Real-Time Computer Vision Processing on a Wearable System

Kevin DeMarco
CS-6235 : Real-Time Systems
Project Proposal

September 18, 2011

Purpose

The proposed project involves the development of a wearable system that is capable of processing visual imagery from a camera in real-time and providing the user with actionable information. The system will be evaluated based on the image processing frame rate and its ability to detect obstacles.

1 The Problem and Its Importance

Ubiquitous computing has become a reality in modern society. Almost all individuals carry a cell phone for communication and many carry a smartphone capable of connecting to the internet and providing important information to the user. While smartphones are great for providing the user information on demand, wearable systems can provide the user with context-sensitive information even before the user requests the data. A wearable system, which is usually composed of sensors, cameras, or a head-mounted display, is continuously monitoring the environment and assessing the current scenario. If the wearable system determines that the user should be made aware of an event or object in the environment, the wearable system will provide the user with information in the form of video, sound, or physical feedback. There are a number of practical applications for wearable systems ranging from soldiers using wearable systems to detect threats in the environment to helping the blind navigate unknown locations. This research will focus on developing a wearable system that can be used to assist the blind in navigating foreign environments. The wearable system will consist of a small embedded computer, a camera, and a haptic feedback system.

2 Related Work

There have been individuals developing wearable systems for several decades. Researchers at MIT's Media Lab have been renowned for their development of wearable systems before wireless cell phone networks provided data plans. For example, Steve Mann placed several radio transceivers at strategic locations around Boston in order to provide himself access to the internet throughout the city [1]. He attributed miniaturization to his ability to continuously reduce the size of his head-mounted displays. Thad Starner has utilized wearable systems to localize the user within an indoor environment with the use of accelerometers and infrared detectors [2].

2.1 Non-Visual Feedback Systems

Researchers at the Georgia Institute of Technology have developed a method for assisting the blind in navigation through the use of an auditory feedback system in the System for Wearable Audio Navigation (SWAN) [3]. SWAN uses a number of sensors to accurately locate the user and then uses a pair of headphones to assist in navigation and inform the user of obstacles. While auditory feedback might make sense for the blind, the blind often rely heavily upon their hearing. Covering their ears with headphones might degrade their own awareness of the environment when their hearing is extremely acute. Thus, other researchers have implemented haptic feedback with the use of motors and other vibrating devices [4].

2.2 Wearable Simultaneous Localization and Mapping (SLAM)

The SLAM problem has been thoroughly explored in ground and aerial robotics systems over the past ten years [5]. The SLAM problem makes use of the platform's motion model, estimations of sensor noise, and feedback from the world to probabilistically build a world map and localize the robot. There has been thorough research on the real-time vision-based monocular SLAM problem utilizing indoor features [6–8]. While most of the SLAM research has focused on robotic implementations, there are some researchers that have focused on developing SLAM techniques for wearable systems [9].

3 Proposed Approach

The wearable system will be composed of three components: a robust hardware architecture, a software publish-and-subscribe communication architecture, and a user feedback system.

3.1 Hardware Architecture

The central piece to the hardware system is the Gumstix Overo Air computer-on-module (COM) [10]. The Overo Air is an ARM Cortex-A8 OMAP3503 embedded system that has the capability to communicate via Bluetooth and the 802.11g wireless networking protocols. A version of Ubuntu Linux will be burned to a micro SD card, which will be used to boot the Overo Air. The Gumstix will process video imagery from a Logitech Webcam C525 via USB 2.0 since the Overo Air has USB host capabilities [11]. The Overo Air will also communicate with the HTC EVO 4G Android smartphone via an ad hoc WiFi network [12]. If the Android phone is unable to connect to an ad hoc network, the Bluetooth communication channel will be used instead. The complete hardware architecture is shown in Figure 1. Figure 1 depicts a user with a shoulder-mounted webcam that is connected to an Overo Air COM via USB 2.0. The Overo Air COM communicates debug data to the Android smartphone for development purposes.

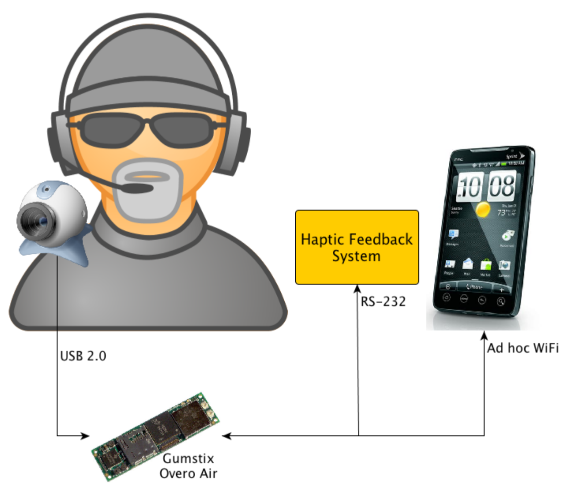

Figure 1: Wearable Hardware Architecture

3.2 Software Architecture

The software for this research will be written mostly in C++ and Python and will be targeted for an Ubuntu Linux system. The communication between software nodes will take place via the Robot Operating System

(ROS) publish-and-subscribe communication system which is targeted for TCP/IP network applications [13].

