2006-2357: STEREO VISION ON A SMART ROVER

Zekeriya Aliyazicioglu, California State Polytechnic University-Pomona

Kathleen Hayden, California State Polytechnic University-Pomona

Yi Cheng, California State Polytechnic University-Pomona

Tim Lin, California State Polytechnic University-Pomona

Stereo Vision on a Smart Rover

1. Introduction

In 2002, California State Polytechnic University, Pomona (Cal Poly Pomona) was awarded NASA PAIR (Partnership Awards for the Integration of Research into Undergraduate Education) contract. The purpose of this four years NASA PAIR program is to integrate cutting-edge NASA-related research into the undergraduate curriculum. Cal Poly Pomona chose to incorporate the Jet Propulsion Laboratories (JPL) robotic technology research into the undergraduate curricula of the Electrical and Computer Engineering Department, the Engineering Technology Department, Mechanical Engineering Department, and the Computer Science Department. We proposed to conduct an interdisciplinary project, "Deep Space Exploration using Smart Robotic Rovers", and develop an autonomous robotic rover.
During the last three years, students and faculty participating in this program have developed a robotic rover that has successfully accomplished the initial goals of the project: (1) semi-autonomous navigation systems for remote robots, (2) processing of 2-dimensional images and data transmission, and (3) modified operations in a degraded communication environment. At this time the rover is capable of climbing 30° inclines, rotating about its center axis, and maneuvering diagonally while maintaining stability. It has also been designed to protect the vital internal components from outside contaminants and provides mechanical support for all externally mounted equipment including sensors, a robotic arm and a stereo camera.

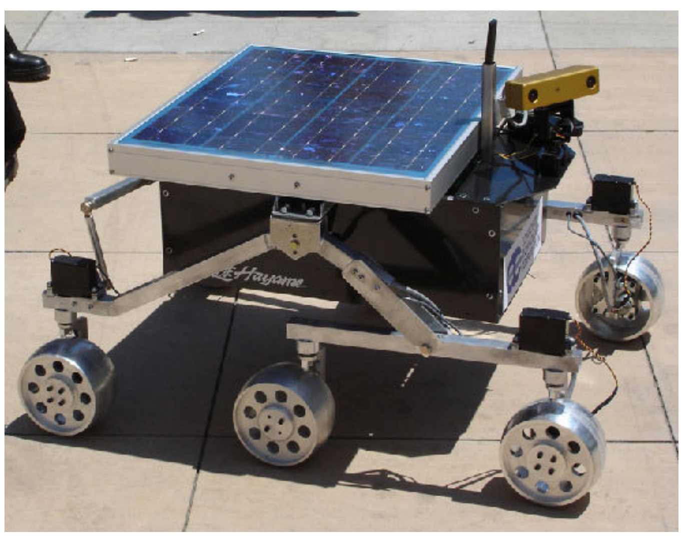

Figure 1. Smart Robotic Rover Prototype

The robot uses images captured by the stereo camera to move from one location to another location while avoiding any possible obstructions in its path. Stereographic images are used to garner depth information of objects in the path of the rover. Currently, target recognition, the ability of identifying and locating a designated target object, is under investigation.

Our multi-disciplinary team is using a Point Grey Bumblebee stereovision camera on the rover to perform image processing and assist the rover in autonomously navigating through its designated terrain. Our initial investigation of the stereographic images captured by the camera, has spawned new research areas, which include target size and distance measurement, shadow effects, and geographical data collection and analysis.

2. Educational Objectives

Interdisciplinary teams are composed of faculty and students from the Electrical and Computer Engineering Department, the Computer Science Department, the Engineering Technology Department and the Mechanical Engineering Department. We are engaged in the development of a robotic rover, capable of being deployed in an unfriendly environment, equipped with an autonomous navigational system, designed to collect scientific data and capable of communicating through a wireless network with the base station during its deployment. Each year junior and senior students are recruited to participate in this effort. Participating students are asked to actively engage in the project activities for six quarters, three while they are juniors and three while they are seniors. Through their participation and contributions towards the mission of this project, students receive degree credit. Juniors are awarded "Upper Division Elective Credit" and seniors fulfill their degree capstone requirement, "Team Senior Project". Each year more than 70 students and 14 faculties have participated in this effort. The students are sub-divided into team specializing in stereovision, object recognition, hardware, mapping and path-planning software, camera on a chip design, and GIS. Each sub-team has one or more faculty advisors to supervise the team's activities.

