2010 The 3rd International Conference on Machine Vision (ICMV 2010)

Tracking Methods for Augmented Reality

Erkan Bostanci, Nadia Kanwal, Shoaib Ehsan, Adrian F. Clark
School of Computer Science and Electronic Engineering
University of Essex
Colchester, United Kingdom
{gebost, nkanwa, sehsan, alien}@essex.ac.uk

Abstract—**Augmented Reality has been an active area of research for over two decades. This paper presents a comprehensive review of the recent literature on tracking methods used in augmented reality applications for both indoor and outdoor environments. After reviewing tracking methods, the paper identifies limitations of the state-of-the-art techniques and suggests potential future directions.**
Keywords- augmented reality; tracking; SLAM

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 a user's environment with synthetic content [1]. For over two decades, AR has been of interest in 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 ability to run on platforms such as mobile computers and cell phones [4].

Tracking, the process of locating a user's position and orientation in an environment, is critical to AR as more realistic results can be obtained when the registration of real and synthetic is accurate. Generally, the user wears a head-mounted display (HMD) upon which the augmented images of the real world are displayed, so tracking it is paramount. Improved accuracy of tracking also prevents problems such as visual capture [2] and prevents visual sensors from gaining priority over other sensors. For instance, inadequate registration accuracy can cause augmentation to be displayed incorrectly and the user to reach the wrong part of the real environment. The user becomes used to errors in the virtual environment and eventually accept them as correct.

This paper reviews state-of-the-art tracking methods used for AR, identifies the bottlenecks involved and proposes future research directions. It is structured as follows: Sec. II discusses tracking techniques for indoor and outdoor environments, fusion methods and a recent set of methods which were well known in the robotics community but are new to computer graphics. Limitations of current methods are identified in Sec. III. Future research directions are proposed in Sec. IV, followed by conclusions in Sec. V.

II. TRACKING METHODS

A variety of tracking methods are found in the literature [5]. This section provides a review of tracking for AR under four main categories: indoor methods; outdoor methods; fusion strategies; and recent approaches (Fig 1).

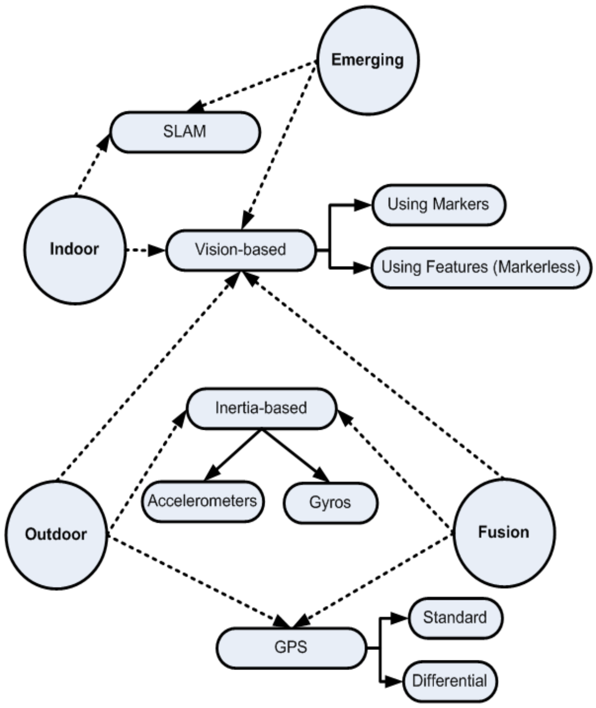

Figure 1. Tracking Methods for Augmented Reality

A. Indoor Techniques

Indoor environments provide a structured domain for an AR application, with the user's movements limited to a specific region, making them more predictable [6]. 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* (or *fiducial*, *beacon* or *landmark*) in the context of these methods. Markers are distinguishable elements put into the environment so that they can be distinguished from other objects present. Markers can be categorized as *active* or *passive:* active markers emit a signal (*e.g.,* magnetic, light) which can be sensed, while passive markers tend to be a pattern which can be easily identified in the environment (e.g. QR codes). In this case, computer vision methods are appropriate for recognizing the marker.

