AR Paper Review

by Abigail McKinstry

September 2014

1 Abstract

Two papers on augmented reality (AR) user-interfaces and gesture recognition were analyzed. The papers in review were chosen for their relevance to an AR gesture interaction project at the University of Canterbury. The aim of this project is to use gesture recognition to control a three-dimensional painting application, where the user can perform extra actions such as erasing and scaling after "painting" their design with gestures.

2 Table-top virtual objects

2.1 Description

The paper *Virtual object manipulation on a table-top AR environment*[1] addresses issues faced in user-tracking and virtual object interactions for AR surfaces. The paper describes in detail how these issues affect user-collaboration and the intuitiveness of the interface. An accurate vision-based tracking method with a *tangible user-interface*, or *TUI* was used for the application.

A TUI turns real objects, which conveniently have constraints on their manipulation, into input and output devices that are easy and intuitive to use. Their physical constraints—such as their lack of visual and verbal cues in interaction—can be overcome using AR techniques. A *tangible augmented reality*, or *TAR* user-interface should have concurrency in the manipulation of multiple interface elements.

The *Shared Space Siggraph 99*[1] TAR interface required users to view the space through a head-mounted camera/display. The position of the camera was calculated with respect to square tracking markers, which was problematic when users obstructed the markers.

An improved method involved the use of multiple, circular fiducial markers ("blobs") for *global* coordinate tracking. However, three "blobs" were required for pose/position estimation, and "blobs" were less distinct from other visible features. So a combination of large square markers and blobs was used, and the tracking of pose/position was estimated from *all* available features. If square markers were available, the estimations were combined with those calculated using the blobs, which resulted in quantifiable errors. If only blobs were available, previous frames were compared to the current frames, and the blobs' positions relative to previously-visible square markers were used. This method was effective only for slow head motions.

2.2 Relevance

For the gesture-interaction project, a TAR interface could include a physical paintbrush that has AR "paint" overlaid onto it. Other tools such as erasers could be used concurrently with the paintbrush.

To fix the occlusion problem, fiducially marked "tools" could be held by users to manipulate objects. This would enable software to correctly display occlusion that would be caused by the tool. However, as the presence of hands would still not result in AR content occlusion, all AR objects could be made transparent to reduce the unnaturalness of the interface. Another method not discussed in this paper could be to track hand position and compute the occlusion this way; or even completely remove the user from the AR content—with their gestures not "touching" any of the objects (perhaps the objects could move with respect to the user).

2.3 Critique

The paper described, in great detail, the limitations of tangible user interfaces as a reason for employing augmented reality. There is then a verbose section on the short-comings of another table-top AR user-interfaces, and people's "rating" of the user interface, which seems largely unnecessary. It isn't until the fourth page/-section that the implemented "accurate vision-based tracking method" is discussed—again without much concision. If the paper had been written more concisely, perhaps more of the mathematical equations used could have been described. The pictures in the last sections, however, were very enlightening, and the final implementation impressive.

3 Real-time finger/hand tracking

3.1 Description

The paper *Lightweight Palm and Finger Tracking for Real-Time 3D Gesture Control* [2] describes the development of real-time finger and palm recognition software for use with a time-of-flight camera (i.e. a camera that gathers depth information). The technique works regardless of lighting conditions and background content. The software is able to detect hand position without prior knowledge of its existence, determine poses from image cues, and recognise gestures from the poses over time. The paper describes common techniques and their respective short-comings, which the new system has remedied.

A basic technique for pose recognition is to estimate the palm as the centroid of the hand "blob", and use specialised mappings of gestures. However, noise causes problems with the blob approach, and an extensive database of "training samples" is required for the gesture mapping. Thus, the authors recommend an approach using circular mask convolution, curvature analysis, or palm-edge distance transform analysis. Fingers are detected by classifying pixels by their "pipe-likeness" and connection to a large body (palms).

3.2 Relevance

The methods used for detailed hand motion detection could be used for the gesture-interactive AR project, and would enable a greater variety of user-interaction features to be explored. Without finger pose recognition, it will be difficult, for example, for the user to "turn off" the paintbrush whilst moving to a new position. At this stage, due to time-constraints, it is unlikely that the project team will implement a finger-recognition strategy of this level of complexity; however if no open-source alternatives are available, a few of the techniques described in this paper could be explored.

3.3 Critique

The paper was well-written and the procedure was relatively easy to follow. However, with no experience with the image-processing techniques used, many of the algorithms described were too complex to understand fully. This level of depth in the mathematical explanations would, however, enable experienced readers to replicate some of the processes.

References

[1] H. Kato, M. Billinghurst, I. Poupyrev, and K. T. K. Imamoto, "Virtual object manipulation on a table-top ar environment," [Online]. Available: http://codesyntizer.tk/wp-content/uploads/2013/09/isar2000.pdf.

[2] G. Hackenberg, R. McCall, and W. Broll, "Lightweight palm and finger tracking for real-time 3d gesture control," [Online]. Available: http://rodmc.com/wp-content/uploads/2012/04/vr2011paper.pdf.