

A Survey on Marker-less Augmented Reality

Sandip Kote^{#1}, Bharat Borkar^{*2}

Department of Information Technology, Amrutvahini college of Engineering,
Sangamner,(M.S) India-431005

Abstract— Augmented reality (AR) is a live and combines direct or indirect view of a physical, real-world environment it's realtime data whose elements are augmented by computer-generated virtual content such as sound, video, graphics or GPS data. An AR system poses unique challenges including requiring a high quality camera pose estimate and operating on resource-limited platforms. There is several hybrid approach using ORB binary features and optic flow that is able to real time performance result is possible with platform specific optimizations, improve speed and extend the usable tracking range.

Keywords— Augmented Reality; Markerless; Real Time; Feature Detectors; ORB; Hybrid Tracking.

I. INTRODUCTION

An augmented reality system generates a composite view for the user. It is a combination of the real scene viewed by the user and a virtual scene generated by the computer that augments the scene with additional information. In all those applications the augmented reality presented to the user enhances that person's performance in and perception of the world. The ultimate goal is to create a system such that the user can not tell the difference between the real world and the virtual augmentation of it.

In order to perform the object tracking, markerless augmented reality systems rely in natural features instead of fiducial marks. Therefore, there are no ambient intrusive markers which are not really part of the environment. Furthermore, markerless augmented reality counts on specialized and robust trackers already available. Another advantage of the markerless systems is the possibility of extracting from the environment characteristics and information that may later be used by them. However, among the disadvantages we can consider for markerless augmented reality systems is that tracking and registration techniques become more complex.

old AR systems like ARTag [1] utilize fiducial markers or “tags”, over which virtual objects are rendered. Marker-based systems are ideal for applications involving a static or fixed environment, or situations where the desired virtual defined marker tag does not extend beyond the tag itself. In the performing and using of application, users must print and carry the tags in order to use the AR application. For more complex, large scale applications, significant infrastructure is required in the form of huge no of or hundreds of markers placed throughout the environment [2].

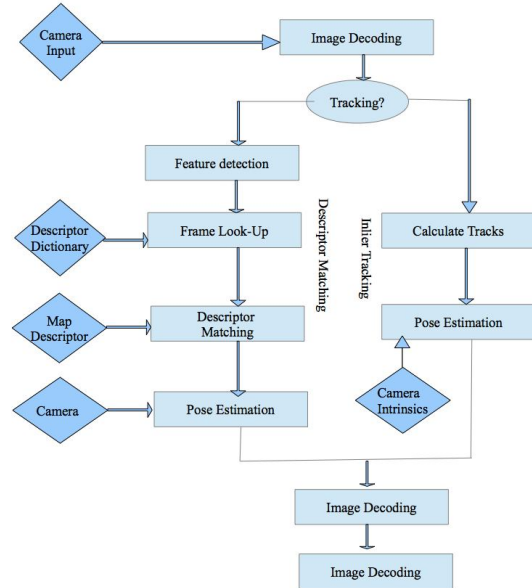


Figure 1 -AR Pipeline overview

II. RELATED WORKS

A. MONOSLAM: REAL-TIME SINGLE CAMERA SLAM:

This MonoSLAM [3] real-time algorithm which can recover the 3D trajectory of a monocular camera, moving rapidly through a previously unknown scene. study system, which we dub MonoSLAM, is the first successful application of the SLAM methodology from mobile robotics to the “pure vision” domain of a single uncontrolled camera, achieving real time but drift-free performance inaccessible to Structure from Motion approaches. The core of the approach is the online creation of a sparse but persistent map of natural landmarks within a probabilistic framework. these add up to an extremely efficient and robust algorithm which runs at 30 Hz with standard PC and camera hardware. This work extends the range of robotic systems in which SLAM can be usefully applied, but also opens up new areas. We present applications of MonoSLAM to real-time 3D localization and mapping for a high-performance full-size humanoid robot and live augmented reality with a hand-held camera. Efficiency is provided by active feature search, ensuring that no image processing effort is wasted—this is truly a Bayesian, “top-down” approach. We have presented experimental implementations which demonstrate the wide applicability of

the algorithm, and hope that it will have an impact in application areas including both low-cost and advanced robotics, wearable computing, augmented reality for industry and entertainment and user interfaces.

B. Parallel Tracking and Mapping on a Camera Phone :

Smart camera phones are a promising new way for hand-held augmented reality. As their computational resources grow, they are becoming large amount of suitable for visual tracking tasks. At the same time, they still offer considerable challenges: Their cameras offer a narrow field-of-view not best suitable for robust tracking; images are often collected at less than 15Hz; long exposure times result in significant motion blur; and finally, a rolling shutter causes severe smearing effects. This work describes an attempt to implement a keyframe-based SLAM system on a camera phone (specifically, the Apple iPhone 3G).

a series of adaptations to the Parallel Tracking and Mapping system to mitigate the impact of the device's imaging deficiencies. Early results demonstrate a system capable of generating and augmenting small maps, albeit with reduced accuracy and robustness compared to SLAM on a PC.[4] Finally, this work has discussed the iPhone's limitations vis-a-vis the PC, but has not attempted to exploit its extra capabilities. The iPhone's camera is high performance and can take high-resolution stills, the phone contains an accelerometer and a GPS receiver, and future devices will contain magnetic compasses and gyroscopes. These extra capabilities could surely be used to benefit visual tracking and AR.

