ON POSITIONING FOR AUGMENTED REALITY SYSTEMS

Stelian Persa
Pattern Recognition Group, Delft University of Technology
Lorentzweg 1, 2628 CJ Delft, The Netherlands
stelian@ph.tn.tudelft.nl

Pieter Jonker
Pattern Recognition Group, Delft University of Technology
Lorentzweg 1, 2628 CJ Delft, The Netherlands
pieter@ph.tn.tudelft.nl

ABSTRACT

In Augmented Reality (AR), see-through Head Mounted Displays (HMDs) superimpose virtual 3D objects over the real world. They have the potential to enhance a user's perception and his interaction with the real world. However, many AR applications will not be accepted until we can accurately align virtual objects in the real world. One step to achieve better registration is to improve the tracking system. This paper surveys the requirements for and feasibility of a combination of an inertial tracking system and a vision based positioning system implemented on a parallel SIMD linear processor array.

1. INTRODUCTION

Position and Orientation Tracking is used in Virtual and Augmented Reality (VR, AR). With a plethora of different graphics applications that depend on motion-tracking technology, a wide range of interesting motion-tracking solutions have been invented [1],[2]. Many applications only require motion over a small region, and these tracking approaches are usable, although there are still difficulties with interference, line-of-sight, jitter and latency.

Inertial sensors measure acceleration or motion rate, so their signals must be integrated to produce position or orientation. Noise, calibration error, and the gravity field impart errors on these signals, producing accumulated position and orientation drift. Hybrid systems attempt to compensate for shortcomings of each technology by using multiple measurements to produce robust results. Inertial gyro data can increase the robustness of a vision system by providing a frame by frame prediction of camera orientation and a vision system can correct for accumulated drift of an inertial system.

The paper is organized as follows. Section 2 describes an inertial tracking system for augmented reality. Section 3 presents the way in which a real-time vision system can enhance the inertial system. Execution times and implementations are also presented. Section 4 concludes the paper.

2. INERTIAL TRACKING FOR AUGMENTED REALITY

A modern six-degree of freedom sensor is the IS-600 [3]. Its inertial sensor simultaneously measures 9 physical properties: angular rates, linear accelerations, and magnetic field components along all 3 axes. Micro vibrating elements are used to measure angular rates and linear accelerations. The angular rate signals are integrated in a direct manner to obtain orientation. The linear acceleration signals are double integrated to track changes in the position. However, this leads to an unacceptable rate of positional drift and must be corrected frequently by some external source. Hence the system uses an acoustic time-of-flight ranging system to prevent position and orientation drift. The accelerations are also fed into an error estimator to help cancel pitch and roll drift. The yaw drift is corrected by the acoustic range measurements simultaneously with the position drift. This leads to a:

- Max. Angular Rate: 1200°/sec [3]
- Angular Accuracy: 1.3° RMS [4]
- Translation Accuracy: 3.6 RMS(mm)[4]
- Prediction: 0-50ms [3]
- Orientation Update Rate: up to 500Hz[3]
- Position Update Rate: up to 150Hz [3]

The question is whether such a tracker is feasible for our Retinal Scanning Display for Augmented Reality [8]. The accuracy needed for registration of the real and virtual images on such a see-through headset involves two aspects, spatial accuracy and latency. The user of a Virtual Reality system will tolerate, and possibly adapt to, errors between his perceived motion and the visual result. However, with an Augmented Reality system the registration is within the visual field of the user and hence the visual effect may be more dramatic.
For spatial accuracy we use the fact that the central fovea of a human eye has a resolution of about 0.5 minute of arc [7] and that it is capable of differentiating alternating brightness bands of one minute of arc. For latency we can reason that at a moderate head rotation rate of 50 degrees per second the system lag must be ≤ 10 ms to keep angular errors below 0.5 degrees.

The IS-600 system is measured as having about 1° RMS static orientation accuracy and a 3° RMS dynamic accuracy. We must conclude that although suitable for applications in virtual reality, this accuracy is probably inadequate for AR tracking.

