

# Tracking Challenges in the Development of a Mobile Outdoor Augmented Reality Application

Fei Xue<sup>1</sup>, Shobhit Chandra<sup>2</sup> and Grace Rumantir<sup>1</sup>

<sup>1</sup>Berwick School of Information Technology, Monash University

<sup>2</sup>Centre for GIS, School of Geography & Environmental Science, Monash University

**Keywords** Mixed Reality, Computer Vision, Hybrid Tracking, GPS, Orientation Sensor

**Abstract** Augmented Reality (AR) technology can be used to augment one's view with information handling capabilities by constructing a mixed-reality environment to acquire intuitive and interactive environmental representation. This paper stems from a work-in-progress development of a mobile outdoor AR application which aims to investigate the potentials and technical challenges of outdoor mixed-reality environment and to optimize the use of on-the-spot urban and architecture planning for planners and local communities. The potentials and challenges are exemplified by blending a live video stream of a suburban streetscape environment and relevant artificial virtual information together to construct a mixed reality that provides informative and intuitive environmental design presentation. This paper focuses only on one technical challenge related to registration and tracking encountered in the development of this type of AR application. The challenges emerge from two aspects. First, accuracy, latency and stability in registering and tracking a view point within an outdoor environment are far from being satisfactorily provided. Second, recent outdoor sensor technologies capturing spatial location and orientation, such as Differential GPS (DPGS) sensing and Inertial Measurement Unit (IMU) sensing, are only used as a compensatory method to complement the results of Computer Vision (CV) based tracking, but not independently adopted due to their accuracy and drift problems. This paper suggests a way of overcoming this challenge by experimenting on a sensor-focused solution without the use of CV based techniques.

## Introduction

Innovative technologies such as Augmented Reality (AR) are highly in demand to enable planning specialists and public participation in the areas of urban and architectural planning to preview and analyze environmental changes associated with proposed structures (Azuma et al. 2001, Bishop et al. 2005). In recent years, some AR systems have been built (Piekarski & Thomas, 2005) however the development of AR systems still has many unsolved problems, especially in the area of providing easy and correct integration of 2D/3D models, GIS datasets and mobile hardware peripherals for outdoor applications. Such integration must register all of software and hardware components within the same 6DOF (6 Degrees of Freedom, including both location and orientation) coordinating system. Here, all components must be determined in terms of their initial coordinates (referred to as registration) and the coordinates must be updated periodically (referred to as tracking).

Our work-in-progress project aims to investigate the potentials and technical challenges of outdoor mixed-reality environment and aims to optimize the use of on-the-spot urban and architecture planning for planners and local communities. Current visualization technologies lack the ability to provide the planners with a tool to view changes in situ. The potentials are planned to be exemplified by the fact that a real-time video stream representing a suburban streetscape environment and relevant artificial virtual information can be blended to construct a mixed reality that provides informative and intuitive environmental presentation. This presentation can possibly be interactive when it allows users to select new 3D models to substitute old models of in-the-street objects. In the mixed reality environment, houses of different architectural models, road lights of different designs and trees of different heights and

species can be tried to instantly preview potential environmental changes and thus help to make informed decisions. It is planned the project will allow planners to observe and analyze the environmental changes from different viewpoints while moving around the spot.

Registration and tracking is one of the most important technical aspects in this on-going AR project. For the time being, in terms of accuracy, stability and latency, the result of tracking is still not good enough for real-time outdoor AR applications. The tracking challenge has been encountered in the initial stage of this project. This paper focuses only on one technical challenge related to registration and tracking. It then suggests a way of overcoming this challenge by experimenting on a sensor-focused solution without much adoption of CV based techniques.

### *Tracking Techniques*

Some mathematical and computing techniques have been adopted to measure the spatial coordinate of a viewpoint for real-time uses in tracking such as Computer Vision (CV) approaches and hybrid approaches that combine CV and other sensory technologies. These two approaches will be discussed in this sub-section. The motivation for our new sensor-centric tracking approach will also be addressed.

