User Tracking Methods for Augmented Reality

Erkan Bostanci, Nadia Kanwal, Shoaib Ehsan, and Adrian F. Clark

***Abstract*—Augmented reality has been an active area of research for the last two decades or so. This paper presents a comprehensive review of the recent literature on tracking methods used in Augmented Reality applications, both for indoor and outdoor environments. After critical discussion of the methods used for tracking, the paper identifies limitations of the state-of-the-art techniques and suggests potential future directions to overcome the bottlenecks.**

***Index Terms*—Augmented reality, tracking.**

I. Introduction

Augmented Reality (AR) is the blending of real-world images with artificial objects or information generated by a computer. It is also defined as the extension of user's environment with synthetic content [1]. For more than two decades, AR has been a topic of interest in the field of computer graphics as it enriches human perception and facilitates understanding of complex 3D scenarios [2], [3]. AR applications are now becoming more popular due to their capability to run on a variety of platforms such as mobile computers and even cell phones [4].

Tracking, the process of locating a user in an environment, is critical to the accuracy of AR applications as more realistic results can be obtained in the presence of accurate AR registration. It usually includes determining the position and orientation of the AR user. Generally, the most important part is tracking the head as the user wears a Head Mounted Display (HMD) from which the augmented images of the real world are displayed. The improved accuracy of the AR system due to tracking also prevents problems such as visual capture [2] and does not allow visual sensors to gain a priority over other sensors. For instance, inadequate registration accuracy can cause the user to reach wrong part of the real environment because the augmentation has been displayed on another part. The eyes of the users get used to the error in the virtual environment and after some time of usage they start to accept these errors as correct which is not desirable.

This paper reviews the state-of-the-art tracking methods used for AR, identifies the bottlenecks involved and proposes future research directions. The paper is structured as follows: Section 2 discusses the tracking techniques for indoor and outdoor environments, fusion methods and a recent set of methods which were very well known in the robotics community but are new to computer graphics. The limitations of the currently used methods are identified in Section 3. Future research directions that can be explored are proposed in Section 4 which is followed by conclusions in Section 5.

Manuscript received June 1, 2012; revised July 2, 2012.

The authors are with the School of Computer Science and Engineering, University of Essex, UK (e-mail: gebost@essex.ac.uk, nkanwa@essex.ac.uk, sehsan@essex.ac.uk, alien@essex.ac.uk).

II. Tracking Methods

A variety of tracking methods for different applications are found in literature [5]. This section provides a review of different tracking methods used for AR applications under four main categories: indoor methods, outdoor methods, fusion strategies and recent approaches as shown in Figure 1.

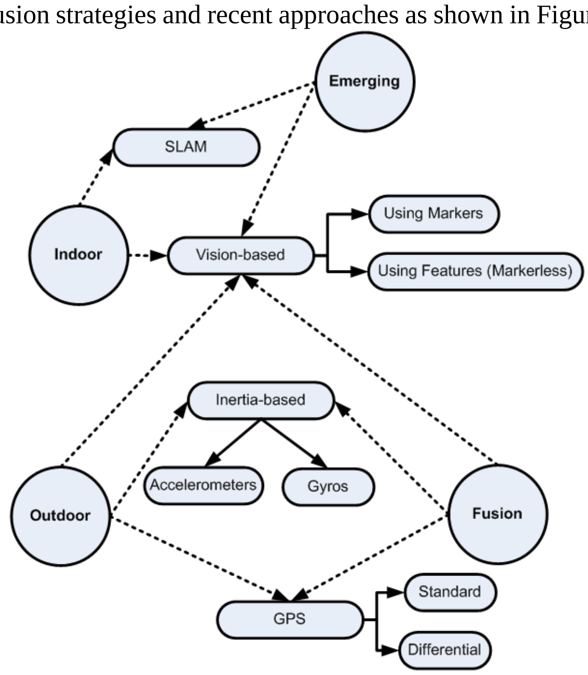

Fig. 1. Tracking methods for augmented reality.