ISBN: 978-1-4244-8889-6 © ICMV 2011

Indoors tracking is generally achieved by *outside-in* or *inside-out* methods [8]. These names indicate the location of the sensor, which can be magnetic, ultrasonic, radio frequency identification (RFID) or visual, and how the tracking is achieved. For outside-in, the sensor is fixed in the environment and user wears headgear on which the markers are mounted. For inside-out, the user carries the sensor, and the markers are fixed at known locations in the environment.

Although indoor tracking can use magnetic or ultrasound sensors, such systems generally use expensive, complex hardware [9,10]. Although GPS is a good option for tracking outdoors, indoor environments generally attenuate these signals too much, so vision-based tracking systems prevail.

A wide-range tracking system named HiBall was presented in [8]. The aim of that work was an accurate, robust, flexible system for use in large indoor environments. The HiBall device was designed as a hollow dodecahedron with its upper part fitted with 6 lenses. LEDs, controlled by an interface board, were mounted on the ceiling.

A complete fiducial-based tracking system named Video Positioning System (VPS) was designed and developed in [11]. An innovative fiducial design was introduced, with unique patterns for accurate position and orientation calculation. The design employed region adjacency graphs, with a 'key' to indicate that a region was a marker and a separate identifier to differentiate markers. VPS was also applied on a parallel architecture in [12], and it was shown that this produced real-time performance for parts of it. In [13], a comparison of VPS and ARToolkit showed that VPS provided more accuracy than ARToolkit for moderate changes in viewpoint and distance. However, ARToolkit performed better when distance is increased and the authors suspected that this was due to the fiducial design.

Chia *et al.* [14] developed a camera tracking system based on natural features. It used pre-captured reference images of the scene, RANSAC being used for robust matching to achieve motion invariance of the feature points. The system was able to run at 10Hz using some markers too.

Park *et al.* tracked several 3D objects simultaneously, robustly and accurately in real time [15]. Frame-to-frame tracking was found to be computationally undemanding but prone to failure while detection was more robust but slower. For each target object, a 3D CAD model and a small set of reference images ('keyframes') were available. Keypoint detection and pose estimation were performed in one thread, keyframe detection in another. The system was able to work at 15-20fps on a 3.2 GHz multicore CPU, though performance deteriorated as the number of objects increased.

Bekel [16] presented a viewpoint-based approach for AR using a Self-Organizing Map (SOM) trained as a classifier which was then used to label different objects in the scene.

Adams *et al.* [17] developed a method for aligning viewfinder frames obtained from a mobile phone camera and they applied it to 3 different cases: 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 diagonal directions and a set of point features. Second was the alignment of edges. These gave the translation between two frames. Using feature correspondences, the initial translation was obtained. The alignment algorithm was fast and robust against noise but fragile against rotations >1.5^0.

A different approach for pose tracking with a built-in camera of a mobile phone was followed by Wagner *et al.* [18]. They used SIFT for robust features and Ferns, a fast classification method. To reduce the computation in SIFT, the calculations for scale invariance were replaced with a database of the features at different scales. FAST was used for corner detection for its high repeatability.

VisiTrack was developed in [19] for tracking with mobile devices using point and edge extraction together with colour segmentation. Although the system was claimed to provide marker-less 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.

B. Outdoor Techniques

Outdoor environments are generally less predictable than indoor ones as there is less chance for instrumenting it. Pre-defined artificial landmarks cannot be used and natural landmarks must be found. Varying light poses a problem for camera tracking which is not an issue for indoors. As mentioned earlier, GPS is a good tracking option for outdoor AR. A comparison of GPS receivers, including brands such as Trimble, Garmin and DeLorme, is given in [6].

Differential GPS and a compass were 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.

In [22], reference images were used for video tracking, with matching performed to find these reference images for an outdoor AR system. A video image was compared with the reference images and a matching score obtained. For the best matching image, the 2D transformation was calculated and the camera position and orientation deduced. This was used to register the model on the video frame. The matching technique employed Fourier transformation, to be robust against changes in lighting conditions, limiting it to 2D transformations such as rotation and translation. This had a fixed number of computations, making it suitable for real-time operation at10Hz, a low rate for real-time display.

