IPT-EGVE Symposium (2007)
B. Fröhlich, R. Blach, and R. van Liere (Editors)

Short Papers

Advanced Inertial-Optical Tracking System for Wide Area Mixed and Augmented Reality Systems

D. Wormell (deanw@intersense.com), E. Foxlin (ericf@intersense.com), and P. Katzman (pkatzman@intersense.com)

InterSense, Inc., 36 Crosby Drive, Suite 150, Bedford, MA, USA

Abstract

InterSense is introducing a vastly improved version of its wide-area tracking inertial-optical tracking system (IS-1200 VisTracker). The new VisTracker technology has several enhancements including device miniaturization, an embedded image processor, improved optics, a higher resolution imaging CMOS monochrome sensor, power reduction, and ruggedization. A review of the IS-1200 architecture and new components is discussed. Examples of the IS-1200 initial deployment are given for AR applications.

Keywords:

Wide Area Motion Tracking, Mixed Reality; Augmented Reality; Ubiquitous Tracking, Inertial Tracking

Categories and Subject Descriptors:

1. Introduction

InterSense pioneered 6-DOF inertial based motion tracking systems in 1999 with its IS-900 wide area inertial-acoustic tracking system. Although the inside-out acoustic configuration was originally conceived and prototyped using wireless ultrasonic transponder beacons to allow very wide-area tracking of mobile AR users [1], the IS-900 product that was brought to market was designed for room-scale immersive display applications or simulators, and thus uses a hard-wired array of beacons. To address the original requirement for building-wide tracking in mobile AR applications, we embarked on development of a vision-aided inertial self-tracker that combines an inside-out smart camera with an inertial sensing core [2]. After several iterations of prototypes and field trials, the IS-1200 VisTracker was completely redesigned. The paper describes the requirements that were established as a result of the feedback from the field trials, and the enhancements that were made to the product compared to the prototypes.

2. IS-1200 Enhancements

2.1. First Generation IS-1200 VisTracker

The first generation of the IS-1200 VisTracker was introduced in 2003 (Figure 1). The Vision-Inertial Self-Tracker (VisTracker) is a hybrid system that fuses data from inertial and vision sensors. This provides a tracker solution for extremely wide areas without the prohibitive cost of an extensive network of sensors or emitters installed in the tracking environment. It was small enough to wear on a belt, and accurate enough for good AR registration.

Figure 1: First Generation VisTracker, comprising InertiaCam sensor head, InertiaCam electronics unit, and wearable Sensor Fusion Processor

The VisTracker was designed to meet the performance requirements of AR, wearable VR systems, and robot navigation. This hybrid vision/inertial self-tracking system distinguished itself from vision-only trackers such as ARToolKit [3] because:

1. It tracks its pose relative to a unified world coordinate system, fusing together measurements from all the fiducials that are visible.
2. It provides 180 pose updates per second, GENLOCKED to the video rendering rate,

with guaranteed constant latency and inertial-quality prediction.

3. It uses a novel passive fiducial design and image processing strategy which allows it to quickly decode over 30,000 different markers under more varied lighting conditions and with higher centroid accuracy.
4. It relies primarily on inertial tracking, so it can handle rapid camera motions and occlusions much more robustly.

2.2. New Generation IS-1200 VisTracker

Based on customer feedback from the first generation IS-1200 VisTracker, the goal of the new VisTracker is to bring users greater mobility, performance, simplicity, industrial ruggedness, and reliability.

Figure 2: New Generation VisTracker

2.2.1. Reductions in size & power

The new IS-1200 InertiaCam (Figure 2) consolidates the inertial/optical sensors and the image processing DSP into a single unit about the size of the original sensor head. A 40% power reduction (from 2.5 W to 1.5W) leads to extended mobility and battery life. Also, the acceptable input voltage range for the new InertiaCam is 4-9V, which is significantly more flexible than the 8-12V of the prior version, and allows the product to run directly off 5V USB bus power, eliminating the need for an extra power supply and power cord. Another major improvement is that the sensor fusion software has been reduced to a background process that can run in the host PC and consume only a few percent of its resources. This enables the elimination of the bulky and power-hungry wearable Sensor Fusion Processor shown in Figure 1, and reduces the VisTracker system from 3 boxes to one small sensor unit.

2.2.2. Imaging and Processor Performance Improvements

A Blackfin DSP [4] is embedded in the InertiaCam sensor, providing both miniaturization and 5-10 times increase in image processing computational power. This also provides a platform for future expansion of image processing capabilities, such as faster fiducial recognition or eventual incorporation of natural feature tracking. The resolution of the CMOS monochrome image sensor was doubled to 1280 x 1024, allowing applications to use smaller fiducials at the same distances or same fiducial sizes at longer distances (Figure 3). The Blackfin is responsible for all the image acquisition and processing, as well as interleaving messages coming from the inertial sensor at 180 Hz between the messages describing target centroid data from the camera. It uses an interrupt driven scheme to assure all inertial data is passed on with no latency, since the tracking updates depend directly on the inertial data, with optical data used on an as-available basis through Kalman filter updates to correct the gradual drift of the inertial sensors. Although the increased sensor resolution results in the need to acquire and process four times more pixels per target, the combination of the faster processor and a CMOS imager which allows selection and download of small regions-of-interest around the desired targets results in an average rate of optical measurements of about 50 Hz, compared to 30 Hz in the previous version.