Computer Vision is considered as a general method suitable for most AR registration and tracking algorithms, although it also brings some disadvantages. Many CV based tracking solutions rely on either fiducial markers or natural features. The fiducial markers approach is less suitable for outdoor applications due to its limited feasibility of physically marking multiple objects. The natural features approach, often referred to as markerless approach, is more adoptable in outdoor environments because it is powered by some other technologies such as visual servoing and computers model recognition. Most CV approaches require a number of computing processes to recognize the patterns of natural features. Many researchers adopt such approaches to solve the task of tracking the object model of a natural feature (Koller et al, 1997; Klinker et al, 1999; Stricker & Kettenbach, 2001, Behringer et al, 2002, Comport et al, 2006). In the computing process of recognizing natural features, the results of CV based tracking algorithms can be weakened by some factors, such as frequent and arbitrary changes of view angle, clarity of objects and brightness of sunlight. These factors cause many current tracking systems to suffer from high latency, low accuracy and low stability, and thus outdoor tracking is more difficult than its indoor use. These limitations have been experienced in many previous works. You et al (1999) track natural features to correct registration errors but too much computing processes are required which decreases their system's performance and causes a high latency. Kouroggi et al. (2000) uses panoramic image database for the comparison of natural feature patterns and Behringer et al (2002) uses a CAD model of buildings in an urban environment to correct the results of their traditional tracking system, but both cannot ensure the stability of tracking due to the complexity of running comparison algorithms and the unreliability of image processing at real-time. Satoh et al. (2001) summarize that heavy dependence on CV algorithms causes a balance between precision and frame rate, i.e. the higher the precision is achieved, the worse the frame rate is applied. For all vision-based methods, it is almost inevitable that environmental changes such as varying sunlight and air transparency may lead to unsatisfactory results. Thus, vision-based tracking systems are unreliable unless they can correctly and efficiently identify and track arbitrary features from frame to frame.

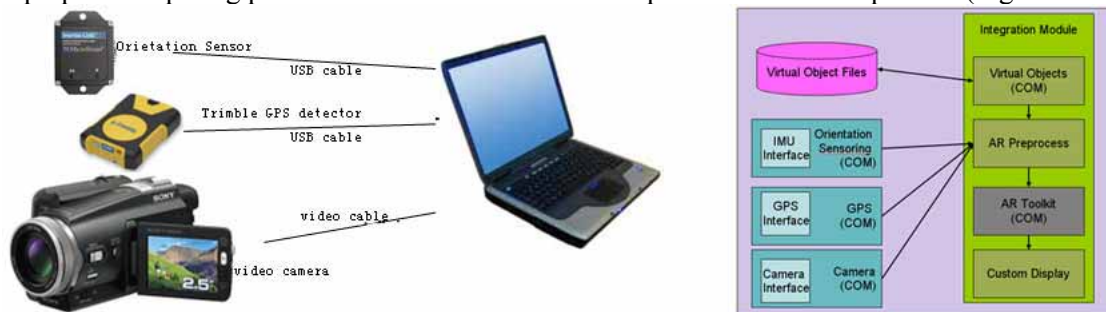
Some researchers study hybrid tracking/registration which combine vision-based methods with sensing technologies, however this hybrid process still requires further improvement. Satoh et al. (2001) have proposed a combination of gyroscope-based tracking with vision-based methods to compensate for measurement drift, but their tracking capability relies on already-known features that must be small enough to create an apparent discriminator. Wagner (2002) overcomes this restriction by using transitional markers but it only suits indoor applications. Piekarski et al. (2003) integrates standard outdoor orientation trackers with indoor fiducial trackers using GPS style world coordinates and however they states that such a hybrid system still suffers from a number of problems such as transitions between indoor and outdoor,

false marker detection and varying lighting condition. Ribo et al (2002) implements a hybrid system combining a vision-based tracker (using natural landmarks) with an inertial tracker. Although their system works with satisfactory performance in their prototype scenarios, it has to encounter some limitations caused by imprecise localization, the number of tracked landmarks, restrictions on camera technology and drawbacks of vision-based algorithm. Furthermore, Jiang et al. (2004) successfully present a robust hybrid systems comprised of gyroscopes and line-based vision tracking algorithms. This system only chooses high quality measurements for tracking, which consequently limits its application.

Both vision-based approaches and hybrid approaches suffer from the drawbacks of vision-based methods. Even in recent hybrid approaches, sensor-based approaches have been used but merely to complement vision-based approaches for better accuracy, stability and latency. Completely sensor-based tracking methods have been studied in some previous work. Azuma et al. (1999) presents one tracking system that combines a compass and tilt sensor with three rate gyroscope. Yuan et al. (2004) explores the motion parameters of a camera to register virtual objects through a homographic matrix provided from coplanar points. Baillot et al. (2003) presents a general calibration method to calculate the transformation among tracking system, virtual world and graphic display by manually aligning virtual objects over the display. With the recent advance of sensor technologies such as Differential GPS (DPGS) sensing and Inertial Measurement Unit (IMU) sensing, it is worthwhile to further explore sensor-centric tracking without involving much vision-based methods.