Inertial sensing is widely used since its operation is similar to the otolith stones in the human ear [23]. Accelerometers are used for translation and gyros for rotation. This method is generally used in combination with other tracking methods; see *C* below.

For tracking in the ancient city of Paestum in Italy, the use of WiFi instead of GPS was planned [24]; however, this was not implemented due to the opposition to changes in the original archaeological site by archaeologists [22].

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

C. Fusion Strategies

When the above-mentioned methods are used alone, tracking accuracy may be low. More accurate systems can be obtained by using combining sensor signals. Fusion methods may be loosely or tightly coupled [23]. In loosely coupled systems, the sensors act separately and perform calculations independently; however, in tightly coupled systems, sensor fusion is used to generate improved position estimation.

Visual inertial tracking is popular as the sensors have complementary characteristics. Vision allows estimation of the camera position directly from the images observed [26] but is not robust against 3D transformations, and the computation is expensive. For inertial trackers, noise and calibration errors tend to accumulate so they have long-term stability problems [27]. When used together, faster computation can be achieved with inertial sensors, while inertial drift errors can be corrected with vision. Applications generally use low frequency vision data and high frequency inertial data [28] since visual processing is more expensive. Estimates at 550Hz are possible on custom hardware [29].

A hybrid tracking system was developed for mobile outdoor AR in [3], combining vision-based and inertial trackers. The hardware included 3 accelerometers, 3 gyroscopes and a digital signal processor, yielding 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], using several cameras. A fusion filter combined visual and inertial measurements, employing markers arranged in a circular matrix.

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

You *et al.* [26] developed a hybrid system for accurate AR registration using a prediction-correction method. Data obtained from the inertial sensor was used to estimate 2D feature motion (prediction) and then visual feature tracking corrected the estimate (correction). Finally, 3D correction was performed using gyros from the 2D motion residual.

Tracking in [6] used GPS and head tracker, a camera being used only for view. System components included a Trimble AgGPS 332 Receiver, TCM5 3-axis orientation tracker, WristPC wearable keyboard2, Cirque smart cat touch pad3, i-glasses HMD and laptop.

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

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

Piekarski [33] developed an outdoor AR system using a Trimble Ag132 GPS unit and an orientation tracker, achieving an accuracy below 50cm. 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. A different outdoor application which aimed to display an archaeological site was given in [34]. The system used GPS and inertial sensors within the HMD.

Sherstyuk [35] developed a novel method for fast semi-automatic 3D geometry acquisition using motion tracking equipment, intended for quick surface prototyping where quality is not of high priority. A life-size medical mannequin 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.

D. Emerging Approaches

Tracking position and orientation is an issue in both 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 can be applied for a robot, it can also be applied to the AR user.

An interesting, interactive application was presented in [36] where EKF SLAM [37] was applied to an AR game in which a ninja tries to jump from one plane to another until a target plane is reached. Higher-level structures such as planes were created from point feature sets using RANSAC, while the OGRE engine was used to implement the game.

Bleser [38] investigated the robustness and accuracy of real-time, marker-less AR in both known and unknown environments. Sensor fusion was performed on IMU and camera data with particle filtering to match a CAD model of the environment or an object with the actual object. The tests showed operation in a space of 2.1 x 1.5 x 2.5m and a conceptual solution for large environments was presented.

One of the most impressive marker-less tracking for AR using SLAM was presented in [39], using separate threads for tracking and mapping the environment.

Kozlov *et al.* [40] proposed using AR for visualizing the internal state of a robot in order to test and debug SLAM algorithms. The authors proposed visualizing robot pose, state map and data association; cross-correlations could be used to show decreases of uncertainty in the map.

III. PROBLEMS WITH CURRENT METHODS

Current tracking systems still have many problems. Two types of errors were defined in [2], namely *static* and *dynamic* errors. Static errors are due to the inadequate accuracy provided by the actual sensors, while dynamic errors are delays. The end-to-end delay is the time elapsed between the time when the tracking system measured 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 at a low cost. The problem with these methods is the lack of robustness [41]. For some applications, *e.g.* [31], the probability of incorrect matches is large as the texture of the ground is similar in different areas, 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. For visual tracking, the features used as landmarks should be invariant to changes in lighting and viewpoint. As this is not always possible, vision-based tracking outdoors is reported to be fragile [3]. Using a camera as the only sensor was found to be accurate but expensive in computation [43].

