AUGMENTED REALITY APPLICATIONS USING VISUAL TRACKING

Chutisant Kerdvibulvech

Department of Information and Communication Technology, Rangsit University, Thailand

Email: chutisant.k@rsu.ac.th

ABSTRACT

In this 21st century, augmented reality has recently become a quickly developing field of computer science and engineering in which a virtual world is overlaid on top of the real world. This paper provides an up-to-date review of major visual tracking research. We surveyed recent augmented reality applications using visual tracking. A literature review of the most recent visual tracking techniques is presented. We divide methods of tracking into two main categories. The first group is color-based tracking. The second group is feature-based tracking. Description and limitations of each technique are given. Ultimately, we give a summary of the visual tracking.

Index Terms - Augmented Reality; Applications; ARTag; Image Processing; SLAM; Feature; Color

1. INTRODUCTION

In several past decades, virtual reality (VR) was a popular field of computer science and information technology. A conventional virtual reality system seeks to completely immerse the user in a computer generated environment. As the user is looking at a virtual world, sometimes it is not a natural connection. In recent years, augmented reality has rapidly become a very popular field of computer science and engineering in which a virtual world is overlaid on top of the physical world. Augmented reality can help this overlaying issue by generating imagery in live-video streams as a way to expand the physical world by applying image processing technology. In other words, augmented reality (AR) is a growing area in virtual reality work.

Augmented reality applications have been recently presented through research laboratories in many prestigious universities around the world. The task of visual tracking in augmented reality has been actively researched in recent years. In this paper, we review visual tracking for augmented reality applications. These include SixthSense from Massachusetts Institute of Technology (MIT) and Parallel Tracking and Mapping (PTAM) from University of Oxford, MirageTable from Microsoft Research, Facial Feature Tracking from ETH Zurich, and some recent researches in this area.

Basically, we divide methods of visual tracking into two main categories. The first category is color-based tracking. The second category is

 [Published online: 30 September 2013]

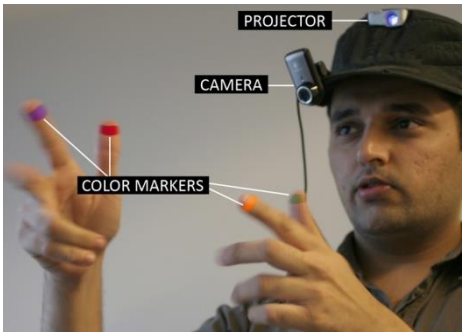

Figure 1. Augmented reality application called SixthSense from MIT Media Laboratory [1].

feature-based tracking. In this section, we discuss each category in detail.

2. AR SYSTEMS WITH COLOR-BASED TRACKING

One of the leading augmented reality researches using color-based tracking is SixthSense [1]. This system was initiatively proposed by Mistry and his group from MIT Media Laboratory. As shown in Figure 1, SixthSense is a gestural system that allows natural hand gestures interaction with digital information. The system projects the digital information onto the surfaces and real objects around us, making any surface into a digital interface. The camera recognizes the user's colored fingertip markers and physical objects using the vision-based tracking algorithm [2]. However, using colored fingertip markers sometimes seems to be unnatural.

Next application which was implemented by Kerdvibulvech [3] is an augmented reality game for tracking a car that is remotely and interactively controlled by players in real time. This technique requires no marker. The interactive application can entertain players by controlling the car to the

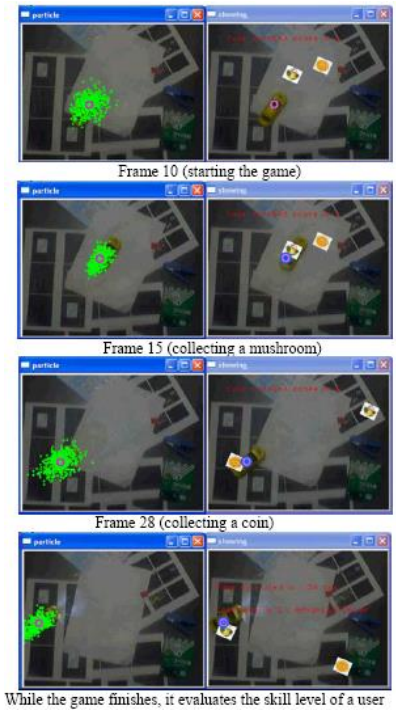

Figure 2. Markerless vision-based tracking for interactive augmented reality game [3].

augmented items. In this research, a framework for colored remote-controlled car tracking is utilized based on a Bayesian classifier and particle filters. Figure 2 shows the experimental results of the application. The yellow car is tracked by the system, while the augmented items are shown to entertain players.

