous vehicle capable of driving in a complex dynamic environment.

On the real-world side, our goals were to: (1) Commission mobile robots and integrate them into the simulation and control systems; (2) Develop and evaluate adaptive tracking systems; (3) Develop new multi-view 3D scene reconstruction algorithms for "virtualizing" complex physical objects and environments from photographs; (4) Parallelize our robust, general 3-D object recognizers, analyze their behavior, and integrate them into real-world mobile robotic systems.

In the virtual world, goals were to: (1) Close the loop using real-time visual perception and control algorithms in a realistic simulated world; (2) Use uncalibrated methods of fusing artificial graphics with real video (augmented reality); (3) Generate and manipulate photo-realistic, non-metric object models from video; (4) Build a brain-computer interface (BCI) using virtual reality environments to expand the capabilities of current BCI's by improving EEG signal recognition techniques; (5) Include "haptic virtual reality" in our simulations, using four Phantom haptic devices, two custom-made for increased working volume.

Systems support goals were to: (1) Develop an integrated compile-time and run-time system for efficient shared memory parallel computing on distributed memory machines (Cashmere and Treadmarks); (2) Support an interactive client-server computing model; (3) Apply results to applications in data mining and object recognition.

3.2 Components and Materials Required

In the virtual world, the upgrades to the SGI Onyx Infinite Reality were needed for adequate servo and refresh rates for the photo-realistic driving simulator, the Brain Computer Interface, and for the haptic virtual reality work. It computing needs of the haptic interface (1kHz servo rate) and visual refresh rate of 60 Hz.

In the real world, performance tests and validation of our statistical performance model were done using the 4-processor Enterprise 3000 (purchased with RI funds) with 2 Gigabytes of main memory. The size of main memory permitted experiments to be performed in parallel and in core, representing a speedup of approximately 40X over a standard desktop, which would require use of virtual memory to handle large databases. Since large-scale performance tests took several days to run, speed was critical to obtaining

the results that we did. It was also essential for testing our photorealistic reconstruction algorithms on geometrically-complex 3D objects, allowing the construction and manipulation of volumetric data sets containing more than 200 million voxels.

In systems support, our new AlphaServer will have about 4 times the aggregate compute power, 4 times the memory, half the communication latency, and 10 times the aggregate communication bandwidth of the current cluster. We need leading-edge machines to ensure the relevance and applicability of our computing assumptions, particularly as different architectural parameters change over time at different rates, thereby changing all the tradeoffs. It is crucial for one of our top research priorities in the coming year: a thorough evaluation of the design space for software distributed shared memory (S-DSM) on low-latency networks.

3.3 Indications of Success

More robotic, virtual reality, augmented reality, and parallel processing resources have been merged into our laboratory for research on intelligent action in real-world environments.

Haptic virtual reality, including an interactive interface for the control of multiple virtual reality experiments. We started several projects using the phantom device in motor control related tasks. E.g. we use a grasp and repositioning task to study behavior when the weight, density, and appearance of manipulated (virtual) objects are changed in unexpected ways. The augmented reality project created **augmented reality goggles**, using a Virtual Reality helmet fitted with two cameras.

Computerized object recognition improvements: 1) New algorithm that uses the object model to direct a search for object features, which in turn can be used to control a manipulator, so that we can command the system to "pick up the cup", say. 2) Additional performance tests of system have confirmed that it has the current world's best reported performance for a general 3-D recognition system robust to clutter and occlusion. 3) A statistical performance model accurately predicts recognition performance as the number of objects in the database grows, and as a function of the severity of clutter and occlusion.

VR-control interface: Extensions to the "automated driver" project, our first demonstration of real-time visual control using graphics video of a realistic virtual world. To last year's stop-light and stop-sign behaviors we added a car-following behavior based on looming detection in log-polar coordinates. All routines have been integrated into a driving program that controls a vehicle moving in traffic in a simulated town. The routines have also been tested on real video sequences, in order to show the generalizability of the simulation. Last, we plan to control the mobile wheelchair robots with the driving system, using wireless connectivity for TV outputs and command inputs.

Reconstruction of photorealistic 3D scene models: We are developing new 3D reconstruction techniques that take as input arbitrary collections of photographs of a 3D scene (e.g., multiple views around an unknown object or multiple views inside a room or inside an entire building) and generate a 3D scene model whose shape is consistent with all input photographs. Unlike previous work in shape-from-stereo, the resulting algorithms provably handle arbitrary camera and scene geometries and open up the possibility of turning complex real scenes into virtual environments for visualization and robot control simulation experiments.