To illustrate this we map this error onto the 2D-image domain of our see-through headset. Considering the focal length of a video camera (in pixels), $L_x$ be the horizontal image resolution and $\theta_x$ the Field-Of-View of the camera, then the ratio of image pixel motion to the rotation angels (in pixels/degree) is:

$$\frac{L_x}{\theta_x} = \frac{L_x}{2 \arctan \left( \frac{L_x}{2 f_x} \right)}$$

E.g. with an image resolution of 640x480 and a focal length $f_x = 600$ pixels, the 1° RMS static accuracy of the tracker leads to 11 pixels misalignment between real and virtual objects on our see through headset.

### 3. REAL-TIME VISION FOR THE CORRECTION OF POSITION DRIFT

To correct positional drift in a 6-DOF inertial tracking system, some type of range measurements to known points in the environment is required. For outdoor operation, only a vision-based system is suitable, because for all the other systems the resolution is inadequate (GPS, the range is too small and/or the sensors are not suitable to operate outside (Acoustic). Hence we investigated image processing algorithms and architectures for fast feature and fiducial tracking. We foresee a vision system that searches and tracks known features in outdoor scenes, such as lines, corners and man made fiducials to calculate a pose to be used to correct the drift of the inertial tracker.

In order to obtain fast pose estimations we implemented the algorithms on a parallel architecture for real-time image processing, the IMAP-Vision board, a frame grabber and Linear Processor Array (LPA) of 256 8 bit processors in SIMD mode on a single PCI board. The board is programmed in a data parallel version of C [5].

Our line detection is a parallel Sobel edge detector followed by a parallel Hough Transform. Sobel operator and Hough transform over 1300 points takes 49 ms on the IMAP-Vision system. The error in estimation of the line segment is about one pixel for the radius and 0.7 degree for the angle. These errors can be reduced by increasing the parameter space, however, this will increase the execution time. We implemented two corners detection algorithms in a data parallel way. Our parallel implementation on the IMAP-Vision system of the Kitchen and Rosenfeld corner detector takes 7ms and about 3ms for a gradient direction corner detector. The last one needs a false corner suppression postprocessing. To regularly calibrate the AR headset on the Campus we can use man-made fiducials onto known positions. The vision system can recognise these and automatically reset position and orientation of the tracker, when sufficient correspondences are found. As fiducials we used small square red blocks (forming a dot-code) on a larger black background. Our fiducial recognition algorithm takes 25msec, and has an error of about 9mm in real world if the camera is 3m apart and the FOV of the camera is 45 degree, due to the resolution of the grabber (256²).

Subsequent pose estimation takes from 15msec. to 2sec. on a Pentium depending on the initial guess of the pose, and with a Kalman filter we can estimate recursively the position and uncertainty of a moving feature point in each next frame. The Kalman filter also fuses the information from inertial measurement unit with the data from vision system. We estimate that on a Pentium this takes about 10msec. With a total pose update rate of 60 msec., and an accuracy of 9mm we conclude that it is feasible to make vision trackers based on parallel Image Processing Architectures, for the correction of position drift in inertial trackers for Augmented Reality applications.

### 4. RESULTS AND CONCLUSIONS

In this paper we surveyed the requirements for and feasibility of a combination of an inertial tracking system and a vision based positioning system for augmented reality systems. At a moderate head rotation rate of 50 degrees per second, the lag of such a system must be ≤ 10 ms to keep angular errors below 0.5 degrees. The dynamic accuracy for rotations of 1°, yields a registration error of 11 pixels, which forms a problem for objects far away from the camera.

We developed a visual system based on a parallel image processing array. A fiducial tracker takes 25 msec and the error is about 9mm. We can conclude from this results that the IMAP-VISION board is suitable as a vision system for outdoor augmented reality applications.

This work has been conducted in the framework of the Ubiquitous Communication Project (UbiCom) from the TU Delft (http://www.ubicom.tudelft.nl/).
5. REFERENCES