Next, adaptive color-based tracking was presented in [4] by using the on-line adaptation of color probabilities. In this research, skin color of Asiatic people is tracked by employing Bayes' theorem and Bayesian classifier as displayed in Figure 3. This method has two main features. It can cope with luminance changes, and it can also process in real time. However, this research focuses only on skin color of Asiatic people.

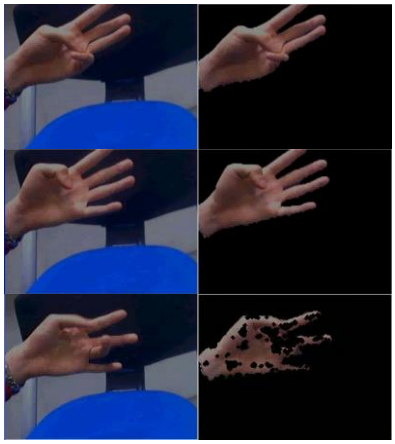

Figure 3. Asiatic skin segmentation based on color tracking adaptively [4].

More recently, Microsoft Research's MirageTable was proposed in May 2012 by Benko et al [5] at the CHI conference. This interactive system was designed to merge physical and virtual worlds into a single spatially registered experience on top of a table. The instrument is composed of a single depth camera, a stereoscopic projector, and a curve screen. Color and Depth images are calculated in each frame. The user's eyes are occluded by the shutter glasses. Therefore, rather than track the eyes, they decided to track the location of the glasses in the depth image instead of tracking the eyes. They used that information to compute the user's viewpoint. This camera tracks the user's eyes and performs a real time capturing of both shape and appearance of any object placed in front of the camera. The system enables perspective stereoscopic 3D visualization to a single user. Figure 4 shows the experimental results. This allows the user to interact with virtual objects through freehand actions without handling. The system shows the potential of using

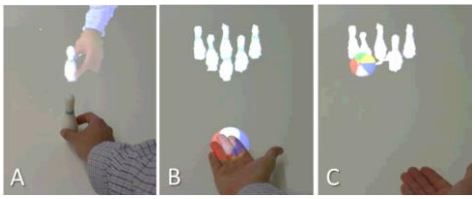

Figure 4. Bowling game on MirageTable from Microsoft Research [5].

the projector/depth camera system to simulate such scenarios.

To the best of our knowledge, this MirageTable research is the state-of-the-art in visual color-based tracking. However, this research may still not be enough to implement the Ultimate Display [6]. Thus, for visual tracking, another possible method is to use the feature for tracking the position of object. The next subsection will discuss about feature-based tracking in detail.

3. AR SYSTEMS WITH FEATURE-BASED TRACKING

One of the famous augmented reality researches using feature-based tracking is Parallel Tracking and Mapping (PTAM) which was presented by Klein and his group from University of Oxford [7]. In fact, the very first method was presented in his PhD dissertation [8] by observing the movement of a user in 2006. A camera was mounted on the head to capture the movement of location. The position captured by camera was used for analysis the movement of location using real objects and virtual objects. The results of this study demonstrated the flexibility of the movements. The equipment used in this research was inexpensive. However, image data from the camera were still blurred.

Figure 5. Parallel Tracking and Mapping (PTAM) from University of Oxford [7].

Figure 6. Guitar imaging augmented reality application [12].

In 2009, Klein proposed a keyframe-based simultaneous localization and mapping (SLAM) system on a camera phone [7]. Figure 5 displays the sample results. Nevertheless, this system still suffers from the issues of speed and frame-rate. This means that the computational time for the algorithm is high. By applying PTAM, an egomotion estimation algorithm for an Android Smartphone was recently proposed by Porzi [9] in September 2012. This research used an Extended Kalman Filter for improving localization accuracy integrating the information from inertial sensors for supporting his augmented reality applications. Similar research for data association and finding camera trajectory using SLAM for outdoor augmented reality can be also found in [10] by Erkan et al.