3.2.1 Publish-and-Subscribe

The ROS publish-and-subscribe system has established itself within the academic ground and aerial robotics research domain over the past two years. However, its communication architecture can enable any set of parallel tasks to communicate over networked channels. Figure 2 shows how ROS nodes communicate in order to exchange data.

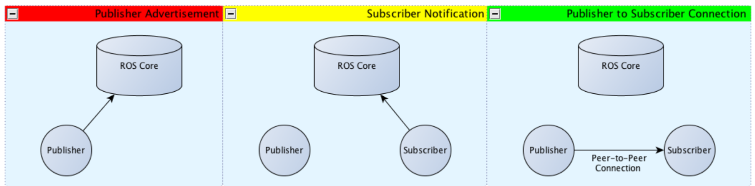

Figure 2: Publish-and-Subscribe Architecture

First, a node publishing data notifies the ROS core database of its intention to publish a certain type of data. When another node comes online and wishes to subscribe to that data, the ROS core informs the subscriber of the publishing node's location in the network. Finally, the two nodes create a peer-to-peer communication channel and begin exchanging data. This peer-to-peer topology is ideal for transferring large amounts of data without creating a bottleneck at the core database, which is often the case with other publish-and-subscribe systems. Thus, as is shown in Figure 3, a ROS node will be reading data images from the USB webcam and then publishing the imagery to the ROS node that will be executing the computer vision object detection algorithms. The objects will then be published to the decision engine that computes the navigation commands based on the importance of obstacles and the goal location. The navigation commands will be sent to the haptic feedback system and the Android smartphone.

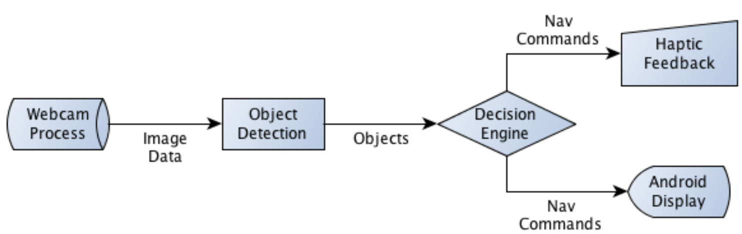

Figure 3: ROS Nodes

3.2.2 Computer Vision / Object Detection

The main purpose of the computer vision node will be to use the OpenCV computer vision package to run the Scale-invariant feature transform (SIFT) algorithm on image data [9, 14]. The SIFT algorithm will be tuned to extract features of indoor obstacles. After the features are extracted, they will be sent to a decision engine that will determine if the user should be notified of the upcoming obstacles. There will be a great deal of trial-and-error testing during this phase of the research to fine tune the obstacle detector.

3.3 User Feedback System

The third aspect of this research will be the construction of a haptic feedback system. The system will use five vibrating actuators arranged in the configuration shown in Figure 4. The vibrators will be controlled with a small 8-bit microcontroller that will communicate with the Gumstix via RS-232. The haptic feedback system will be placed on the back of the user and will indicate the direction of movement. For example, when the wearable system wants the user to move straight and then turn to the right, the top vibrator will actuate and then the right vibrator will actuate. The "Stop" command will be indicated by the center vibrator. While this is a fairly crude navigation system, it is not the purpose of this research to determine the psychological effects that the haptic feedback system will have on the user. In the future, this haptic feedback system can easily be replaced by another feedback system and still interface with the rest of the wearable system without any problems. During development, the Android smartphone will also serve as a feedback system for the developer.

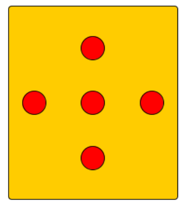

Figure 4: Vibrating Actuator Configuration

3.4 Hardware Resources

- Gumstix Overo Air COM with Tobi Expansion Board
- Logitech Webcam C525
- Android HTC EVO 4G Smartphone
- Vibrating Actuators
- Microchip 8-bit PIC Microcontroller

3.5 Software Resources

- Robot Operating System (ROS)
- OpenCV
- Android API

4 Evaluation

The wearable system will be judged based on its ability to detect and notify the user of upcoming obstacles. The system will also be evaluated based on the frame rate at which the system can process images and extract features. The final demonstration of the system will involve guiding an artificially visually-impaired user through a room or hallway filled with obstacles.