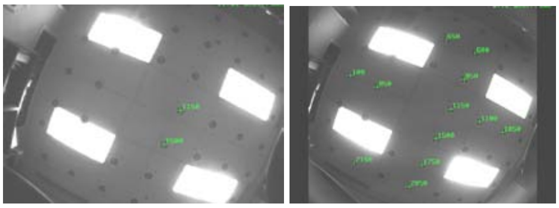

Figure 3: Ability of VGA (left) and XGA (right) InertiaCams to decode 12.7 cm fiducials from 2.7 m distance

2.2.3. Packaging and Factory Calibration

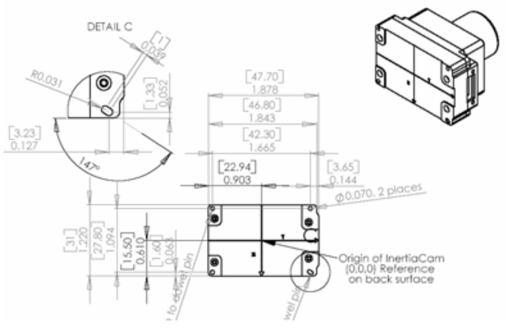

Figure 4: Mechanical Package of VisTracker

Each system comes factory calibrated allowing a simple installation procedure with a one time initial registration. Factory calibration involves three separate procedures: calibration of inertial sensors relative to the package, intrinsic calibration of lens parameters, and extrinsic calibration of optics relative to package. Furthermore, the system will operate over a wide range of natural lighting conditions and volumes by utilizing auto exposure algorithms that can ignore bright ceiling lights. In addition, the new alignment foot makes it possible for users to replace one InertiaCam with another and not have to perform any re-calibration because each unit can self-align to the mounting pins with sub-milliradian accuracy (Figure 4).

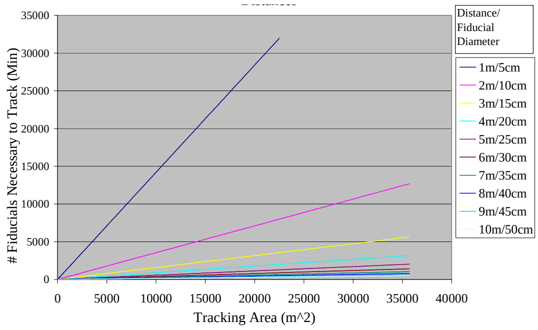

Figure 5: Minimum # Fiducials at Large Tracking Areas and at Various InertiaCam Distances

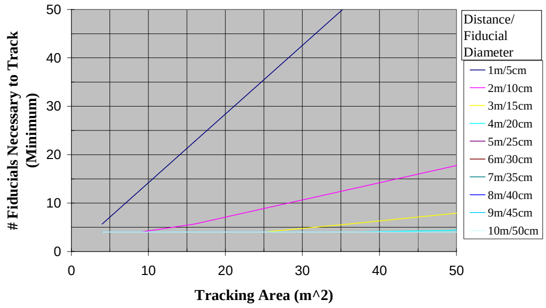

Figure 6: Minimum # Fiducials at Tracking Areas Under $50m^2$ and at Various InertiaCam Distances

2.2.4. Ruggedness and Reliability

The IS-1200 VisTracker can withstand the rigors of a wide range of challenging mobile environments. Designed for industrial conditions, the precision machined aluminum packaging seals out surrounding contaminants and protects the precision optics and electronics. Drop tests of VisTracker (Video1) at 3 feet above a linoleum floor were performed and demonstrated to show continued performance. The cable connects to the InertiaCam through a secure 15-pin connector which is both exceptionally small and rugged to military standard MIL-DTL-32139.

3. VisTracker use in AR Applications

3.1. Virtual Production

Arguably, the first commercial deployment of augmented reality was with film and video producers. By tracking the production or film camera, virtual scenes and elements can be digitally mixed into the recorded scene. While the most accurate and consistent way of tracking camera motion has been with the use of mechanically encoded systems, free roaming camera tracking has been deploy with varied success using optical outside in tracking, through the lens tracking and inertial acoustic tracking [5]. The VisTracker is currently being integrated with Cinital's Previzion [6] virtual production system. To provide rapid deployment, Cinital has built a "fiducial mat" that is deployed quickly on the studio floor. The VisTracker tracks the studio camera and its data is combined with lens encoders to an image processing engine which renders virtual and real camera data at high definition resolutions. Figure 7 shows a mixed reality video frame from Cinital's Previzion System.

