Keystone Mixed Reality

John Anderson, Jacky Baltes, Michael de Denus, Jeff Allen, and Daniel Troniak

Autonomous Agents Laboratory
Department of Computer Science
University of Manitoba
Winnipeg, Manitoba, Canada
http://aalab.cs.umanitoba.ca
andersj,jacky@cs.umanitoba.ca

Abstract. Keystone Mixed Reality is an applicant to the Physical Visualization (Mixed Reality) competitions at RoboCup-2008, from the Autonomous Agents Laboratory at the University of Manitoba, Canada. The team placed third in the competitions in 2007, and performed well across the full range of events. This paper outlines the current and planned future research using the mixed reality approach and the Citizen micro-robots, the improvements to league infrastructure that will be submitted, planned applications for the mixed reality platform, and the application of the approach to education at the team's home University.

1 Introduction

The Keystone Mixed Reality team competed in RoboCup-2007 in Atlanta, and placed third in the physical visualization subleague, which was the first time the Citizen micro-robots were officially used at RoboCup. Our achievement at RoboCup-2007 was important because it was the result of good performance in all three of the major elements comprising the competition (league development, innovative application, and soccer tournament).

Each of the elements of the competition closely matches work being done in our laboratory, which continues to stimulate our interest in the physical visualization (mixed reality) competition. While we use soccer as a challenge problem combining many important areas of artificial intelligence and robotics, we also do a broad range of work in artificial intelligence, multi-agent systems, robotics, and computer vision. The Citizen micro-robots allow for smaller playing fields and greater robot interaction than other platforms. The mixed reality approach also provides an ideal means of combining the challenges of physical robotics with the experimental control of simulation, resulting in an ideal research platform able to provide repeatable and quantitative results for robotics research.

We are also interested in developing better approaches to affordable robotics education (in terms of expense, learning curve, and expertise). By focusing on the mixed reality approach, we find that students are more motivated and can explore applications (e.g. those that involve altering the environment for the perception of others) in a much more effective way than physical robotics alone.

Moreover, the infrastructure for such educational approaches can, if approached correctly, be largely maintainable by the students themselves.

This paper describes our past and current work in mixed reality, including our plans for this year's competition. We outline our team background, and then describe our current and intended contributions to the league infrastructure, work toward innovative applications, and previous and proposed research using this infrastructure. We also describe our work in applying the citizen micro-robots and mixed reality platform for graduate and undergraduate education.

2 Background and Team Experience

The Autonomous Agents Laboratory at the University of Manitoba has been competing in RoboCup and other robotics competitions since 2002. Prior to this, Jacky Baltes, one of the directors of the laboratory, competed in RoboCup since 1999, while based in Auckland, NZ. In RoboCup, we have had entries in the small-size, E-league, humanoid, and rescue leagues, and have fielded up to three teams in any given year. We have won awards in the humanoid league, the e-league, and most recently, third place in the physical visualization subleague, the precursor to this year's competition.

We were also two of the founders of the U-league [1], which later became the E-league. This was a global-vision RoboCup league sharing many of the same goals as the current competition: to allow entry-level participation to a broad audience, while shifting the focus away from specialized team hardware and toward strong AI. We later promoted this approach for education [2], which has led to our current work on using the mixed reality approach for education.

Our laboratory focuses on four main research areas: computer vision, intelligent adaptive control, multi-agent systems, and humanoid robotics. All but the last are extremely relevant to the mixed reality competition, and this background also allows us to contribute to many areas in terms of advancing the state of the art

3 Contributions to League Infrastructure

The 2007 Citizen micro-robots were a revolutionary robot design because of their extremely small size. One of the promises of the physical visualization league was that we would be able to play 11 vs. 11 soccer on a reasonable sized playing field. However, further inspection showed that this dream will remain elusive a while longer. The two main issues currently preventing this level of play are:

1. The vision system is not robust and fast enough to track 22 robots and a ball over a large playing field;
2. The infrared transmitter used to send control commands from the clients to the robots is restricted to a small number of robots because of the methodology it employs.

We have therefore decided to put our efforts into solving these two problems by contributing two very important pieces of league infrastructure, and describe our approaches to these issues in subsections 3.1 and 3.2.

3.1 Ergo Vision System

The improvement in global vision can be made through the addition of our system, *Ergo* [3, 4]. While a functional vision system was provided for all teams in the initial year's competition, that vision system suffered from two problems that are common to almost all global vision systems. First, like most (e.g. [5, 6]), it assumed a perfectly mounted overhead camera, requiring a table organization that placed the camera perfectly above the table. This interacted with its second issue, a reliance on fine color calibration, to make its use under variable lighting conditions difficult. Teams in the practice area had to set up umbrellas or other shading devices to stop lighting reflections, and calibration of the system was nontrivial.