III. Indoor Techniques

Indoor environments provide a structured domain for an AR application and movements of the user are limited only to a specified region. In [6], it is stated that for an indoor space, the dimensions of the environment are fixed and the user's possible movements are more predictable. The structured domain also provides power for the tracking equipment and presents a controlled environment [7].

Before proceeding, it is important to understand the term "marker" used in the context of these methods. Fiducial markers are distinguishable elements put in the environment so that they can be identified apart from other objects in the same environment. These markers can be categorized as active or passive markers. Active markers emit a signal (e.g. magnetic, light) which can be sensed by the sensor. Passive markers tend to be a pattern which can be easily isolated from the texture of the environment (e.g. QR codes). In this case, computer vision methods must be applied out to recognize the marker. Markers are sometimes also referred as beacons

DOI: 10.7763/IJCTE.2013.V5.654

and landmarks in different applications.

Indoors tracking is generally achieved by two methods: outside-in and inside-out as described in [8]. Names of these methods present clues about the location of the sensor, which can be magnetic, ultrasonic, radio frequency identification (RFID) sensors or a camera, and how the tracking is achieved. In the first method, the sensor is fixed to a place in the environment. The user wears a hat-like item on which the fiducially markers are mounted. As the name suggests, the sensor is placed somewhere outside the user (outside) but is sensing the markers on the user (in). Vice versa, the user carries the sensor and the fiducially markers are mounted around the environment (certainly within the sensor's range or field of view) in the second method. As the locations of these markers are well-known in advance, tracking can be achieved.

Although there are many different types of indoor tracking methods using magnetic or ultrasound sensors, these systems generally use both expensive and complex hardware [9,10]. Although GPS is a good option for tracking user position outdoors, indoor environments like laboratories or buildings generally block these signals. The uncertainty of GPS signal in the indoor environments calls for more reliance on vision-based tracking systems.

A wide range tracking system named HiBall was presented in [8]. The aim of the study was an accurate, robust and flexible system to be used in large indoor environments. The HiBall device was designed as a hollow ball with a dodecahedron shape. Upper part of the device was fitted with 6 lenses. LEDs controlled by an interface board were mounted on the ceiling in the laboratory.

A complete tracking system named Video Positioning System (VPS) was designed and developed in [11]. This system used fiducial markers. A new fiducial pattern design was introduced which allows using unique patterns for accurate position and orientation calculation. The pattern design was based on RAG (region adjacency graph) with two parts namely key and the identifier one of which indicated that this shape is a marker and the second differentiated between different markers in the system.

VPS was also applied on a parallel architecture in [12] and it was shown that parallelization improved the performance of some parts in the system for real-time operation.

In [13], a comparison of the VPS system given in [11] and the ARToolkit was presented. The results showed that VPS provided more accuracy than the popular ARToolkit system against moderate changes in viewpoint and distance. However ARToolkit performed better when distance is increased and the authors suspected that this was due to the design of the fiducial markers.

Chia *et al.* [14] developed a camera tracking system based on natural features. The system used precaptured reference images of the scene and then RANSAC was used for robust matching to achieve invariance in motions of the feature points. The system was able to run at 10Hz using some fiducial markers as well.

Park *et al*. tracked several 3D objects simultaneously, robustly and accurately in real-time [15]. Frame-to-frame tracking was found to be less computationally demanding but it was prone to fail and detection was more robust but it was slower. These techniques were combined to benefit from their advantages. For each target object a 3D CAD model and a small set of reference images, named as keyframes, were available. Key point detection and pose estimation was performed in one thread and key frame detection was performed in another thread working in parallel. The system was able to work at 15-20fps on a 3.2 GHz multicourse CPU, though the performance deteriorated when the number of objects increased.

Bekel [16] presented a viewpoint based approach for AR. This method used Self-Organizing Map (SOM) to train as a classifier which is later used to label different types of objects by overlaying in the scene.