## Methodology

The aims of the project are (1) accuracy, latency and stability in the process of outdoor registering and tracking mainly relying on the adoption of vision-based methods; (2) thoroughly utilizing recent outdoor sensing technologies, such as Differential GPS (DGPS) sensing and Inertial Measurement Unit (IMU) to facilitate registration and tracking. The goal for tracking in this AR project is to achieve an accurate, reliable and efficient tracking which should allow for dynamic 6DOF movement of view points. It is not desirable to involve vision-based methodology because we intend to avoid all the limitations associated with vision-based or/and hybrid methods. A new tracking prototype system was proposed to address the possibility to achieve the goal. The tracking system mainly involves a DGPS detector, an orientation sensor, a laptop as computing platform and a video camera to capture live video sequences (Figure 1-left)



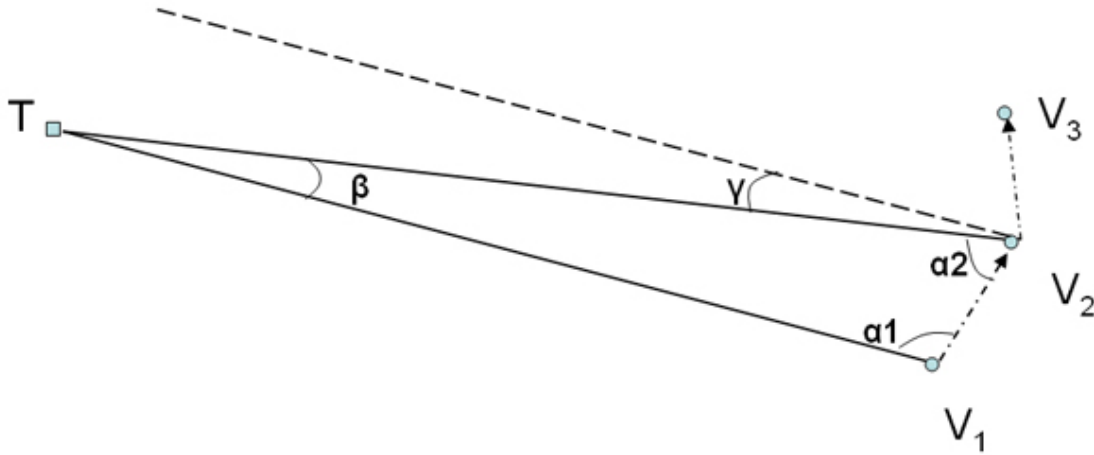
**Figure 1 The new tracking system (Left: hardware architecture, Right: software architecture)**

The application for which this system is being experimented is size-sensitive– it is to be carried and tested while moving around a selected outdoor spot; and all sensors are physically bound with the video camera. For these two reasons, the sensors are required to be small, lightweight and low-power. In order to ensure the accuracy is relatively good enough for our sensor-centric tracking, we studied many existing orientation sensors and selected “Inertialink”, an orientation sensor from Microstrain Inc., which supports 360° orientation range in pitch, roll and yaw. Pitch and roll detection are based on two accelerometers and their accuracy is at  $\pm 0.010$  g for  $\pm 15$  g range  $\pm 0.003$  g for  $\pm 2$  g range. The yaw/heading is based on a gyroscope of Inertialink at an accuracy of  $\pm 0.2^\circ/\text{sec}$  for  $\pm 300^\circ/\text{sec}$  until  $\pm 1200^\circ/\text{sec}$ .

A component-oriented software architecture is proposed to make all components within the integration of this system interchangeable and reliable (Figure 1–right). At the present stage,

COM (Component Object Model) based software modules have been chosen for this purpose. Orientation COM, GPS COM and Camera COM can communicate with the orientation sensor, the GPS detector and the camera through three components as communication interface – IMU interface, GPS Interface and Camera interface respectively. Prepared virtual objects are managed and imported to the system by Virtual Object COM. AR Toolkit COM is designated to mix all data together and then output an augmented view to the Custom Display process.

