Augmented Reality: Tracking Methods

Archita Dad
Computer Department,
ACE, Mumbai

Bhavna Arora
Computer Department,
ACE, Mumbai

Nida Parkar
Computer Department,
ACE, Mumbai

Tejal Rachh
Computer Department,
ACE, Mumbai

***Abstract*— Augmented reality has been an interesting area of research for the last two decades or so. This paper presents a brief overview of the recent literature on tracking methods used in Augmented Reality applications, both for indoor and outdoor environments.**

Key Words* — *Augmented reality , tracking, Indoor, Outdoor.

I.INTRODUCTION

Augmented reality (AR) is direct or indirect live, view of real-world environment which elements are augmented by computer-generated sensory input such as images, video, sound, graphics or GPS data. It is also defined as the extension of user's environment with synthetic content [1]., AR has been a topic of interest in the field of computer graphics or virtual reality more than two decades as it improves human perception and facilitates understanding of complex 3D scenarios [2], [3]. AR applications and games like Pokémon Go are now becoming more popular due to their interactive user friendly environment and capability to run on a variety of platforms such as mobile computers and even cell phones [4].

Accuracy of AR applications or realistic results is depending on tracking method which is the process of locating a user in an environment. It takes account of the position and orientation of the AR user. Generally, tracking the head is the most important part as the user wears a Head Mounted Display (HMD) by which the augmented images of the real world are displayed. Due to tracking accuracy of the AR system is improved and prevents problems such as visual capture [2] and does not allow visual sensors to gain a priority over other sensors. For example, insufficient registration accuracy can cause the user to reach wrong portion of the real environment because the augmentation has been displayed on another portion. The eyes of the users get used to the error in the virtual environment and after some time they start to accept these errors as correct which is not desirable

This paper reviews the tracking methods used for AR, identifies the difficulty involved and proposes future research work. The paper is structured as follows: Section 2 discusses the tracking methods used for indoor and outdoor environments and a recent set of methods. The restrictions of the currently used methods are described in Section 3. Future research area that can be explored are proposed in Section 4 which is followed by conclusions in Section 5.

II.TRACKING METHODS

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

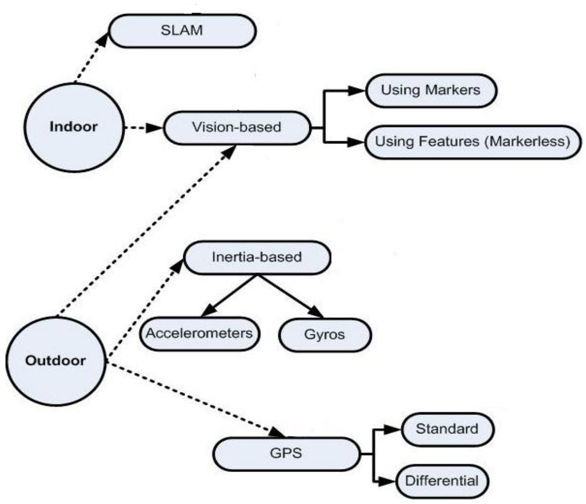

Fig. 1. AR tracking methods.

III. INDOOR TECHNIQUES

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

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

Although there are many different types of indoor tracking methods with magnetic or ultrasound sensors, these systems generally use both costly and difficult hardware [9,10]. Although GPS is a good decision for tracking user position outdoors, indoor environments like laboratories or buildings generally obstruct these signals. The uncertainness of GPS signal in the indoor environments calls for more depend on vision-based tracking systems.

A complete tracking system named Video Positioning System (VPS) was proposed and evolved in [11]. This system used fiducial markers. Markers are discernible elements put in the environment so that they can be recognized apart from other objects in the same environment. These markers can be classified as active or passive markers. Active markers produce a signal (e.g. magnetic, light) which can be sensed via the sensor. Passive markers tend to be a pattern which can be easily separated from the texture of the environment (e.g. QR codes). A new fiducial pattern design was introduced which permits using distinctive patterns for accurate position and orientation calculation. The pattern blueprint was based on RAG (region adjacency graph) with two elements namely key and the identifier one of which tell that this shape is a marker and the second distinguished between different markers in the system.