Students enrolled in the project agree to commit a minimum of six hours per week to the project. Sub-team meetings are held each week of the quarter. All teams meet together three times each quarter to present their progress towards the goals specified for the sub-team and to coordinate events and tasks for the project. Each spring, the College of Engineering holds a "Symposium Day" to showcase the activities of all "Team Senior Projects" that have been completed during that year. Each sub-team in the NASA PAIR group presents their accomplishments and demonstrates to the invited students, faculty members, and industrial guests including members of the technical staff from JPL.

Students are selected by their interest and academic training which includes software programming, microcontroller design and interfacing, FPGA design and control systems. All engineering students participating in the project have completed a minimum of 2 years of calculus, one year of physics, several courses in electrical engineering including programming, electronics, circuit analysis and control systems. As members of the NASA PAIR Team, they are encouraged to take additional elective classes in related fields such that they may make greater contributions to the project. Many students take upper division classes in Robotics, Advanced DSP, Microprocessor Applications in Process Control, Digital Design using Verilog HDL and Computer Networks.

Students are challenged to envision a system solution requiring contributions from many different disciplines. As a result they are exposed to topics, tools and technology that would not otherwise be a part of their undergraduate training. They benefit from participating as a team member on this project by applying the knowledge they have gained in the classroom to solving real-world engineering problems. Additionally, they benefit by interacting with students and faculty from other disciplines such that they acquire a greater appreciation of both the technical and social complexities of participating in a team effort to solve a complex problem.

3. System Architecture of the Rover

The mission requirements for a vision-based control system are: (1) to move the rover from one specified location to another specified location autonomously, (2) to move the robot to a specified target using an on-board digital camera and computer, and (3) to avoid obstacles in the path of the rover without human intervention

Figure 2 depicts the system diagram of our smart rover including the stereo camera. The main processor consist of a Mini-ITX motherboard used to collect and compute the disparity image map, calculate a path between two specified locations, avoid obstacles in the path of the rover and to locate a target within the terrain. The FPGA is used to monitor information collected from the externally mounted sensors, receives movement commands from the PC and controls all propulsion elements of the rover. Ultimately, it is the responsibility of the FPGA, based on the commands sent by the PC and the information collected from the external sensors, to either permit movement or enforce an emergency halt due to the detection of an obstacle not detected by the image collection and processing system.

When employing the target recognition system, the rover searches the terrain for a specified target and then moves towards the target; it will detect obstacles along its path and circumnavigate around these obstructions. Obstacle avoidance will be accomplished by moving to the left or right of the obstacle. Each time the rover encounters and obstacle and deviates from the original path a new path will be automatically calculated and the rover will continue towards its initial target destination. The robot will automatically stop when it nears the target. The system does not require the position coordinates of the target, nor does it require real-time position measurements. In order to move the robot to the target, the controller attempts to align and maintain the target image at the center of the camera's focus while progressing towards the target.