In contrast to this, our Ergo system (like its predecessor Doraemon, which is still used in the RoboCup small-sized league) allows a more flexible positioning of the camera. This is made possibly by using the standard Tsai camera calibration algorithm with our own calibration method [7]. Thus the camera can be moved to whatever angle and position minimizes reflections.

The left side of Figure 1 shows a sample image from Ergo which shows the use of an oblique angle (the upper right of the figure shows a low-resolution interpolation–a reconstructed overhead view–that is actually used for object tracking). More than this, the figure also shows the more significant advantage of using this system for vision: it is extremely robust under poor lighting conditions, and can handle variation in lighting conditions without significant or frequent recalibration.

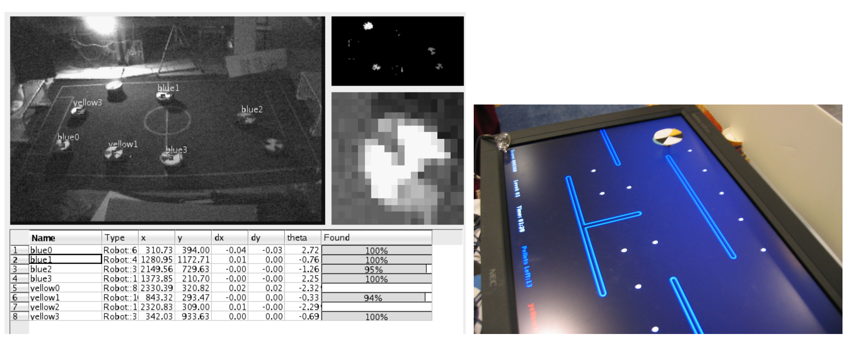

Fig. 1. Left: Using Ergo under very poor lighting conditions [3]; Right: Ergo hats used in Pac Man

Ergo does this by eliminating dependence on color calibration. Rather than performing direct color thresholding of camera images, Ergo thresholds for motion across pixels in each frame compared to a background image. An adaptation of $\Sigma\Delta$ background estimation [8] is used, which provides a computationally inexpensive means of recursively estimating the average color and variance of each pixel in a camera image.

Detecting motion involves setting a threshold above which variation across pixels will be considered to be motion. A combination of local and global thresholding is employed. A threshold is set for each pixel by examining the variance for each pixel in the background image, then a convolution is applied to consider a pixel's variance across its 9-pixel neighborhood. This local threshold is then scaled by a global threshold. To detect motion, each incoming image has its sum-squared error calculated across all pixels against the background image, the same convolution is applied to the result, and each value is compared to its corresponding pre-computed threshold. The use of the convolution has the effect of blending motion in small areas to eliminate noise, while making the movement of small objects such as the ball more obvious by also considering small changes in neighboring pixels. The individual motion pixels are then merged together into regions.

Objects to be identified and tracked, such as individual robots, still require a patterned marking, but we employ a bit-pattern represented by three different values: black, white, and *any color other than black and white* (right side of Figure 1. Two black segments are used to identify orientation, and also serve to allow the center of the tracked object to be found after these are identified.

Tracking objects involves taking areas of motion that fit the size of the patterned marking, examining these areas in the uninterpolated image, and looking for black and white areas. Other segments can be found and identified based on positions relative to the black orientation segments.

If two or more moving objects appear in close proximity to one another (or even partly occlude one another), motion analysis will view this as one large region of motion, with a center that will not be helpful in identifying anything. Ergo deals with this by tracking objects across visual frames and predicting their next position, which provides for a goal directed application of the object tracking algorithm. Any predicted objects recognized during this phase are masked during motion analysis, and may serve to remove one of a group of objects that may appear together in a moving region. In the case of multiple objects occupying the same space, the regions of interest will be those that are too large for any one object. If any of these regions were to contain more than one object, at least one recognizable object will be touching the edge of the region, and so the edge is where recognition efforts are focussed.

3.2 IrDA Transmitter Design

The second component we are contributing is the design of an infrared communications system. The first generation of the Citizen micro-robots (Eco-Be) were supplied with infrared transmitters; this will not be the case in the second generation, and a common design must be employed.

In developing a common design, we first note that the infrared transmitter used in 2007 proved to be a major bottleneck for the communication between the clients and the robots. In fact, some people may argue for the use of infrared as compared to RF technologies (413MHz, 900MHz, Bluetooth, WLAN). This would have major disadvantages: there is already much interference from the various RF technologies in use in RoboCup, and even if a channel could be devoted to the league, interference would limit the ability of teams to practice while games were being played. Infrared, on the other hand, can be easily blocked via a curtain.