Brain-Computer Interface: We are concentrating on a particular brain evoked potential (EP) phenomenon, the "P300", which is the potential that occurs 300 milliseconds after a rare and salient stimulus occurs. We need reliable detection of this EP to achieve reliable control, and of course we want to link the real and virtual worlds. Single trial P300 evoked potential recognition in a virtual environment now seems feasible and EEG signal recording is possible from within a VR helmet.

System support: Our Cashmere S-DSM system provides an efficient implementation of shared memory for programs running on a cluster of small-scale multiprocessors. Two innovations provide significant performance improvements over previous approaches: (1) a very low-latency network (the Digital Memory Channel) allows us to perform synchronization and directory updates very quickly; (2) a novel coherence protocol allows us to integrate hardware coherence within nodes with software coherence between nodes, in a highly asynchronous fashion. We are developing an interface to the S-DSM system for use by a paralleliz-

ing compiler. The compiler improves code efficiency by directing the run-time system to prefetch data and to move large blocks of data whenever possible, thereby avoiding significant amounts of fixed, per-block overhead.

Applications: (1) We designed and analyzed the performance of a parallel version of our object recognition algorithm, which uses a large database (several gigabytes) of object features. The parallelization follows a manager-worker model, where the work is distributed by the manager when needed. (2) Association data mining is the process of identifying commonly-occurring subsets in a database whose elements are sets. We found that for effective parallelization, attention needed to be paid to memory placement for improved locality and false sharing avoidance.

Altogether there were 31 refereed publications and 17 other publications directly related to the infrastructure. Six undergraduate and 25 graduate students, three of them women, made direct use of the infrastructure, including six who received the PhD this year.

4 Evaluation

NSF funding has allowed us to build a laboratory that has unique capabilities to support research in psychophysics, cognition, alternate human-computer interfaces (e.g. the BCI could enable people with diseases like Lou Gehrig's disease to live more independently), and photo-realistic and touch-realistic simulations for advanced computer control. Almost all our infrastructure funding has been spent on highly-leveraged, unusual hardware that helps us attract and retain the best students and faculty as well as carry out our work. We were pleased that after two years of infrastructure building we were able easily to develop several successful applications in which the real and simulated worlds were effectively married, realizing improved testing, performance, debugging, and human experimental interfaces.

4.1 Unmet Goals

A continuing problem is the lack of robotics and computer vision research hardware, namely of flexibly computer-controllable manipulators and flexibly-programmable vision hardware. We are luckily able to pursue much the same scientific objectives by using other equipment (the Phantoms, the mobile robots, etc.).

4.2 Outcome

The grant has helped us retain our position as a leader in active paradigm, biologically-motivated vision and robotics research, and in shared-memory parallel computing. It has helped us recruit and train the very best graduate students (consistently the best in the College on standardized tests), and place them at the best institutions upon their graduation.

At the national and global level, work supported by the grant has its impact through the influence of our papers on other researchers, and through the work of our alumni. Our work is heavily cited. Some of our laboratory techniques have been widely copied (e.g. verging binocular cameras) and others are still on the leading edge (eye trackers in VR helmets, oversize haptic hardware, uncalibrated augmented reality, real-time vision algorithms implemented on special hardware, hybrid hardware/software shared memory.)

4.3 Immediate Impact

Undergraduate students supported (and bachelor's degree if awarded in 1998): Henry McCauley, Yasser Mufti, Craig Harman, Elliot Barnett (BS), Kari Sortland (BA), Josh Drake (BA).

Graduate students supported (and master's degree if awarded): Zhenlei Cai (MS), Steve Haley (MS), Robert Stets, Sotirios Ioannidis (MS), Srinivasan Parthasarathy, Melissa Dominguez, Andrea Selinger (MS),

Jessica Bayliss, Garbis Salgian, Rodrigo Carceroni, Christopher Eveland (MS), Isaac Green, Greg Sharek, Galen Hunt, Yiyang Tao, Rahul Bhotika, Markus von der Heyde, Mike VanWie, Mohammed Zaki, Umit Rencuzogullari, Angkul Kongmunvattana.

1997-98 PhD graduates supported: Michal Cierniak, Intel Corp: "Optimizing Programs by Data and Control Transformations". Martin Jagersand, Yale University: "On-Line Estimation of Visual-Motor Models for Robot Control and Visual Simulation". Wagner Meira, UFMG, Brazil: "Understanding Parallel Program Performance Using Cause-Effect Analysis". Maged Michael, IBM T.J. Watson Research Center: "Reducing the Overhead of Sharing on Shared Memory Multiprocessors". Raj Rao, Salk Institute: "Dynamic Appearance-Based Vision". James Vallino, Rochester Institute of Technology: "Interactive Augmented Reality".

~

~