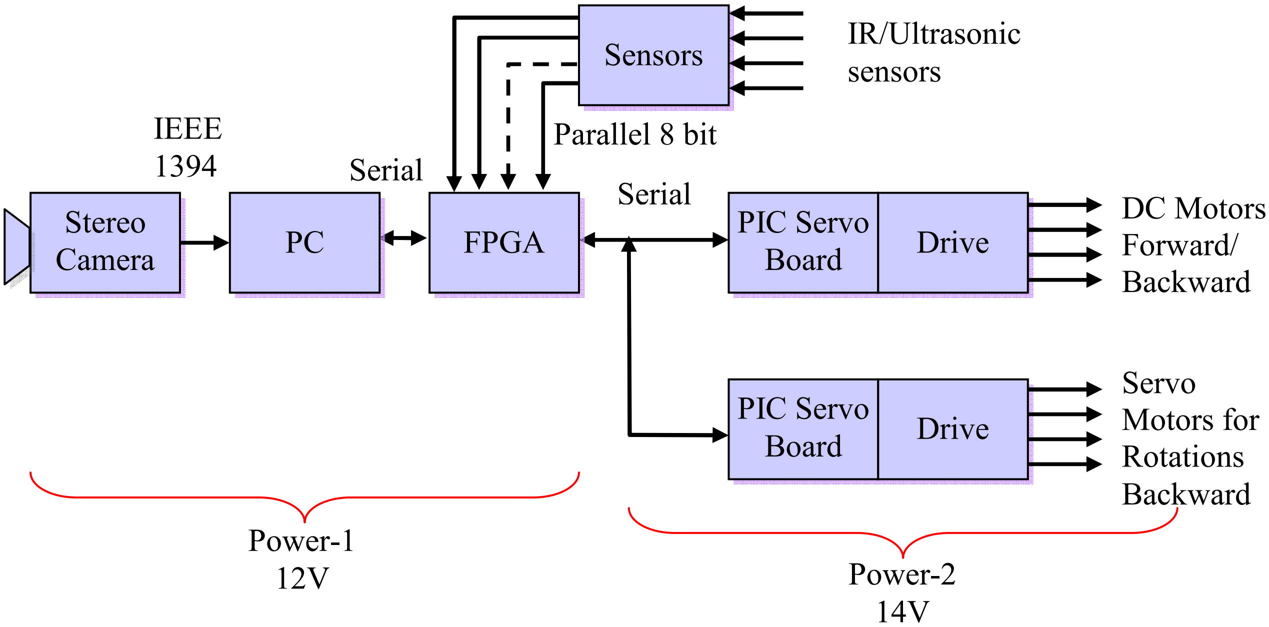

Figure 2. System diagram of A Smart rover with Stereo Vision

The PC receives the image data from the stereovision camera via a FireWire IEEE 1394 connection. It then locates the target object in the image and plots a path towards the target. As the rover proceeds towards the target it continues to collect image data to align its path towards the target and to avoid obstacles detected in its path. The processing unit communicates the calculated path for the rover via a set to movement commands to the FPGA. The FPGA and PC communicate through the PC serial port. The FPGA receives the PC movement commands and compares these requests to the information collected from the external sensors. If path communicated to the FPGA is free of obstacles then it will attempt to maneuver the rover as directed by the PC. However, if obstacles are detected by the external sensors but not recognized in the image data then the FPGA will make the final decision as to whether the rover can continue on the directed path or halt movement and communicate the status to the PC.

Stereovision camera

For this project we selected the Bumblebee Stereo Vision camera made by Point Grey. The camera is ideal for applications such as people tracking, gesture recognition, mobile robotics and other computer vision applications. The Bumblebee is pre-calibrated for lens distortions and camera misalignments. It does not require in-field calibration and the left and right images are aligned within 0.05 pixel RMS error. The calibration information is pre loaded on the camera, allowing the software to retrieve the image correction information. This feature allows seamless swapping of the cameras, or retrieving the correct information when multiple cameras are connected together.

Mini-ITX PC Board

The VIA EPIA Mini-ITX motherboard was selected as the central processing unit for the rover. It is a very compact platform with a square dimension of 170mm x 170mm. It includes an on-board x86 processor, USB ports, Firewire capability, and serial ports. It was selected due to its low power consumption and cooler operating conditions. We included a 20 GB hard drive and used the Microsoft Windows XP operating system.

FPGA:
A Xilinx Spartan-3 FPGA Starter Board was selected for sensor monitoring and propulsion control. It features a 200K gate Spartan-3, on-board I/O devices, and 1MB fast asynchronous SRAM, making it the perfect platform to experiment with our design. The board contains a Platform Flash JTAG-programmable ROM, so designs can easily be made non-volatile. The Spartan-3 Starter Board is fully compatible with all versions of the Xilinx ISE tools.

PIC servo Board:
The PIC servo board is designed as a controller for the D.C. motors, which includes incremental encoder feedback and motor drivers. It provides a serial interface connection, which communicates with our FPGA via the software designed on-board UART. The PIC Servo controller is capable of 32-bit position control, velocity control, acceleration, and Step & Direction inputs.

4. Stereo Images processing with disparity map

4.1 Disparity map of image

Composing a disparity map of the image is an efficient method for storing the depth of each pixel in the image. Depth is defined as the distance from the camera. Each pixel in the map is directly related to the corresponding pixel in the actual image. A grey scale value from 0 to 255 is used to specify the corresponding depth of that pixel rather than depicting a color.

We analyze two photos, one taken from the left lens and one taken from the right lens of the stereo camera to create a raw stereo image pair. The left photo is then superimposed onto the right photo. Corresponding points of the objects in the image are then identified. Objects closer to the camera appear to have greater separation between the corresponding points on the left and right images. Whereas, object farther away from the camera appears to have less separation between the corresponding points of the left and right images. The software correlates every pixel in the left image with every pixel in the right image using a correspondence algorithm.