There are several methods in which an infrared signal can be encoded and transmitted. These include Pulse Position Modulation (PPM), which is robust to interference but slow, and serial communication, which is still limited to 2400 baud. Neither of these is promising for controlling a team as large as 22 robots. In contrast, a new IR communication standard, *IrDA*, supports very high speed IR communication up to 16 MBits/sec with a planned extension to 100 MBits/second. The IrDA standard is able to achieve these high speeds by using a completely different encoding scheme where the timing of short IR bursts is used to encode logical zero and ones. We are developing an IrDA transmitter circuit for the physical visualization league. This is very important to support these robots, since reasonable team sizes are necessary to make the best use of this small platform.

There are several reasons why it is non-trivial to develop a transmitter for the physical visualization league based on IrDA. The IrDA standard assumes a point to point communication between two machines which establish a link after exchanging IDs and capabilities. In the physical visualization competition, we require a broadcast transmission from a server to many robots. Therefore, we cannot use the higher level link management IrLAN and IrCOMM protocols and are limited to using the IrPhys layer. In addition, the maximum range of an IrDA compatible device is 1m (but best results are achieved for distances shorter than 60cm) and the angle of an IrDA device is +/- 15 degrees of the center line. This angle is too small to cover the whole playing field given a transmitter mounted at a height of 60cm over the playing field, while the distance is fine for the current field but will have to be extended for larger screens.

There are a number of IrDA transceiver chips available. For example, the Microchip MCP 2150 implements a small IrDA stack in hardware and is a nice device for point to point communication. However, since the chip handles discovery and link management automatically, it is not suitable for our broadcast scenario. We therefore chose the Vishay IrDA transceiver which implements the IrDA physical layer specification. Our current design is based on a AVR ATMega128 microcontroller connected to a serial port. We plan on using a USB interface for our final version.

In addition to these components, we have also served as consultants from time to time on the design of the second generation of the Citizen robots.

4 Relationship to Current and Planned Research

The Citizen micro-robots and the mixed reality approach are central to a number of important ongoing research projects in our laboratory:

Adaptive Formation Control. The large population sizes supportable on a large flat-panel monitor, combined with our ability to track objects the size of Citizen Robots in real time provides possibly the best platform available today to do work in exploring physical robot formations. Most schemes use simulation, or smaller populations of robots in order to estimate the efficacy of formation control algorithms.

Grounded Communication in Multi-Agent Systems. Rather than assuming common coordinate systems such as GPS, we are working with teams of agents that develop their own consistent schemes for referring to locations in the domain during group actions [9]. We are currently working on using the mixed reality platform to display and alter landmarks to move to make this approach more robust in the face of errors in perception.

Imitation Learning. Learning by imitating others is an important type of social learning. We are working on an approach to imitation learning that allows the use of heterogeneous demonstrators [10]. Working with the Citizen robots is an important part of this approach, since their small size makes for significant issues in heterogeneity compared to other robots.

Robust Visual Tracking. As described above, robust vision is one of our major interests. We are working to improve Ergo for use in mixed-reality applications under poor/variable lighting conditions, including more work in efficient interpolation and goal-directed tracking

5 Innovative Applications

For applications, we prefer domains that allow students to explore innovative elements made possible by the mixed reality approach. In 2007, we developed a mixed-reality Pac Man game (right side of Figure 1, which both placed highly in this competition, and won a technical innovation award at the 2007 AAAI Robot Exhibition. This application both made use of physical and virtual elements (Pac Man can subsume virtual pellets and powerups) and physical elements (agents contact one another to score points by killing a ghost or killing Pac Man) and required both deliberative and reactive reasoning by multiple agents.

We have decided to base our current application around ice hockey, a game that is fast-moving and dealing with the physics surrounding movement on ice rather than solid ground. While a team-based sport like hockey might appear at first glance to have much in common with soccer, there are important difference. First, the play is with a stick, which affords a much finer level of agent-puck interaction than simply trapping or bumping a ball with the player's body. The puck must be in a certain position on the stick relative to the angle of stick placement and speed of both the robot and puck for the shot to be in the least accurate. The sticks added to soccer players by the University of Osaka to play with a physical ball showed the potential for stick-based games last year, while also demonstrating the far more difficult control problem in terms of accuracy using a stick to move an object rather than the robot itself. Our application will differ from this by using a virtual stick and puck, which will both serve to emphasize elements of both the physical and virtual while allowing a greater interaction between the robot and puck.

The second major difference is in the physics of the game. The puck, for example, will bounce off boards and will move much further than a physical ball could be pushed or shot by the robot itself. This requires much better prediction on the part of players as to where a shot will end up. From the perspective of the robots, the fast play also requires a significantly different physics. An ice hockey player cannot simply turn nearly as quickly to the side as a soccer player. The game also has interaction between players beyond soccer: players can legitimately interfere with one another, and physics will have to be in place for these much more complex player interactions.