Adams *et al.* [17] developed a method for aligning viewfinder frames obtained from the camera of a mobile phone and they applied their system to 3 different cases including night-view, panorama and an input method for the camera instead of shaking. The authors stated that two algorithms were required for alignment. First was the generation of the digest by extracting edges in horizontal, vertical and two diagonal directions and a set of point features. The second part was the alignment of edges. This gave the translation between two frames. Then by using the point feature correspondences the confidence of the initial translation is obtained. The alignment algorithm developed in this study was fast and robust against noise however it was very fragile against rotations where rotations greater than 1.5 degrees were reported to create problems.

A different approach for pose tracking in a built-in camera of a mobile phone was followed by Wagner *et al.* in [18]. They have used SIFT for robust features and Ferns for classification. Ferns is a fast classification method however it requires a great amount of memory. To alleviate the computational expense of the SIFT method it was altered by removing the calculations required for scale invariance. Instead, a database of the features at different scales was used for the same purpose. FAST was used for corner detection for its high repeatability.

VisiTrack system was developed in [19] for tracking in mobile devices using point and edge extraction together with colour segmentation. Although the system was claimed to provide markerless localization a marker can be seen in the test sequences in the system running at 25fps.

For the indoor AR system in [20], visual tracking was used. The system recognized image views of the environment acquired beforehand. Processing was carried out by remote PCs via a wireless LAN.

IV. Outdoor Techniques

Indoor environments are generally more predictable whereas the outdoor environments are usually limitless in terms of location and orientation. As opposed to indoors, there is less chance for preparing the environment to track the user while working outdoors. Moreover, predefined artificial landmarks cannot be used and natural landmarks need to be found. Also, varying light poses a problem for camera tracking which is not an issue for indoors.

As mentioned earlier, GPS is considered a good tracking option when working outdoors. A comparison of different

GPS receivers including different brands such as Trimble, Garmin and DeLorme is given in [6].

A differential GPS and a compass was used for position and orientation estimation in [21]. Latitudes and longitudes of several viewpoints were stored in a database along with the set of images taken at different times of the year with varying light conditions.

Reference images were used for video tracking and matching was performed to find these reference images for the outdoor AR system in [22]. A video image was compared with the reference images and a matching score was obtained. For the best matching score, the 2D transformation was calculated and the current camera position and orientation were deduced. This transformation was used to register the model on the video frame. The matching technique was based on Fourier Transformation to be robust against changes in lighting conditions hence it was limited to only 2D transformations such as rotation and translation. This technique had a fixed number of computations therefore it was suitable for real-time operation without using markers, however it worked on 10Hz which is a low rate for real-time display.

Inertial sensing is a widely used method since its operation is similar to the otolith stones in human ear [23]. Accelerometers are used for translational motion and gyros for rotational motion. This method is generally used together with other tracking methods as it will be given in more detail in Fusion Strategies.

For the tracking of the user's location in [24] in the ancient city of Paestum in Italy, a WiFi system was planned by installing antennas instead of using GPS. However, this was not implemented due to the opposition to changes in the original archaeological site by the archaeologists [22].

Tracking for a walking or running user was performed using a custom tracking system in [25]. The system used an inertial sensor, electromagnetic sensor, push-buttons in the heels of the user and trackers in the knees so that the actual motion of the legs could be obtained. The transmitter was mounted above user's waist so that the relative motion of the legs could be extracted when the user's foot does not ground.

V. Fusion Strategies

When the above mentioned methods are used as the only sensors, accuracy of the tracking may be low. However, very accurate systems can be obtained by using different sensors together.

Fusion methods are classified as loosely and tightly coupled systems [23]. In loosely coupled systems, the sensors act separately and perform calculations regardless of each other. However, in tightly coupled systems, a sensor fusion is used i.e. calculations are performed together to generate a single and improved position estimation.