Figure 7: Cinital use of VisTracker with their Virtual Production System (background is virtual)

3.2. Advanced Interactive Role Playing Games

Researchers in the College of Computing at the Georgia Institute of Technology are using VisTracker in a number of projects including the ARFacade [7], an interactive augmented reality drama. Players move through a physical apartment and use gestures and speech to interact with two autonomous characters. Few other entertainment experiences can combine interactive virtual characters, non-linear narrative, and unconstrained embodied interaction as well as ARFacade. The benefit of VisTracker is its ability to track multiple wide areas, including obstructed views, without losing performance, and within the constraints of research budgets. Figure 8 and Video2 shows ARFacade.

Figure 8: VisTracker used for AR Role Playing

4. Summary

4.1. IS-1200 VisTracker for mobile autonomous computing applications

The VisTracker is ideal in mobile autonomous applications. When combined with a mobile PC, any number of users or objects can be tracked in the same environment. Figure 9 shows the InertiaCam mounted on a small handheld tablet PC (Sony UX-180) to produce a highly portable AR "magic lens" that can be used in maintenance and inspection applications.

Figure 9: USB tracker on mobile PC

4.2. Advantages over pure optical outside-in systems

Pure optical outside-in systems require two or more cameras looking at a target with a minimum of three non-collinear targets on the tracked object to determine both location and orientation. The targets need to be separated by a distance great enough for the camera to resolve each target in order to calculate orientation. Greater target separation is needed to improve angular accuracy. Optical occlusions of targets will cause tracking failure, even for brief periods of time. Larger tracking areas are achieved by placing more cameras in the environment, resulting in a very expensive solution.

4.3. Range of hybrid inside-out systems

Inside-out systems require only one small camera on the tracked object, and a constellation of fiducials in the environment. Deployment of fiducials over building-size tracking areas is low cost and can be done very quickly, especially in buildings with a drop-ceiling grid. For example, by creating a set of fiducials mounted onto CDs with magnetic discs glued to the back, they can be stuck at the intersections of the metal ceiling grid, and by using a magnet whose diameter just fits inside the corner, they will self-align to the grid with sufficient accuracy that in most cases no further mapping is required (Figure 10). Another improvement in the new VisTracker is the incorporation of new algorithms that allow it to acquire and track with just two fiducials in the field-of-view, which halves the required fiducial density. Figures 6 and 7 give guidelines for the recommended number of fiducials to cover a certain amount of area at various ceiling heights. Figure 10, left, shows a large conference room being tracked by just 12 magnetically-mounted fiducials that were put up in a couple minutes. The high ceiling allowed good results with low density (4-foot spacing = 0.67 fids/m^2). Figure 10, right, shows a good density for use in a typical office, in which fiducials are staggered to produce 2.8-foot spacing (1.37 fids/m^2). The combination of a wide 90° lens FOV, increased camera resolution that can decode 2.5" fiducials printed on 4.7" compact disks, and new algorithms that can acquire and track with just two fiducials in view, has resulted in a system which is quite practical to scale over very large office building areas using reasonable numbers of fiducials that are quick and easy to install. It is hoped that the ready availability, low cost and convenience of this new type of wide-area tracking system will lead users to develop many new mobile AR and robotics applications ranging from building maintenance to entertainment.

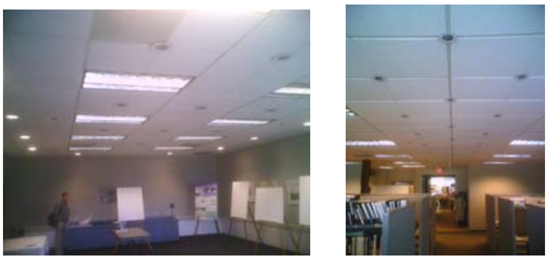

Figure 10: Magnetic-mount fiducials made from CDs allow very rapid deployment

References

1. E. Foxlin, M. Harrington, and G. Pfeiffer. (1998). "Constellation™: A Wide-Range Wireless Motion Tracking System for Augmented Reality and Virtual Set Applications", SIGGRAPH 98 Conference Proceedings, ACM Annual Conference Series, Orlando
2. E. Foxlin and L. Naimark. (2003). "VIS-Tracker: A Wearable Vision-Inertial Self-Tracker" IEEE Virtual Reality 2003
3. ARToolKit. http://www.hitl.washington.edu/research/shared_space/.
4. Analog Devices Blackfin web site. http://www.analog.com/processors/blackfin/
5. D. Wormell and M. Read (2000) "Unified Camera, Content and Talent Tracking in Digital Television and Movie Production", NAB 2000, Las Vegas, NV.
6. Cinital Previzion Virtual Productions System. http://www.cinital.com.
7. Georgia Institute of Technology, GVU Center http://www.gvu.gatech.edu/arfacade/index.html.
8. E. Foxlin and L. Naimark. (2003). "Miniaturization, Calibration & Accuracy Evaluation of a Hybrid Self-Tracker", IEEE/ACM International Symposium on Mixed and Augmented Reality, Tokyo.