6.191 Thesis Proposal: Intelligent Camera and 6.270

Jonathan Lee

December 7, 2001

1 Computer Graphics Group

The Computer Graphics Group (CGG) at the Laboratory for Computer Science (LCS) currently has several research groups dedicated on the latest graphics technology. Professor Leonard McMillan oversees many of the projects being researched, including Projective Drawing, Real-Time Visual Hull, and Computational Video.

2 The Problem: 6.270 and Machine Vision

6.270 is a hands-on, learn-by-doing class open only to MIT students, in which participants design and build a robot that will play in a competition at the end of January. The goal for the students is to design a machine that will be able to navigate its way around the playing surface, recognize other opponents, and manipulate game objects. 6.270 robots are totally autonomous, so once a round begins, there is no human intervention.

A robot can sense its surroundings through different types of sensors. Physical switches can act as bumper that allow the robot to feel walls, game pieces, and its opponent. CdS cells and phototransistors are used to detect colors. Several of these sensors can be used simultaneously to help the robot navigate by detecting lines painted on the game board. Distance sensors are used to detect the space between the rsensor and an object or a wall. Shaft encoders can be used to determine the distance travelled by the robot, and servos help determine the robots direction. Potentiometers can take analog measurements, such as how much the wheels underneath the robot have rotated. All robots are required to have an infrared (IR) beacon transmitter and receiver. This allows a competing robot to locate its opponent.

Unfortunately, all these sensors still cannot accurately depict the situation the robot is in. For example, a robot may have a shaft encoder that counts the number of revolutions a tire makes to determine distance travelled. If during the course of a round the robot runs into its opponent and gets wedged off the ground, the tire could possibly still be turning, telling the robot that it is still moving, when in fact it is stuck against the other robot. The IR beacon alone can only detect the opponent's direction, but not its distance, nor what possible obstacles are in the way, such as a plateau or a game piece.

Machine vision could help the robot understand its surroundings better by giving the controller board a more accurate depiction of where objects lie and where the robot's opponent is located. A camera mounted on the robot could also locate objects more quickly and navigate the board dextrously. Color recognition can also be done more robustly, which would help in games in which colors of different objects represent different point values.

Several long-term goals exist for this thesis:

- Currently the 6.270 robots use the Handyboard controller board, which does not provide enough computing power to process a real-time video feed. A new hardware architecture must be found.

- Image-processing algorithms must be chosen and coded. Depending on the hardware solution, the language needed to program the algorithms may create unwelcome software overhead.

- A useful user interface (UI) must be created so that students can take advantage of the camera's capabilities. Because students who compete in 6.270 have never taken the course before and are not expected to have any prior knowledge about robotics, programming, or engineering, simplicity and ease-of-use are mandatory.

- The robots in the end should exhibit more intelligent behavior. They should be able to sense their surroundings better. The robots should also be able to become more competitive. It is possible that the existing sensors can direct the camera to help the robot navigate. One long-term hope is to get rid of some of the sensors completely, such as the IR beacon.

3 Proposal

Several possible hardware designs have been proposed. One possibility would be to use the IPAQ. The PDA's OS is Linux, and contains two sleeves. One of the sleeves would be occupied by a FPGA that processes the images that come from the camera. It would require, however, the author to learn VHDL to program the vision algorithms.

Another possibility would be to use the Bitsy card. Funding is a problem here: the Bitsy card costs on the order of thousands. The camera also provides limited capability, since it can send images between 5 and 30 frames per second. The card is also bulky.

Professor McMillan had an idea of an intelligent camera that takes advantage of the XESS' Prototype Boards. This prototype board has a 100 MHz oscillator and an 32-KByte 8-bit SRAM. It uses the XC95108 CPLD from XILINX, and a 8031 microcontroller. The prototype board contains a parallel port connection to the PC, and the user can download code to the CPLD. This way the camera can be programmed to process the images. One side of the board, however, has PS/2 and VGA monitor ports that are needed by the camera. The first part of the thesis would involve redesigning the prototype board without these connections, and instead include an imager. Other tweaks to the board can be discussed and planned before the design phase begins.

The second part of the thesis involves software development, although that has not been discussed and finalized yet. It could involve creating an SDK for the camera, or delving into the algorithms themselves. The SDK could be tailored for 6.270, or for more general use.

This Intelligent Camera could be used in many applications. Professor McMillan had this idea for a while, and was looking for a student to design the board. This thesis proposal would provide for CGG the camera for general application use. For the author, the camera could be used to solve the problem regarding low-resource machine vision. If the camera works well, a further continuation of the thesis could merge the functionality of the Handyboard with the camera.

4 Resources

A Prototype Board from XESS would be needed for investigation, and a computer workstation to design the circuit board would be needed. Financial funding would purchase the appropriate XILINX CPLD and Kodak or Motorola imager, as well as fabricate and assemble the circuit board. Although Professor McMillan has no problem with funding, he prefers to take students first as UROPs to see their work ethic. Due to the nature of the author's schedule, he may be required to TA a class for at least one term.

The author also needs to acquire the appropriate knowledge to complete the project. Being a computer science major, having to design hardware is a challenge. Fall 2002 the author plans to take 6.111 to familiarize with hardware and circuit design, as well as machine vision to assist with the software portion of the project. He has already taken 18.06, 6.046, and 6.837, which provide a solid foundation for development of the image algorithms.

Other professors in other groups can also provide much help. Trevor Darrell of the Vision Interface Group works on more complicated video recognition algorithms than what is required here, but could provide useful help with software algorithms. Dr. Bove of the Object-Based Media Group has developed similar hardware with the Eye Society project. Jamie Hicks of COMPAQ Research Labs assisted in the development of the Bitsy board, and could help with the hardware research and design.

5 Timeline

Because the author will be taking the spring 2002 semester off, his schedule is pushed back effectively a whole year. This gives him three semesters to work on the project, two summers, and two IAPs. By summer of 2003, the hardware implementation should be completed, and the board fabricated. By IAP 2004, the software side of the project should be finished, and the following spring term can be used to write the thesis.

6 Risks

This project is not set in stone, and so it is possible that Professor McMillan could back out of this project if a more qualified candidate takes on this project. The author still, however, wishes to do a project based on machine vision and graphics, so he may have to look around for projects unrelated to 6.270 but still relevant to the field.