The issue of a more constrained physics in hockey compared to soccer requires placing constraints on the command server. For example, the server would have to refuse to send turn commands when the robot is moving faster than constraints on turning would allow at the robot's given speed. We also intend our approach to this to be generic, in the sense that further constraints can be imposed or relaxed. Thus, our application will be designed to allow a broad range of games to be programmed and played, through the use of constraints placed on how players can act on the environment and vice versa. In addition to these constraints, virtual actions will be definable, so that a slap-shot that raises the puck off the ground can be differentiated from a more basic wrist-shot pass, for example. Families of such actions will allow combinations of puck angle and arm placement for the robots, and will be customizable for various types of games. A set of such constraints would then be used to reprogram the system from one game to another.

These three components: modifiable command server constraints, virtual actions, and XML-based descriptions, currently do not exist within the PV architecture, and thus can also be considered architectural contributions.

6 Educational Approach

An important element of all of our work in mixed reality is that we expect it to be applicable to teaching students. The course for which we employ this approach is a fourth year course involving a small set of students (12–15) working in groups [11, 4]. We begin by covering basic concepts in vision (e.g. color models, perspective geometry) while students learn to use the mixed reality environment in a laboratory setting. They then write an interface to control the robots manually, while learning about control algorithms (e.g. fuzzy logic controllers, Egerstedt's and Balluchi's controllers). Students then implement these control algorithms to run a series of laps on a racetrack to implement path following, while covering path planning methods (e.g. quad-tree decomposition)

Students then embody the latter topic by performing a treasure hunt in the virtual world, where a series of locations must be visited using path planning, while students cover dynamic path planning in class. Students apply this knowledge by running races across the field, where moving virtual must be avoided, while larger agent architectures are covered in class (e.g. behavior-based approaches). Each of these steps involves applying the skills learned at the previous stage, and at this point students can demonstrate sophisticated interacting behaviors, such as passing a physical ball between quadrants marked on the virtual world, and playing simple games involving obstacle avoidance (such as Pong). Finally, these are combined into a capstone assignment that requires combining all the skills they have learned, such as two-on-two soccer or Pac Man (Fig. 1).

We find that the mixed reality approach is ideal for capturing the students' imaginations over and above working with robots in general.

References

1. Anderson, J., Baltes, J., Livingston, D., Sklar, E.: Toward an undergraduate league for RoboCup. In: Proceedings of RoboCup-2003, Padova, Italy (2003)
2. Baltes, J., Sklar, E., Anderson, J.: Teaching with RoboCup. In: Accessible Hands-on Artificial Intelligence and Robotics Education. AAAI Spring Symposium Series, Stanford, CA (2004) 146 – 152
3. Furgale, P., Anderson, J., Baltes, J.: Real-time vision-based pattern tracking without predefined colors. In: Proceedings of CIRAS-2005), Singapore (2005)
4. Anderson, J., Baltes, J.: A pragmatic global vision system for educational robotics. In: Robots and Robot Venues: Resources for AI Education. AAAI Spring Symposium Series, Stanford, CA (2007) 1–6
5. Bruce, J., Veloso, M.: Fast and accurate vision-based pattern detection and identification. In: Proceedings of IEEE ICRA-2003, Taipei, Taiwan (2003) 567–570
6. Ball, D., Wyeth, G., Nuske, S.: A global vision system for a robot soccer team. In: Proceedings of the 2004 Australasian Conference on Robotics and Automation (ACRA). (2004)
7. Tsai, R.Y.: An efficient and accurate camera calibration technique for 3d machine vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL (1986) 364–374
8. Manzanera, A., Richefeu, J.: A robust and computationally efficient motion detection algorithm based on sigma-delta background estimation. In: Proceedings of the 4th Indian Conference on Computer Vision, Graphics and Image Processing, Kolkata, India (2004) 46–51
9. Wiebe, N., Anderson, J.: A local approach to developing grounded spatial references in multi-robot systems. In: Proceedings of IEEE IROS 2007, San Diego (2007) 1357–1364
10. Allen, J., Anderson, J.: A vision-based approach to imitation using heterogeneous demonstrators. In: Proceedings of the AAAI Workshop on Plan and Intent Recognition, Vancouver (2007) 9–16
11. Anderson, J., Baltes, J.: A mixed reality approach to undergraduate robotics education. In Holte, R., Howe, A., eds.: Proceedings of AAAI-07 (Robot Exhibition Papers), Vancouver, Canada, AAAI Press (2007)