A sensor-based algorithm is proposed to calculate the orientation of view points relative to a target object as the camera moves. This way, the tracking system can know exactly when and where to display the target object on the monitor. 3DOF orientation can be projected onto three projected planar views: the pitch plane, the roll plane and the heading plane. We assume the following calculation only addresses the orientation changes on the heading plane and its planar view is illustrated in Figure 2. The same calculation can be similarly used for the projection on the pitch and roll planes.



**Figure 2 the calculation of the changes on the baseline orientation**

Let T be the target place of a virtual object in our mixed reality environment. We assume the spatial information (longitude, latitude and height) of T, i.e. T(LLH), has already been known. Let V be the view point where the mobile camera is located. The current direction  $\overrightarrow{VT}$  is defined as the direction from V to T.  $\overrightarrow{VT}$  changes as V moves, as the following steps describe.

Step (1): Let V1 be the initial view point of the mobile camera. At V1, set the orientation's pitch P and roll R to 0 and heading H to the yaw value sampled at the starting time.

Step (2): Assuming V moves from V1 to V2 at a time interval (saying two seconds),  $\overrightarrow{VT}$  changes from  $\overrightarrow{V1T}$  to  $\overrightarrow{V2T}$ . As illustrated in Figure 2,  $\overrightarrow{V2T} = \overrightarrow{V1T} - \gamma$ . Geometry tells  $\gamma$  should be equal to  $\beta$  where  $\beta = 180^\circ - \alpha1 - \alpha2$ .

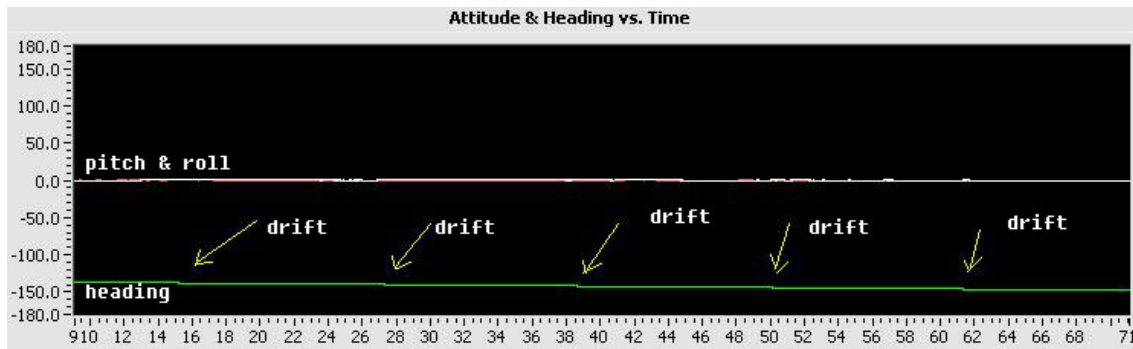
Step (3): Get  $P_1(LLH)$  and  $P_2(LLH)$  from the GPS detection. As T(LLH),  $P_1(LLH)$  and  $P_2(LLH)$  are all known, use Vincenty's Inverse formulae (Vincenty, 1975) to calculate  $\alpha1$  and  $\alpha2$ . Thus,  $\gamma$  can be worked out.

Step (4): Update the current direction with  $\overrightarrow{V2T}$  where  $\overrightarrow{V2T} = \overrightarrow{P1T} - \gamma$ .

Step (5): If V moves to next point V3, then set V1 to V2 and V2 to V3, and loop from Step (2) until tracking is stopped.

### Preliminary Findings and Discussions

The new tracking prototype system is currently being developed. However, at the initial stage of the development, some preliminary findings have been achieved. We experimentally tested the orientation sensor and found that a drift problem always occurs in the gyroscope-based detection of heading, as illustrated in Figure 3. Drift may be eliminated if a magnetometer-based detector is on-board to detect the heading direction by measuring Earth's local magnetic field. Because it makes the heading referenced to the magnetic poles not the true poles, we could use declination to achieve a true polar reference when necessary.



**Figure 3 the drift of heading**

The proposed algorithm has also been experimented by manually testing some sample data. It is found that the offset between  $d$  and  $w$  must be limited to a large number relatively close to the value of  $d$ . The larger it is, the more isosceles-like the triangle would be. However, if the sampling rate of GPS detector is high, this number will be very close to the value of  $d$ . This implies that the view point moves too shortly at a very short time and the whole calculation has to repeat frequently. This may cause a performance issue from frame to frame in the video sequence. As a result, the sampling rate should be carefully set up. According to one's normal walking speed, we suggest a rate ranging from 0.25Hz to 0.50Hz. In addition, if  $d$  is less than 30 meters, the algorithm may not provide enough accuracy unless the view point doesn't move too much. The final result of this situation is to be identified.