Standard GPS has an error in order of 10m. This is improved when differential GPS or real-time kinematic (RTK) correction is used. The lack of a line of sight to satellites is a problem in urban environments [23] and under dense tree canopy [44]. Other problems with GPS were explained in detail in [45]. The system developed in [46] reported tracking problems when GPS was used as the only sensor; its authors suggested using sensor fusion or GPS dead-reckoning. Double integration of inertial tracker data causes errors to accumulate rapidly [31], reducing long-term stability [48]. Active tracking systems require calibrated sensors and signal sources in a prepared environment, while tracking equipment can be affected by signal noise, degradation with distance and by interference [41]. Magnetic trackers interfere with nearby ferro-magnetic objects [47].

Other problems also arise. Data association (finding a correspondence between the feature model and the observed feature) problems occur due to low precision and recall rates [49]. Linearization, due to the characteristics of current SLAM methods, affect filter stability and convergence, resulting in less accurate localization [50].

IV. FUTURE DIRECTIONS

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

AR has the potential for many different and interesting applications, including entertainment such as games [1,51] or cultural heritage applications as in [22,34]. Most outdoor methods use GPS and inertial trackers [51,34,46,6]. Vision-based tracking has also been employed [41,22,21]. Inertial sensors can be used for stability and robustness in the presence of rapid motion or occlusion [43]. Current applications of SLAM for AR [52,53,43,36] are limited to desktop or laboratory environments, though with accurate tracking. In [52], localization was performed according to known 3D junctions, AR tests being carried out with a rectangular pattern in view at all times. Similar results can be seen in[43].

Considering the methods used today, several ideas are worth following up in future research. Firstly, sensor fusion within a SLAM framework is promising. Secondly, vision-based tracking is useful because images of the environment are needed for augmentation. As this information source is at hand, it is wise to use it for both tracking and augmentation. Thirdly, the introduction of robust detectors such as SIFT or SURF will improve the visual tracking process, when real-time implementations are achieved. [54]. Finally, for performance considerations, GPU-based or parallel implementations are suggested.

V. CONCLUSION

This paper has presented several examples of tracking methods for AR, for both indoor and outdoor environments. They were examined critically, considering their advantages and disadvantages for stand-alone use or in combination. Emphasis was placed on visual tracking methods due to their increasing popularity. With new methods being developed by the computer vision community and the fusion of vision-based methods with other sensors, the authors believe that the accuracy of tracking will soon be good enough 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, “Generalpurpose 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,” Int. Symp. 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,” in Third International Symposium on Wearable Computers, 1999, pp. 29–36.

[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, p. 445458, 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," Int Conf on Patt Rec, vol. 4, 2006.

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

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

[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," Int. Symp. on Mixed and Augmented Reality, 2008, pp. 125–134.

[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," in 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," in 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," in Eurographics Italian Chapter Conference, 2006.

[25] Y. Yamanaka, M. Kanbara, and N. Yokoya, "Localization of walking or running user with wearable 3d position sensor," in 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," in 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," in IEEE Instruments and Measurement Technology Conference, 2002.

[30] E. Foxlin and L. Naimark, "Vis-tracker: A wearable visioninertial self-tracker," in 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, Department of Electronic systems Engineering, University of Essex, 2006.

[35] A. Sherstyuk, A. Treskunov, and B. Berg, "Fast geometry acquisition for mixed reality applicaitons using motion tracking," in 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," in 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," in 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, l. G. Schal, D. Wagner, I. Barakonyi, G. Reitmayr, J. Newman, and F. Ledermann, "Managing complex augmented reality models," IEEE Comp 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," in 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," Symp. on VR 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 vision based slam," Int. 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 CompSci, 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," in Workshop on Reasoning with Uncertainty in Robotics, 2005.