C. Hybrid tracking method :

Taehee and Hollerer also propose a multi-threaded approach in [5]. They use a hybrid tracking method which extracts SIFT features [6] (instead of image patches) and then tracks them using optic flow. They achieved real-time performance by only extracting and matching SIFT features periodically in a thread separate from the tracker. They also perform scene recognition by recognizing previously recorded SIFT features.

On the mobile front, attempts have been made to adapt SIFT and its speedier counterpart SURF [8] to mobile devices. Wagner et al proposed a hybrid between FAST corners [9] and a reduced version of the SIFT descriptor in [10]. Rather than compute descriptors every frame, they also adopt a hybrid approach and track existing features using SAD patch correlation. This method achieves upwards of 20Hz while extracting ~150 features per frame (320x240). Chen et al propose a modified version of SURF in [11] that achieves a roughly 30% speed-up over the original SURF algorithm, but nevertheless falls far short of real-time operation on mobile devices. The first self-contained AR system to run on a consumer-grade cell phone was presented in 2004 by Mehring et al [12]. Their system recognizes different markers via circular bar codes and detecting gradient changes in the red, green, and blue color channels. Their entire pipeline

achieved a frame rate of 4-5fps, at a camera resolution of 160x120.

D. Parallel Tracking and Mapping :

A reduced version of the PTAM system has been adapted to the iPhone [13], but results showed severely reduced accuracy and execution speed. PTAM is intended for use in small AR workspaces, and suffers reduced performance as the map gets bigger and bigger due to the bundle adjustment process being cubic with respect to the number of features in the map ($O(n^3)$). As a continuation of their work in [10], Wagner et al propose a more complete version of their feature detection and matching system in [14]. They use a similarly modified version of SIFT and outlier rejection that runs at approximately 26Hz. However, when AR-related overhead (image retrieval from camera, rendering, etc.) is taken into consideration, the speed drops to 15Hz. They combine this with patch tracking to greatly improve speed to 8ms per frame (not including AR overhead) instead of using SIFT on every frame. Their system runs on 320x240 imagery during the SIFT phase, and down

III. APPLICATIONS :

- Facilitating Education And Training
- Augmented Reality Technology to the Service of Industrial Maintenance
- Augmented Reality games, a new way to uplift your product communication.
- Medical services.

IV. CONCLUSION :

This paper presents an extensive survey on markerless augmented reality system designed to operate in real time on current generation mobile devices. Also work on mobile devices to operate at speeds of approximately 30Hz while tracking, and 15Hz while extracting and matching features. Many research issues have been highlighted and direction for future work have been suggested. Many open issues have been highlighted by the researchers such as dealing with optimisations, speed with dynamic scene.

V. ACKNOWLEDGMENT

I would like to express my special appreciation and thanks to my advisor Professor Mr. Bharat Borkar, you have been a tremendous mentor for me.

REFERENCES

- [1] M. Fiala, "ARTag, a fiducial marker system using digital techniques," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, pp. 590- 596 vol. 2.
- [2] M. Fiala and G. Roth, "Magic Lens Augmented Reality: Table-top and Augmentorium," presented at the ACM SIGGRAPH 2007 posters, San Diego, California, 2007.
- [3] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-time single camera SLAM," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, pp. 1052- 1067, 2007.
- [4] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," in *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, 2007, pp. 225-234.
- [5] L. Taehee and T. Hollerer, "Hybrid Feature Tracking and User Interaction for Markerless Augmented Reality," in *Virtual Reality Conference, 2008. VR '08. IEEE*, 2008, pp. 145-152.
- [6] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, 1999, pp. 1150-1157 vol.2.
- [7] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," *Computer Vision–ECCV 2006*, pp. 404-417, 2006.
- [8] E. Rosten and T. Drummond, "Machine Learning for High-Speed Corner Detection," presented at the European Conference on Computer Vision, 2006.
- [9] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg, "Pose tracking from natural features on mobile phones," presented at the Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 2008.
- [10] W.-C. Chen, Y. Xiong, J. Gao, N. Gelfand, and R. Grzeszczuk, "Efficient Extraction of Robust Image Features on Mobile Devices," presented at the Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007.
- [11] G. Klein and D. Murray, "Parallel Tracking and Mapping on a camera phone," in *Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on*, 2009, pp. 83-86.