Chia et al. [12] developed a camera tracking system depend on natural features. The system used pre-captured reference images of the scene and then RANSAC was used for robust matching to achieve invariance in movements of the feature points. The system was able to operate at 10Hz using some fiducial markers as well.

Bekel [13] proposed a viewpoint based approach for AR. This method used Self-Organizing Map (SOM) to instruct as a classifier which is later used to label different types of objects by overlaying in the scene.

A different approach for pose tracking is a built-in camera of a mobile phone was followed by Wagner et al. in [14]. They have used SIFT for robust characteristics and Ferns for classification. Ferns is a fast classification method however it needs a great amount of memory.

VisiTrack system was implemented in [15] for tracking in mobile devices by point and edge extraction jointly with colour segmentation. Although the system was stated to provide markerless localization a marker can be viewed in the test sequences in the system running at 25fps.

For the indoor AR system in [16], visual tracking was utilized. The system recognized image views of the environment grabbed beforehand. Processing was carried out by remote PCs using a wireless LAN.

IV.OUTDOOR TECHNIQUES

Indoor environments are usually more predictable whereas the outdoor environments are limitless in terms of location and orientation. As contrasting to indoors, there is less chance for preparing the environment to track the user while working outdoors. Moreover, already specified artificial landmarks cannot be used and natural landmarks need to be found.

As stated earlier, GPS is a good tracking option when working outdoors.

A differential GPS and a compass was used for position and orientation judgement in [17]. Latitudes and longitudes of several viewpoints were collected in a database along with the set of images captured at different times of the year with varying light conditions.

Reference images were utilized for video tracking and matching was performed to discover these reference images for the outdoor AR system in [18]. A video image was examined with the reference images and a matching score was achieved. For the finest matching score, the 2D transformation was measured and the current camera position and orientation were deducted. This transformation was utilized to register the model on the video frame. The matching technique was based on Fourier Transformation to be robust against variation in lighting conditions hence it was limited to only 2D transformations like rotation and translation. This technique had a fixed number of computations therefore it was appropriate for real-time operation without using markers, it yet worked on 10Hz which is a low rate for real-time display.

Inertial sensing is a widely used method because its operation is same to the otolith stones in human ear [19]. Accelerometers are utilized for translational motion and gyros for rotational motion. This method is usually used together with other tracking methods as it will be given in more specific in Fusion Strategies.

For the tracking of the user's position in [20] in the ancient city of Paestum in Italy, a WiFi system was designed by installing antennas alternatively of using GPS. However, this was not implemented because of the opposition to changes in the original archaeological site via the archaeologists [18].

Tracking for a walking or running user was obtained using a custom tracking system in [21]. The system utilized an inertial sensor, electromagnetic sensor, push-buttons in the heels of the user and trackers in the knees in order to obtain the actual motion of the legs. The transmitter was mounted above user's waist so that the relative motion of the legs could be calculated when the user's foot does not ground.

V. PROBLEMS WITH URRENT APPROACHES

There are still many problems in current tracking system. Two types of errors were mentioned in [2] namely static and dynamic errors. Before going into details about problems, these two important terms will be clarify.

The errors in tracking systems are known as static errors due to the insufficient accuracy provided by the commercial systems at the time being. Delays are known as dynamic errors. The end-to-end system delay is the time occupied in between the time when the tracking system to compute the position and orientation of the user to the time when the images show on the display.

Vision methods permits both tracking and managing residual errors. They are not expensive. The problem with these methods is the lack of hardiness [24]. For some applications e.g. [23], there may be a huge probability of incorrect matches since the texture of the ground is mostly similar in different areas as a repeating pattern. It was expressed in [25] 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.