Visual inertial tracking is a very popular technique, due to the complementary characteristics of both sensors, used in many different applications. Vision allows estimation of the camera position directly from the images observed [26]. However, it is not robust against 3d transformations, and the computation is expensive. For inertial trackers, noise and calibration errors can result in an accumulation of position and orientation errors in inertial trackers. Vision is good for small acceleration and velocity. Inertial sensors have long term stability problems [27].

When these sensors are used together, faster computation can be achieved with inertial sensors and the drift errors of the inertial sensor can be corrected with vision. Applications generally use low frequency vision data and high frequency inertial data [28] since visual processing is more expensive and trackers today can generate estimates at rates up to 550Hz using custom hardware [29].

A hybrid tracking system was developed by [3] for mobile outdoor AR. The system combined vision-based methods and inertial trackers. Developed inertial tracker hardware includes 3 accelerometers, 3 gyroscopes and a Digital Signal Processor (DSP). Three devices were used to track accelerations in x, y and z coordinates. The vision system used point features and calculated the 6 degree-of-freedom (DOF) camera pose using Perspective-N-Points (PnP).

Another visual-inertial tracker system was developed by Foxlin *et al.* [30]. One or more cameras could be added to the system. The authors developed a fusion filter to combine visual and inertial measurements. Artificial fiducial markers were arranged in a circular matrix and used by the system.

A self-contained tracking system for outdoor using inertial and visual sensors was developed in [7]. The system used a fiducial design based on colours for indoor AR application.

You *et al.* [26] developed a hybrid system for accurate registration in AR. A prediction-correction method was used in this system. The data obtained from the inertial sensor was used to estimate 2D feature motion (2D prediction) and then visual feature tracking was employed to correct the estimate (2D correction). Finally, a 3D correction was performed on the by the gyroscopes from 2D motion residual.

User tracking in [6] was performed with GPS and head trackers and a camera was only used for view. The system components included Trimble AgGPS 332 Receiver, TCM5 3-axis orientation tracker, Wristpc wearable keyboard2, Cirque smart cat touch pad3, i-glasses SVGA HMD and a laptop.

Inspired by a desktop optical mouse and based on the "Anywhere Augmentation" paradigm introduced by the authors for outdoor AR, a tracking system with a camera aiming directly to the ground and an orientation tracker was developed in [31]. The system additionally used GPS to prevent long term drift of the system.

Haala *et al.* [32] used a low-cost GPS and a digital compass for positioning in an urban environment. The authors applied shape matching with the 3D model of the building and the actual building. When the system found a match, the 3D model was overlayed in the video.

Piekarski [33] developed an outdoor AR system by using a Trimble Ag132 GPS unit and an orientation tracker and achieved an accuracy of less than 50cm. In the software, the user was able to define the corners of 3D model to be drawn with a pinch-glove. The marker on the glove was tracked by the system to define the corners of 3D model to be drawn.

A different outdoor application which aimed to display an archaeological site to users was given in [34]. The system used GPS together with inertial sensors provided within the HMD.

Sherstyuk [35] developed a novel method for fast semiautomatic 3D geometry acquisition using motion tracking equipment which was intended for quick surface prototyping where quality is not of high priority. A life size medical mannequin (articulated doll) is used by additional touch sensitivity to arbitrary locations of the mannequin. Then a surface scanning method was used to track the motion of the user and generate the 3D reconstruction of the mannequin for medical visualization.

VI. Emerging Approaches

Finding position and orientation of the agent is an issue in both tracking in AR and robotics. There has been a vast amount of research in the robotics field about this topic. Algorithms known as Simultaneous Localization and Mapping (SLAM) have been developed to localize a robot based on the map which it creates by observing its environment.

Since SLAM algorithms can be applied for a robot, they can also be applied to a camera mounted on the user in an AR context. Conventional tracking techniques explained before had their own advantages and disadvantages in different situations. Application of SLAM algorithms on tracking brings new initiatives to current state-of-the-art systems.

An interactive and interesting application was presented in [36] where Chekhlov *et al.* applied EKF SLAM given in [37] to an AR game in which a ninja tries to jump from one plane to another until he reaches the target plane. Higher level structures such as planes were created from point feature sets using RANSAC and OGRE game engine was used to implement the game.