5 Deliverables

- Final Report
- Project summary on website
- Image processing source code
- Wearable System User Manual
- Project Presentation

6 Schedule

The project schedule is defined by the milestones in Table 1.

Date	Milestone
9/24	Complete study of Monocular SLAM techniques
10/7	Complete bootable Gumstix SD card image
10/21	Perform basic SIFT feature extraction with webcam attached to Gumstix
10/28	Interface Gumstix with Android smartphone using ROS
11/4	Complete construction of wearable webcam harness
11/11	Complete construction of wearable haptic navigation vest
11/18	Finish interfacing Gumstix to haptic vest
11/19	Begin final testing of entire system
11/28	Begin final report and presentation
12/5	Final report complete and project demonstration

Table 1: Project Milestones

7 Possible Work

If time permits, the following research topics will be explored.

7.1 Simultaneous Localization and Mapping (SLAM)

SLAM has been researched heavily in the robotics community. Many of the SLAM techniques could be utilized to develop a wearable SLAM system. One of the biggest challenges will be establishing the camera's pose since it will be loosely attached to the user where in robotics platforms, the camera is rigidly attached to the platform.

7.2 Facial Recognition

The OpenCV computer vision package has tools within it to perform facial recognition. This could be a great way of helping the user know with who he is talking and for helping the user to remember people he meets in public.

7.3 Computer Simulation

Finally, a computer simulation of a visually-impaired individual navigating the environment could be developed. There are a number of simulators that exist for testing robotic SLAM algorithms. The Blender / Modular OpenRobots Simulation Engine (MORSE) simulator creates a 3D game environment that accurately models sensor noise and motion for the purpose of robotics simulation [15]. The developers of MORSE also created a model of a human being which could be used to test some of the algorithms discussed in this proposal.

References

[1] S. Mann, "Wearable computing: A first step toward personal imaging," *Computer*, vol. 30, no. 2, p. 2532, 1997.

[2] G. Schindler, C. Metzger, and T. Starner, "A wearable interface for topological mapping and localization in indoor environments," *Location-and Context-Awareness*, pp. 64–73, 2006.

[3] B. PhD and J. Lindsay, "Using virtual environments to prototype auditory navigation displays," *Assistive Technology*, vol. 17, no. 1, pp. 72–81, 2005.

[4] S. Ertan, C. Lee, A. Willets, H. Tan, and A. Pentland, "A wearable haptic navigation guidance system," in *Wearable Computers, 1998. Digest of Papers. Second International Symposium on*, 1998, pp. 164–165.

[5] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part i," *Robotics & Automation Magazine, IEEE*, vol. 13, no. 2, pp. 99–110, 2006.

[6] J. Mir, G. Dissanayake, and W. Zhou, "Vision-based SLAM using natural features in indoor environments," in *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2005. Proceedings of the 2005 International Conference on*, 2005, pp. 151–156.

[7] A. Davison, "Real-time simultaneous localisation and mapping with a single camera," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, 2003, pp. 1403–1410.

[8] A. Davison, I. Reid, N. Molton, and O. Stasse, "MonoSLAM: real-time single camera SLAM," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1052–1067, 2007.

[9] R. Castle, D. Gawley, G. Klein, and D. Murray, "Towards simultaneous recognition, localization and mapping for hand-held and wearable cameras," in *Robotics and Automation, 2007 IEEE International Conference on*, 2007, pp. 4102–4107.

[10] "Gumstix small open source hardware," http://www.gumstix.com/. [Online]. Available: http://www.gumstix.com/

[11] "HD webcam c525," http://www.logitech.com/en-us/webcam-communications/webcams/devices/7794. [Online]. Available: http://www.logitech.com/en-us/webcam-communications/webcams/devices/7794

[12] "HTC mobile phones - EVO sprint - overview," http://www.htc.com/us/products/evo-sprint. [Online]. Available: http://www.htc.com/us/products/evo-sprint

[13] "Documentation - ROS wiki," http://www.ros.org/wiki/. [Online]. Available: http://www.ros.org/wiki/

[14] "Welcome - OpenCV wiki," http://opencv.willowgarage.com/wiki/. [Online]. Available: http://opencv.willowgarage.com/wiki/

[15] "morse - openrobots," http://www.openrobots.org/wiki/morse/. [Online]. Available: http://www.openrobots.org/wiki/morse/