### **Conclusion and Future Work**

In this paper, we have introduced our on-going AR project for urban and architectural planning, and have addressed the tracking challenges and techniques in outdoor AR systems. We presented a sensor-based tracking method for outdoor AR systems. The method is novel in two aspects. First, its process is easy and intuitive and second it doesn't rely on vision-based methods and thus avoid their limitations.

The development of the tracking system is ongoing to achieve the aims of the project. The algorithm will be slightly altered for the application on the plane of pitch and roll. It is envisaged that a method to update can be developed with the algorithm for those scenarios that don't have the known location of  $T$ . Other non-vision based techniques will be tested to complement our results.

### **Acknowledgments**

This project has been supported through Monash University's ART/IT Small Grant (2006-2007) jointly funded by Faculty of Arts and Faculty of Information Technology.

### **References**

- Azuma R., Lee J.W., Jing B., Park J., You S. and Neumann U. (1999), "Tracking in unprepared environments for augmented reality systems" *Computers and Graphics*, pp.787-793.
- Behringer R., Park J and Sundareswaran V. (2002), Model-based visual tracking for outdoor augmented reality applications, *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR'02)*.
- Comport A., Marchand E., Pressigout M. and Chaumette F. (2006), "Real-Time Markerless Tracking for Augmented Reality: The Virtual Visual Servoing Framework", *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS*, VOL. 12, NO. 4.
- Jiang B., Neumann U. and You S. (2004), "A robust hybrid tracking system for outdoor augmented reality", *Proceedings of Virtual Reality IEEE*, pp275
- Klinker G, Behringer R and Mizell D. (1999), "International Workshop on Augmented Reality", edited by Mike Bevan, *VR News*, Vol. 8, Issue 1, pp. 18-20.

Koller D.,Klinker G.,Rose E.,Breen D.,Whitaker R.and Tuceryan M. (1997),"Real-time Vision-Based Camera Tracking for Augmented Reality Applications", Proceedings of the Symposium on Virtual Reality Software and Technology (VRST-97),pp. 87–94.

Kouroggi M, Kurata T., Sakaue K., and Muraoka Y. (2000), “ A panorama-based technique for annotation overlay and its real-time implementation”, the Proceedings of International Symposium on Wearable Computers (ISWC 00), pp657-660.

Kouroggi M, Kurata T., Sakaue K., and Muraoka Y. (2000), “ Improvement of panorama-based annotation overlay using omnidirectional vision and inertial sensors”, Proceedings of International Symposium on Wearable Computers (ISWC 00), pp183-184.

Piekarski W.,Avery B.,Thomas B. and Malbezin P. (2003), "Hybrid Indoor and Outdoor Tracking for Mobile 3D Mixed Reality",Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'03).

Piekarski, W. and Thomas, B.H. (2005), “Future Use of Augmented Reality for Environmental and Landscape Planners. In: Visualization in landscape and environmental planning: technology and applications”, Taylor & Francis: Oxon, pp. 234-240.

Ribo M.,Lang P.,Ganster H.,Brandner M.,Stock C.and Pinz a., "Hybrid Tracking for Outdoor Augmented Reality Applications", Computer Graphics and Applications, IEEE, pp.54-63, 2002

Stricker D. and Kettenbach T. (2001), “Real-time and Markerless Vision-based Tracking for outdoor Augmented Reality applications”, the Proceedings of the IEEE and ACM International Symposium on Augmented Reality (ISAR 01), pp. 189.

Satoh K., Anabuki M., Yamaoto H. and Tamura H. (2001), “A Hybrid registration method for outddor augmented reality”, Proceedings of International Symposium on Augmented Reality.

You S. Neumann U. and Azuma R. (1999), “Orientation tracking for outdoor augmented reality registration”, IEEE computer Graphics and Applications, Vol. 19, No. 6, pp 36-42.

Yuan Z., Gong Z, Wu J. and Kanou Y. (2004), “A Registration Method of AR System Applicable to Outdoor Environment”, Proceedings of International Conference on Intelligent mechatronics and Automation, Chengdu, China.

Vincenty T. (1975),"Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations", Survey Review 23, Number 176, pp.88-93.

Wagner M. (2002), "Building Wide-Area Applications with the AR Toolkit", Augmented Reality Toolkit, The 1<sup>st</sup> IEEE International Workshop, pp.7.