Bleser [38] investigated robustness and accuracy of realtime markerless augmented reality in both known and unknown environments for AR. Bleser used sensor fusion of IMU and a camera with particle filtering. CAD model of the environment or an object was matched with the actual object. The tests showed operation in a small environment 2.1m. × 1.5m. × 2.5m. and a conceptual solution for large environments was presented.

One of the most impressive results for AR using SLAM was presented by Klein *et al.* [39]. Their system used markerless tracking in a parallel system. Two threads were executed for tracking and mapping of the environment. They also presented interesting AR applications such as the glass magnifier.

Kozlov *et al.* [40] proposed using AR as an approach to visualize the internal state of a robot in order to test and debug SLAM algorithms. This approach presented a different point of view as the methods mentioned above used SLAM for AR. The authors proposed using visualization for robot pose, state map and data association where cross correlations could be used to show the decrease of uncertainty in the map.

VII. Problems with Current Approaches

Current tracking systems still have many problems. Two types of errors were defined in [2] namely static and dynamic errors. Before giving details about problems when using different types of sensors, these two important terms will be explained.

The errors in tracking systems are considered as static errors due to the inadequate accuracy provided by the commercial systems at the time being. The dynamic errors are delays. The end-to-end system delay is the time elapsed between the time when the tracking system to measure the position and orientation of the user to the time when the images appear on the display.

Vision methods allow both tracking and managing residual errors. They are low cost. The problem with these methods is the lack of robustness [41]. For some applications e.g. [31], there may be a great probability of incorrect matches since the texture of the ground is mostly similar in different areas as a repeating pattern. It was stated in [42] that the structure of AR models is more difficult than Virtual Reality (VR) since the former uses geometry not only for visualization but also for occlusion handling and vision based tracking of scene features.

Also for visual tracking, the features to be used as landmarks should be invariant to changes in lighting and viewpoint. Since this is not always possible vision based tracking outdoors are reported to be very fragile [3]. Using camera as the only sensor was found to be accurate but expensive in computation [43].

Standard GPS has an error in order of 10m. However, it improves when a differential GPS or real-time kinematic (RTK) correction is used. Line of sight can be a problem for the satellites in urban environments [23] or under dense tree canopy due to variance [44]. Other problems with GPS were explained in detail in [45]. The system developed in [46] reported about the tracking problems occurring when GPS was used as the only sensor and the authors suggested using sensor fusion or GPS dead-reckoning.

Double integration in inertial trackers cause drift errors to propagate rapidly [31]. Active tracking systems require calibrated sensors and signal sources in a prepared environment and tracking equipment can be affected by signal noise, degradation with distance and interference sources [41]. Magnetic trackers can be interfered with the ferro-magnetic objects in the environment [47]. A system with gyros and accelerometers provide good bandwidth however it is prone to integrated errors and long term stability is not guaranteed [48].

Other problems also come into consideration as well as other SLAM systems used in robotics. First of all data association which means finding a correspondence between the feature model and the observed feature due to low precision and recall rates [49]. The second problem is linearization due to the characteristics of current SLAM methods. Linearization problems both affect the filter stability and convergence resulting in less accurate localization [50].

VIII. Future Directions

Azuma stated that AR tracking for outdoors in real-time with required accuracy is an open problem [2]. Though more than a decade has passed after his statement, this problem still keeps its validity since AR requires high accuracy, low latency and low jitter [3]. Similarly in [30], it was stated that there was a great need for a self-tracker that can be used in natural outdoor environments as well however a robust implementation of such a tracker was years away due to the challenges in finding robust features in natural environments.

AR has the potential for many different and interesting applications including entertainment such as the games in [1, 51] or cultural heritage applications as in [22, 34] when outdoors are considered.

Most outdoors methods used GPS and inertial trackers for this purpose [6, 34, 46, 51]. Vision based tracking has also been applied for some systems as in [21, 22, 41]. Inertial sensor can be used for stability and robustness in cases of rapid motion or occlusion [43].