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 [19] or under dense tree canopy due to variance [26]. Other problems with GPS were explained in detail in [27]. The system developed in [28] reported about the tracking problems occurring when as a sensor GPS was used and the authors advised for using sensor fusion or GPS dead-reckoning.

VI. FUTURE DIRECTIONS

AR has the potential for lots of different and interesting applications including entertainment like games in [1, 29] or cultural heritage applications in [18, 23] when talk about outdoors are considered.Most outdoor methods utilized GPS and inertial trackers for this purpose [6, 23, 28, 29]. Vision based tracking has also been used for some systems as in [17, 18, 24]. Inertial sensor can be utilized for stability and robustness in cases of rapid motion or occlusion.

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.

VII. 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.

VIII. REFERENCES

[1] W. Broll, I. Lindt, I. Herbst, J. Ohlenburg, A. Braun, and R. Wetzel,"Toward next-gen mobile ar games," *IEEE Computer Graphics andApplications*, pp. 40–48, 2008.
[2] R. Azuma, "A survey of augmented reality," *Presence: Teleoperatorsand 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," *IEEEComputer 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] Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACMComputing Surveys*, vol. 38, no. 4, 2006.
[6] Behzadan, B. W. Timm, and V. R. Kamat, "General purpose modular hardware and software framework for mobile outdoor augmented reality applications in engineering," *Advanced EngineeringInformatics*, 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 MixedReality*, 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 WearableComputers*, 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] K. Chia, A. Cheok, and S. Prince, "Online 6 dof augmented reality registration from natural features," *International Symposium on Mixedand Augmented Reality*, pp. 305–313, 2002.
[13] 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.
[14] 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.
[15] D. Stichling, N. Esau, B. Kleinjohann, and L. Kleinjohann, "Real-time camera tracking for mobile devices: The visitrack system," *Real-TimeSystems*, pp. 279–305, 2006.
[16] J. Kim and H. Jun, "Vision-based location positioning using augmented reality for indoor navigation," *IEEE Transactions onConsumer Electronics*, vol. 54, pp. 954–962, 2008.
[17] P. Dahne and J. Karigiannis, "Archeoguide: System architecture of a mobile outdoor augmented reality system," *IEEE InternationalSymposium on Mixed and Augmented Reality*, 2002.
[18] D. Stricker and T. Kettenbach, "Real-time and markerless vision-based tracking for outdoor augmented reality applications," *InternationalSymposium on Augmented Reality*, 2001.
[19] 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.
[20] R. Chiara, V. Santo, U. Erra, and V. Scanarano, "Real positioning in virtual environments using game engines," *Eurographics ItalianChapter Conference*, 2006.
[21] 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.

[22] S. Diverdi and T. Hollerer, “Heads up and camera down: A vision-based tracking modality for mobile mixed reality,” *IEEETransactions on Visualization and Computer Graphics*, vol. 14, pp.500–512, 2008.
[23] P. D. Ritsos, “Architectures for untethered augmented reality using wearable computers,” Ph.D. dissertation, Dept. Elect systems Engineering, Essex Univ., 2006.
[24] U. Neumann and S. You, “Natural feature tracking for augmented reality,” *IEEE Transactions on Multimedia*, vol. 1, no. 1, pp. 53–64, 1999.
[25] 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.
[26] 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.
[27] T. Bailey, “Mobile robot localisation and mapping in extensive outdoor environments,” Ph.D. dissertation, Australian Centre for FieldRobotics, Department of Aerospace, Mechanical and Mechatronic Engineering The University of Sydney, 2002.
[28] 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.
[29] W. Piekarski, “Interactive 3d modelling in outdoor augmented reality worlds,” Ph.D. dissertation, School of Computer and InformationScience, Division of Information Technology, Engineering, and the Environment, The University of South Australia, 2004.