Recently, Peter et al [11] proposed a feature-based algorithm for detection and tracking eye gaze of human to support game and consumer electronics (CE) system designers. This system was done in real time. The algorithm was potentially used for a means of direct user input into a gaming environment and improving the accuracy of estimation. This can enhance and support user interface design providing smarter modes of gameplay interaction and user interface modalities.

Next, another method for feature-based tracking is to use ARTag (Augmented Reality Tag). realistic guitar imaging application was proposed as depicted in Figure 6 by Kerdvibulvech [12]. This research creates a system to replicate the look of the guitar. The purpose is to reduce probability of damage on a guitar from customer handling. A user selects a preferred guitar, carries the special mock up guitar with a marker and stands in the camera's frame. The application detects the image and replaces the guitar image on ARTag portion. The user can now see how they look holding the guitar of their choice on monitor even without touching one.

Next, another research using ARTag is an augmented reality application for earrings which was implemented and discussed by Kerdvibulvech in [12]. This system allows a user to try virtual earrings like a real earring. The operation of ARTag is to detect a tag (i.e. ARTag marker) by using a

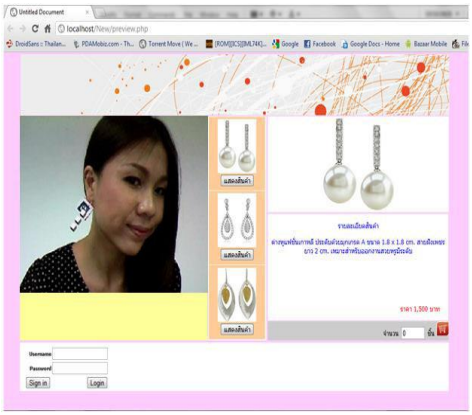

Figure 7. Augmented reality system of virtual earrings [12].

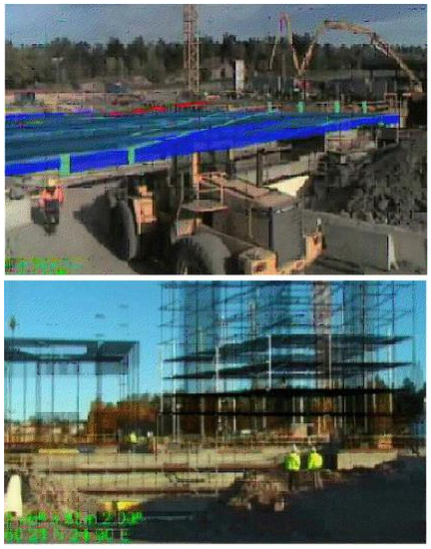

Figure 8. Augmented reality application for construction site visualization in use [13].

single camera for creating simulation of 3D virtual image and displaying it on a computer's screen. Based on this technology, the 3D earrings simulator is developed by ARTag. This system helps users to try many products in 3D simulation by themselves as shown in Figure 7. The users can order the earrings through the online system as a trading process.

Another augmented reality system using the feature-based tracking technique is the mobile system for construction site visualization and interaction which was presented by Woodward et al [13]. As depicted in Figure 8, the system was validated in field tests, covering architectural augmented reality visualization with photorealistic rendering effects, up to construction time applications at a real building site. This system software supported popular standard CAD/Cam 3D-formats, mixing them in real time by using GPS (Global Positioning System). The limitation of this research was that the 3D model building could not be rendered automatically. It had to be uploaded manually. Similar research for a mobile augmented reality framework that achieves scalability and accurate tracking can be also found in [14] by Ha et al.