Current applications of using SLAM for AR as [36, 43, 52, 53] are limited to desktop or laboratory environments though they presented accurate results in tracking. In [52] localization was performed according to known 3D junctions. AR tests were carried out with a rectangular pattern in the view at all times [52]. Similar results can be seen in the work of [43].

Considering the methods used today, we have come up with the following ideas and suggest them for future research:

1) A fusion of different sensors within a SLAM framework is a promising approach to follow.
2) Vision-based tracking is quite useful because we already need images of the environment for augmentation. As we have this information source at hand, it is wise to use it for both tracking and augmentation.
3) The introduction of robust detectors such as SIFT or SURF will improve the visual tracking process. However they are considered as an important barrier to achieve full frame-rate operation in real-time [54].
4) For performance considerations, graphics hardware or parallel implementation are suggested.

We believe that more accurate systems for outdoors AR tracking are feasible using the methods mentioned above.

IX. Conclusion

In this paper, we tried to present several application examples of tracking methods for AR for indoor and outdoor environments. These methods were examined form a critical point of view considering both advantages and disadvantages of using different sensors for the tracking process.

An emphasis was made on visual tracking methods due to their increasing popularity in the literature. With the new methods developed from computer vision community as mentioned earlier and a fusion of vision-based methods with other sensors, we believe that the accuracy of tracking will reach sufficient levels for real-time rendering and augmentation.

References

[1] W. Broll, I. Lindt, I. Herbst, J. Ohlenburg, A. Braun, and R. Wetzel, "Toward next-gen mobile ar games," *IEEE Computer Graphics and Applications*, pp. 40–48, 2008.

[2] R. Azuma, "A survey of augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, pp. 355–385, 1997.

[3] M. Ribo, P. Lang, H. Ganster, M. Brandner, C. Stock, and A. Pinz, "Hybrid tracking for outdoor augmented reality applications," *IEEE Computer Graphics and Applications*, pp. 54–63, 2002.

[4] G. Papagiannakis, G. Singh, and N. Magnenat-Thalmann, "A survey of mobile and wireless technologies for augmented reality systems," *Computer Animation and Virtual Worlds*, vol. 19, pp. 3–22, 2008.

[5] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Computing Surveys*, vol. 38, no. 4, 2006.

[6] A. Behzadan, B. W. Timm, and V. R. Kamat, "General purpose modular hardware and software framework for mobile outdoor augmented reality applications in engineering," *Advanced Engineering Informatics*, vol. 22, pp. 90–105, 2008.

[7] U. Neumann, S. You, Y. Cho, J. Lee, and J. Park, "Augmented reality tracking in natural environments," *International Symposium on Mixed Reality*, 1999.

[8] G. Welch, G. Bishop, L. Vicci, S. Brumback, K. Keller, and D. Colucci, "High-performance wide-area optical tracking: The hiball tracking system," *Presence: Teleoperators and Virtual Environments*, pp. 1–22, 2001.

[9] A. Golding and N. Lesh, "Indoor navigation using a diverse set of cheap, wearable sensors," *Third International Symposium on Wearable Computers*, pp. 29–36, 1999.

[10] A. Cheok and Y. Li, "Ubiquitous interaction with positioning and navigation using a novel light sensor-based information transmission system," *Personal and Ubiquitous Computing*, vol. 12, pp. 445-458, 2008.

[11] D. Johnston and A. Clark, "A vision-based location system using fiducials," *Vision, Video, and Graphics*, pp. 1–8, 2003.

[12] D. Johnston, M. Fluery, A. Downton, and A. Clark, "Realtime positioning for augmented reality on a custom parallel machine," *Image and Vision Computing*, vol. 23, no. 3, pp. 271–286, 2005.

[13] K. Yeung, D. Johnston, and A. Clark, "A comparison of fiducial-based visual positioning systems," presented at International Conference on Pattern Recognition, 2006.