In the disparity map image, objects that are lighter are identified as being closer to the rover and those that are darker are identified as being farther away from the rover. To determine the depth, two images are processed to produce a disparity map.

Figure 3 shows an image pair obtained from the horizontally displaced cameras that composed the Bumblebee camera. It is easy to identify in both images that the horizontal distance between the corresponding two points in both images are related to the actual depth information of the objects. Based on the distance it is possible to determine the disparity of the each pixel. Creating quality disparity maps at 15 frames per second (fps) is not trivial task. In Figure 4, a disparity map of 480x640 pixels and 256 levels in gray-scale depth is shown. The larger disparity values (lighter) correspond to the closer objects and smaller disparity values (darker) correspond to objects farther away. Depending on the depth resolution and object size, objects may appear at more than one image depth value.

Left Image   Right image

Figure 3: Example of a stereo image

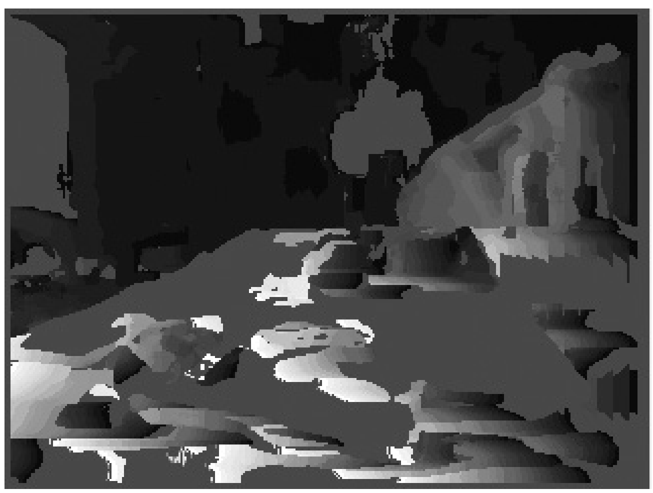

Figure 4: Related disparity image

4.2 Disparity mathematical definition:
Disparity for a feature *a* in the images is defined as

$$d(a) = x_{left}(a) - x_{right}(a)$$

where $x_{left}(a)$ is the *x* coordinate of the point a in the let image. This calculation is done for each pixel.

To establish correspondence between images, we use the sum of the absolute differences correlation method. For every pixel in the image, we select a neighborhood of a given square size (mask) from the reference image. We compare this neighborhood to a number of neighborhoods in the other image horizontally. Then we select the best match. Comparison of neighborhoods or mask is performed by calculating the expression shown below.

$$\min_{d=d_{\min}}^{d_{\max}} \sum_{i=-\frac{m}{2}}^{\frac{m}{2}} \sum_{j=-\frac{m}{2}}^{\frac{m}{2}} \left| I_{right}(x+i, y+j) - I_{right}(x+i+d, y+j) \right|$$

Where:

$d_{\min}$ is the minimum disparity

$d_{\max}$ is the maximum disparity

m is the mask size

I_{right} is the right image

I_{left} is the left image

In some cases, it is not possible to establish correspondence between images. In such cases, the obtained measurement will not be correct. To avoid this incorrect measurement, the pixel can be declared invalid if there are not sufficient correct matches.

5. Conclusions

1. Students who participated in this project learned different technical aspects of an integrated rover system and worked in a large, multi-disciplinary team. They have also improved their communication skills through team meetings and making many presentations. All students who worked in the project in the last three years have graduated and are either working engineers or have started graduate studies. The NASA PAIR project achieved the educational objectives as specified in the original proposal.

2. We have successfully interfaced with a stereo camera, obtained disparity images, and used this video information to navigate a rover to move toward a target and avoid simple obstacles. We are investigating removing shadow information from the images by collecting data from external sensors, including ultrasound and infrared sensors.

References:

[1] "Autonomous Navigation System Design for a Smart Robotic Rover ", Yi Cheng, Kathleen Hayden, Zekeriya Aliyazicioglu, Tim Lin, ASEE Conference 2005

[2] Triclops Stereo Vision System manual Version 3.1

[3] "Processing real-time stereo video for an autonomous robot using disparity maps and sensor fusion" Donald Rosselot and Ernest L. Hall