In July 2011, Park et al [15] presented a method that combines both barcode detection and natural feature tracking methods to track a planar object efficiently on mobile devices. For detecting a region and estimating the initial pose of a QRCode, this method used vertices of position-markers. Typically, a QRCode has an alignment-marker that scans inside the data pattern correctly. It is very challenging to detect a small-scale alignment marker from a camera image. Thus, this method detected the QRCode using only position-markers from the acquired camera image. To do this, QRCode can be used to identify position-markers. From a camera image, the method determined several candidate regions which are produced by closed contour detection. If the candidate regions are possibly position-

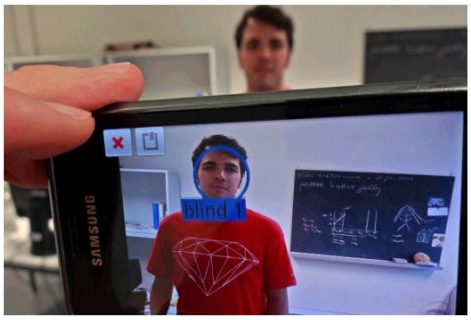

Figure 9. Face augmentation system from ETH Zurich [17].

markers, its own identity will be represented by a hierarchical black-and-white pattern. Similar research for 3-D visual tracking on a mobile phone for augmented reality applications can be also found in [16] by Seo et al in 2011.

Recently, a very interesting research using feature-based tracking is Dantone et al's work [17] which was presented in November 2011. Figure 9 presents some experimental result for tracking the human face. In their research, they implemented an automatic system for face augmentation. This system focused on mobile devices by allowing a user to point his or her mobile phone to a person and the system recognized his or her face. It showed tracked faces in real time and the automatic full crawl information on several social networks, including Facebook. They collected hundreds of thousands of images from Facebook to test the system. However, this system suffered from a security issue because it used anonymous authentication information. The application of this feature based-tracking research was very recently published again in CVPR 2012 for real time facial feature detection [18]. To the best of our knowledge, this method is the state-of-the-art in visual feature-based tracking as it shows the best and promising experimental results.

4. DISCUSSION ON TRENDS OF TRACKING APPROACH IN AR SYSTEMS

In the past ten years, tracking has been a very popular topic for augmented reality researches. This is since it is a basic enabling and challenging technologies for augmented reality. Upon reviewing research published in the last decade in computer science, recent trends and several possible future directions are discussed for further research in this section. Analyzing various tracking methods, two main categories are basically identified.

The first is to incorporate a color-based tracking system to acquire a novel augmented reality system. Tracking methods are mainly based on a color of objects. Basically the accuracy for these color-based algorithms is relatively higher, except when luminance changes immediately. However, this luminance problem has been a long term challenging of computer vision researches. Many algorithms have been proposed to deal with this issue over times, such as an adaptive learning method. Nevertheless, for the adaptive learning algorithm, it is difficult to reinitialize of the tracking system to recover whenever it fails very quickly and suddenly. This is since the color probability used in the adaptive learning method in each frame typically is considered from the previous frames continuously. Thus, if the lighting changes too quickly, the system sometimes cannot

give correct results to converge its probability. This probably leads to tracking failure.

The second is to incorporate a feature-based tracking system to implement many novel augmented reality systems. Typically the computational complexity for these feature-based algorithms is lower, since most of the edge detection algorithms are relatively fast. Also, when changing luminance, this method usually performs well, if comparing to the first relatively. It has been proved in many researches such as [19] that using feature alone cannot produce the robust tracking results.

5. CONCLUSIONS

Visual tracking plays a very important role in computer science and engineering. It offers enormous benefits to research communities in computer science. This paper reviews recent augmented reality works which have been presented in various leading laboratories around the world. We categorized the methods of visual tracking into two main groups. The first group is color-based tracking. The second group is feature-based tracking. Then we discussed each research in detail and suggest the pros and the cons of each technique. We recommend that the selected method should be determined by the purpose of each augmented reality application. This is because color-based tracking and feature-based tracking will work well in different situations. In the future, we believe that the successful visual tracking researches could further shape the promising innovation of computer science society.

REFERENCES

[1] P. Mistry *et al.*, "WUW - Wear Ur World - A Wearable Gestural Interface," in *Proc. ACM CHI Conf. Human Factors in Computing Systems (CHI'09)*, Boston, USA, 2009.

