Shoes as a Platform for Vision

Charles C. Kemp and Paul Fitzpatrick

What: We are exploring the use of a shoe-mounted camera as a sensory system for wearable computing. We are developing a system called `Platform Shoe` which is useful for gait analysis, obstacle detection, and context recognition [2]. `Platform Shoe` uses visual information alone to detect periods of stability and motion during walking – no inertial sensor is required. In the stable phase of walking, the foot can be assumed to be parallel to the ground plane. In this condition, the floor dominates the lower part of the camera's view, and `Platform Shoe` can segment the floor out from the remainder of the scene, leaving walls and obstacles. This in turn facilitates floor surface recognition for context awareness.

Why: Costs for digital cameras and computation continue to be driven lower by technological advances and strong demand. Future wearable computing system can benefit from these trends by applying cameras to new, more specialized, and less traditional sensing tasks. We are interested in the use of a shoe-mounted camera for gait analysis, obstacle detection, and context recognition.

Wearable computing on shoes has been used for a variety of purposes, including user interfaces, power production, and gambling. We show that visual processing can also benefit from this prime location. The planted foot is the only part of the body that is reliably stationary with respect to the world during walking and standing. When we walk, our feet come into contact with the ground in an alternating pattern. Each foot swings swiftly through the air, then is pressed against the ground as the weight of the body is transferred onto it [5]. During these key moments within a person's stride, the planted foot tends to be in a canonical orientation with respect to the floor and relatively motionless, which leads to simplified vision processing.

How: Typically, during the swing phase of walking, the temporal derivative of images from the camera increases (due to motion), the spatial derivative decreases (due to blur), and the mean luminance tends to fall (due to the camera looking towards the floor). These measures allow us to determine when the foot is pressed against the ground, with the camera in a fixed orientation with respect to the ground plane. This is an ideal opportunity for visual processing.

Figure 1: The system. A camera and inertial sensor are mounted on a sandal.

Due to the canonical orientation of the stable images selected by our gait analysis, we know with high probability that the bottom of the image is floor and that the top of the image is not floor. The bottom quarter of these special images corresponds with the area from the toe out to 3 inches on the floor. The top quarter of these images is very far above the horizon line. We exploit this property to perform segmentation and recognition. With segmentation, this observation allows us to collect a significant sample of the appearance of the floor and non-floor parts of the image. With recognition, we are able to reliably sample from the floor without the need for any floor detection algorithms. These images are well suited to wearable computing applications that benefit from detailed sensing of the wearer's nearby environment (for example, detection of walking hazards [4, 6]). If we assume flat floors, we can construct a function that for each pixel gives the distance from the toe to the corresponding point on the floor [3], which could be useful for object avoidance.

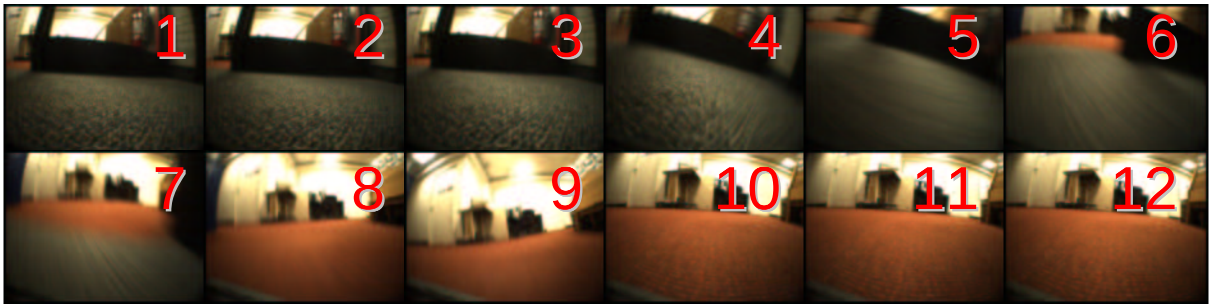

Figure 2: This figure shows a sequence of frames taken during a single step. In that step, the wearer moves from a lobby area into a corridor, and from a blue to red carpet. The main swing phase of the step occurs in frames five to nine. During this phase, there is considerable motion blur, the camera points away from the floor towards the lights, and the view is changing rapidly from frame to frame.

Progress: We have used camera placement to essentially solve the problem of floor detection, and to begin work on floor segmentation and recognition. We have had some initial success with a very simple color model for distinguishing different floors, with an 86.3% success rate (compared with 44% for a classifier that always guessed one particular carpet that dominated the environment). Areas with different functions often have distinct floor surfaces. For example, a wash room floor is unlikely to be carpeted so that it can be easily mopped. Floor recognition is therefore a valuable cue for localization and context awareness [1].

Future: In general, as cameras and computation become less costly we expect for more specialized camera sensing, such as this, to become practical for wearable computing. Issues of privacy and misuse could be mitigated by making a closed sensory system. A camera on each foot would make several applications easier by allowing for nearly uninterrupted acquisition of closely-spaced stable images. Several interesting future applications might be built on top of the results we have presented including automated cartography, localization, detection of nearby people by their feet and legs, more powerful floor recognition systems, outdoor operation, and recognition of common nearby objects such as chairs, tables, walls and trash cans.

Research Support: Funds for this project were provided by DARPA as part of the "Natural Tasking of Robots Based on Human Interaction Cues" project under contract number DABT 63-00-C-10102, and by the Nippon Telegraph and Telephone Corporation as part of the NTT/MIT Collaboration Agreement.

References:

[1] Brian Clarkson, Alex Pentland, and Kenji Mase. Recognizing user context via wearable sensors. In *Proc. of the Fourth Intl. Symposium on Wearable Computers*, pages 69–76, 2002.

[2] Paul Fitzpatrick and Charles C. Kemp. Shoes as a platform for vision. In *Proceedings of the 7th IEEE International Symposium on Wearable Computers*, pages 231–234, White Plains, NY, USA, October 2003.

[3] Ian Horswill. Polly: A vision-based artificial agent. In *Proc. of the 11th National Conference on Artificial Intelligence*, pages 824–829, Menlo Park, CA, 1993.

[4] C. M. Lee, K. E. Schroder, and E. J. Seibel. Efficient image segmentation of walking hazards using IR illumination in wearable low vision aids. In *Proc. of the Sixth Intl. Symposium on Wearable Computers*, pages 127–128, 2002.

[5] Jerry Pratt. *Exploiting Inherent Robustness and Natural Dynamics in the Control of Bipedal Walking Robots*. PhD thesis, MIT, Cambridge, Massachusetts, 2000.

[6] Shraga Shoval, Iwan Ulrich, and Johann Borenstein. NavBelt and the GuideCane. *IEEE Robotics and Automation*, 10(1):9–20, March 2003.