[14] K. Chia, A. Cheok, and S. Prince, "Online 6 dof augmented reality registration from natural features," *International Symposium on Mixed and Augmented Reality*, pp. 305– 313, 2002.

[15] Y. Park, V. Lepetit, and W. Woo, "Multiple 3d object tracking for augmented reality," *IEEE International Symposium on Mixed and Augmented Reality*, pp. 117–120, 2008.

[16] H. Bekel, G. Heidemann, and H. Ritter, "Interactive image data labeling using self-organizing maps in an augmented reality scenario," *Neural Networks*, vol. 18, pp. 566–574, 2005.

[17] A. Adams, N. Gelfand, and K. Pulli, "Viewfinder alignment," *EUROGRAPHICS*, vol. 27, pp. 597–606, 2008.

[18] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg, "Pose tracking from natural features on mobile phones," *IEEE International Symposium on Mixed and Augmented Reality*, pp. 125–134, 2008.

[19] D. Stichling, N. Esau, B. Kleinjohann, and L. Kleinjohann, "Real-time camera tracking for mobile devices: The visitrack system," *Real-Time Systems*, pp. 279–305, 2006.

[20] J. Kim and H. Jun, "Vision-based location positioning using augmented reality for indoor navigation," *IEEE Transactions on Consumer Electronics*, vol. 54, pp. 954–962, 2008.

[21] P. Dahne and J. Karigiannis, "Archeoguide: System architecture of a mobile outdoor augmented reality system," *IEEE International Symposium on Mixed and Augmented Reality*, 2002.

[22] D. Stricker and T. Kettenbach, "Real-time and markerless vision-based tracking for outdoor augmented reality applications," *International Symposium on Augmented Reality*, 2001.

[23] P. Corke, J. Lobo, and J. Dias, "An introduction to inertial and visual sensing," *International Journal of Robotics Research*, vol. 28, pp. 519–535, 2007.

[24] R. Chiara, V. Santo, U. Erra, and V. Scanarano, "Real positioning in virtual environments using game engines," *Eurographics Italian Chapter Conference*, 2006.

[25] Y. Yamanaka, M. Kanbara, and N. Yokoya, "Localization of walking or running user with wearable 3d position sensor," presented at the 17th International Conference on Artificial Reality and Telexistence 2007, 2007.

[26] S. You, U. Neumann, and R. Azuma, "Orientation tracking for outdoor augmented reality registration," *IEEE Virtual Reality*, pp. 36–42, 1999.

[27] S. G. Chroust and M. Vincze, "Fusion of vision and inertial data for motion and structure estimation," *Journal of Robotic Systems*, vol. 21, no. 2, pp. 73–83, 2004.

[28] J. Chen and A. Pinz, "Structure and motion by fusion of inertial and vision-based tracking," *Austrian Association for Pattern Recognition*, vol. 179, pp. 75–62, 2004.

[29] P. Lang, A. Kusej, A. Pinz, and G. Brasseur, "Inertial tracking for mobile augmented reality," presented at IEEE Instruments and Measurement Technology Conference, 2000.

[30] E. Foxlin and L. Naimark, "Vis-tracker: A wearable vision-inertial self-tracker," *IEEE Virtual Reality*, 2003.

[31] S. Diverdi and T. Hollerer, "Heads up and camera down: A vision-based tracking modality for mobile mixed reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, pp. 500–512, 2008.

[32] N. Haala and J. Bohm, "A multi-sensor system for positioning in urban environments," *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 58, pp. 31–42, 2003.

[33] W. Piekarski, "3d modeling with the tinmith mobile outdoor augmented reality system," *IEEE Computer Graphics and Applications*, pp. 14–17, 2006.

[34] P. D. Ritsos, "Architectures for untethered augmented reality using wearable computers," Ph.D. dissertation, Dept. Elect systems Engineering, Essex Univ., 2006.