[2] P. Mistry and P. Maes, "SixthSense – A Wearable Gestural Interface," in *Proc. of ACM SIGGRAPH Asia 2009, Emerging Technologies*, Yokohama, Japan. 2009.

[3] C. Kerdvibulvech, "Markerless Vision-Based Tracking for Interactive Augmented Reality Game," *Int. Journal of Interactive Worlds (IJIW'10)*, vol. 2010, Article ID 751615, 14 pages, 2010.

[4] C. Kerdvibulvech, "Asiatic Skin Color Segmentation Using an Adaptive Algorithm in Changing Luminance Environment," *Rangsit Journal of Arts and Sciences (RJAS)*, vol. 2 No. 1, pp. 33-38, 2012.

[5] H. Benko *et al.*, "MirageTable: Freehand Interaction on a Projected Augmented Reality Tabletop," in *Proc. of ACM CHI Conf. Human Factors in Computing Systems (CHI'12)*, Austin, TX, USA, May 5-10, 2012, pp. 199-208.

[6] I. E. Sutherland, "The Ultimate Display," in *Proc. IFIP Congress*, vol. 2, pp. 506-508, 1965.

[7] G. Klein and D. Murray, "Parallel Tracking and Mapping on a Camera Phone," in *Proc. of IEEE/ACM Int. Symp. Mixed and Augmented Reality (ISMAR'09)*, Orlando, USA, 2009.

[8] G. Klein, "Visual Tracking for Augmented Reality," Ph.D. dissertation, Dept. Eng., Univ. of Cambridge, UK, 2006.

[9] L. Porzi, "Visual-inertial Tracking on Android for Augmented Reality Applications," *IEEE Workshop on Environmental Energy and Structural Monitoring Systems (EESMS)*, Perugia, Italy, pp. 35-41, Sep. 2012.

[10] "Vision-based User Tracking for Outdoor Augmented Reality," *IEEE Symp. Comput. and Commun. (ISCC'12)*, Cappadocia, Turkey, pp. 566-568, 1-4 Jul. 2012.

[11] P. M. Corcoran *et al.*, "Real-Time Eye Gaze Tracking for Gaming Design and Consumer Electronics Systems," *IEEE Trans. Consumer Electron.,* vol. 58, no. 2, pp. 347-355, May 2012.

[12] C. Kerdvibulvech, "Augmented Reality Researches." *Manager 360° Magazine*, Thailand in Column iTech 360°, July 2012. (In Thai)

[13] C. Woodward and M. Hakkarainen, *Mobile Mixed Reality System for Architectural and Construction Site Visualization*, ISBN 978-953-307-422-1, Published: December 9, 2011 under CC BY 3.0 license, in subject Human-Computer Interaction.

[14] J. Ha *et al.*, "Mobile Augmented Reality using Scalable Recognition and Tracking," *IEEE Virtual Reality (IEEE VR'11)*, Singapore, 19-23 Mar. 2011.

[15] N. Park *et al.*, "Barcode-Assisted Planar Object Tracking Method for Mobile Augmented Reality," *IEEE Int. Symp. Ubiquitous Virtual Reality (ISUVR)*, Jeju, South Korea, pp. 40-43, 1-3 Jul. 2011.

[16] B. Seo *et al.*, "3-D Visual Tracking for Mobile Augmented Reality Applications," *IEEE Int. Conf. on Multimedia and Expo (ICME)*, Barcelona, Spain, 11-15 Jul. 2011.

[17] M. Dantone *et al.*, "Augmented Faces," *IEEE Int. Workshop on Mobile Vision for Int. Conf. Comput. Vision (ICCV'11)*, Barcelona, Spain, Nov. 2011.

[18] M. Dantone *et al.*, "Real-time Facial Feature Detection using Conditional Regression Forests," *IEEE Comput. Society Conf. Comp. Vision and Pattern Recognition (CVPR'12)*, Rhode Island, New England, USA, Jun. 2012, pp. 2578-2585.

[19] C. Kerdvibulvech and H. Saito, "Model-Based Hand Tracking by Chamfer Distance and Adaptive Color Learning Using Particle Filter," *EURASIP Int. Journal on Image and Video Processing (JIVP)*, vol. 2009, Article ID 724947, 10 pages, 2009.