[35] A. Sherstyuk, A. Treskunov, and B. Berg, "Fast geometry acquisition for mixed reality applicaitons using motion tracking," presented at IEEE International Symposium on Mixed and Augmented Reality, 2008.

[36] D. Chekhlov, A. Gee, A. Calway, and W. Mayol-Cuevas, "Ninja on a plane: Automatic discovery of physical planes for augmented reality using visual slam," presented at IEEE International Symposium on Mixed and Augmented Reality, 2007.

[37] A. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "Monoslam: Real-time single camera slam," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 1–16, 2007.

[38] G. Bleser, "Towards visual-inertial slam for mobile augmented reality," Ph.D. dissertation, Fachbereich Informatik der Technischen Universitt Kaiserslautern, 2009.

[39] G. Klein and D. Murray, "Parallel tracking and mapping for small ar workspaces," in *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007.

[40] A. Kozlov, B. Macdonald, and B.Wunsche, "Towards improving slam algorithm development using augmented reality," presented at Australasian Conference on Robotics and Automation, 2007

[41] U. Neumann and S. You, "Natural feature tracking for augmented reality," *IEEE Transactions on Multimedia*, vol. 1, no. 1, pp. 53–64, 1999.

[42] D. Schmalstieg, G. Schal, D. Wagner, I. Barakonyi, G. Reitmayr, J. Newman, and F. Ledermann, "Managing complex augmented reality models," *IEEE Computer Graphics and Applications*, pp. 48–57, 2007.

[43] G. Bleser, C. Wohlleber, M. Becker, and D. Stricker, "Fast and stable tracking for ar fusing video and inertial sensor data," presented at International Conferences in Central Europe on Computer Graphics, Visualization and Computer Vision, 2006.

[44] K. Konolige, M. Agrawal, R. Bolles, C. Cowan, M. Fischler, and B. Gerkey, "Outdoor mapping and navigation using stereo vision," *Experimental Robotics*, vol. 39, pp. 179–190, 2008.

[45] T. Bailey, "Mobile robot localisation and mapping in extensive outdoor environments," Ph.D. dissertation, Australian Centre for Field Robotics, Department of Aerospace, Mechanical and Mechatronic Engineering The University of Sydney, 2002.

[46] G. Schall, E. Mendez, E. Kruijff, E. Veas, S. Junghanns, B. Reitinger, and D. Schmalstieg, "Handheld augmented reality for underground infrastructure visualization," *Personal Ubiquitous Computing*, vol. 13, pp. 281–291, 2009.

[47] D. Koller, G. Klinker, E. Rose, D. Breen, R. Whitaker, and M. Tuceryan, "Real-time vision-based camera tracking for augmented reality applications," *Symposium on Virtual Reality Software and Technology*, 1997.

[48] H. Rehbinder and B. Ghosh, "Pose estimation using linebased dynamic vision and inertial sensors," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 186–199, 2003.

[49] R. Sim, P. Elinas, and J. Little, "A study of the raoblackwellised particle filter for efficient and accurate visionbased slam," *International Journal of Computer Vision*, vol. 74, no. 3, pp. 303–318, 2007.

[50] P. Pathirana, A. Bishop, A. Savkin, S. W. Ekanayake, and T. Black, "A method for stereo-vision-based tracking for robotic applications," *Robotica*, pp. 1–8, 2009.

[51] W. Piekarski, "Interactive 3d modelling in outdoor augmented reality worlds," Ph.D. dissertation, School of Computer and Information Science, Division of Information Technology, Engineering, and the Environment, The University of South Australia, 2004.

[52] M. Pupilli, "Particle filtering for real-time camera localisation," Ph.D. dissertation, Department of Computer Science, University of Bristol, 2006.

[53] G. Klein, "Visual tracking for augmented reality," Ph.D. dissertation, Department of Engineering, University of Cambridge, 2006.

[54] R. Sim, P. Elinas, and M. Griffin, "Vision-based slam using the rao-blackwellised particle filter," *Workshop on Reasoning with Uncertainty in Robotics*, 2005