Interaction in Marker-less Augmented Reality Based on Hand Detection Using Leap Motion

Juncheng Zhao
School of Electrical and Electronic Engineering Nanyang Technological University
Hock Soon Seah
School of Computer Science and Engineering Nanyang Technological University

Abstract

In this paper, a novel interaction framework for marker-less Augmented Reality is introduced. Model-based detection is one of the solutions for marker-less Augmented Reality. Instead of a marker, human hand is used as a distinctive object on which the augmented object placed. Leap Motion, a Virtual Reality (VR) hand tracking device, is used to detect the hand in our interaction framework for marker-less Augmented Reality. 3D hand position, gesture and direction can be obtained by using Leap Motion and passed to Unity 3d game engine. The main task for this framework is to calibrate the actual hand and the virtual hand generated by computer so that they can overlay each other. With the help of marker-less Augmented Reality Framework, a user can experience intuitive interaction with virtual object and natural occlusion which will be the core functionality for next generation game, education, user interface and industrial.

Keywords: Augmented Reality, Interaction, Leap Motion, Marker-less, Interaction Framework, Hand Detection

Concepts: • Computing methodologies~Mixed / augmented reality; Image processing;

1 Introduction

Recently there has been a wide range of research in the application of Augmented Reality. Advanced AR technology allows users to experience augmented objects together with the real world simultaneously. Augmented objects are able to overlay physical elements in real-world environment according to how the users interact with them. Both marker-less and marker-based AR can achieve this sort of interactive manipulation.

Most existing AR applications, including Pokémon go and Snapchat, make use of 2D image processing and detecting approaches [Hlavac and Boyle 2014]. However, these 2D methods cannot achieve accurate recognitions of complex gestures and objects. In contrast, integrated 3D devices using holograph [Bimber and Raskar 2005], like HoloLens, are able to provide vivid rendering and gesture recognition. However, these devices require multiple cameras and sensors including world-facing camera, environment understanding cameras and mixed reality capture [Billinghurst and Kato 1999].

One of the main issues for 3D hand recognition today is how to reconstruct the object captured and naturally overlay the real world object. Since 3D hand recognition using inferred camera often experience distortion which impedes the reconstruction of objects.

In this paper, hand detection and reconstruction are achieved by using 3D and 2D hand recognition. To develop 3D hand and gestures detection, a Leap Motion device is used as the tracking device. To overlay the real hand with virtual hand, a FOV 150° wide range camera is used to detect fingertips and map them to the virtual hand. The rest of the paper includes related works, the objective, the proposed approach, the result and the conclusion and future works.

2 Related Works

A. Leap Motion

In order to detect human hands in the real world, Leap motion [Jakus et al. 2014] is used to track the hand position, gestures and orientations at real time. Analysis and evaluation on the robustness and accuracy [Weichert et al. 2013] of Leap Motion device have been done before using it. The Unity 3D plugins for Leap Motion enables users to create hand skeletons, new gestures and display in Unity 3D. Using two monochromatic IR cameras and three infrared LEDs, the device observes a roughly hemispherical area, to a distance of about 1 meter. The LEDs generate pattern-less IR light and the cameras generate almost 200 frames per second of reflected data.

B. 2D Fingertip Tracking

In optical-based interactions, 2D hand and fingertip tracking is a cheap solution that is widely used to provide an easy way to provide interaction between hand and augmented objects in AR. In Marker-less Augmented Reality system, hands can be the base on which augmented objects can be placed upon successful detection [Lee and Hollerer 2007].

3 Objective

The objective of this paper is to introduce a Marker-less Augmented Reality framework which enables users to interact with the augmented objects using their hands by using both 2D and 3D hand recognition. Recently Augment Reality applications lack interactions between users and the virtual objects. The main propose of this framework mentioned in this paper is to make Augmented reality more interesting by adding interactions. With this framework, future Augmented Reality applications in the fields

*e-mail:zhao0192@e.ntu.edu.sg.

VRCAI '16, 03-04, 2016, Zhuhai, China
ISBN: 978-1-4503-4692-4/16/12
DOI: http://dx.doi.org/10.1145/3013971.3014022

of game, education, UI design, and tele presentation are able to provide users with more hand-based interactions.

4 Proposed Approach

Overall Steps— Figure 1 shows the overall steps for the Marker-less Augmented Reality Framework which consists of four main steps. The first step is to obtain the image from a wide range camera and import into game engines. In the second step, the hand model obtained by Leap Motion will be used to draw a hand skeleton in the game engine. The third step for this framework is to calibrate the image obtained by the wide range camera and the skeleton so that they are able to overlay each other. Fingertip Tracking and calibration method is used to adjust the transform of hand skeleton. The last step is to augment virtual objects and add interactions based on gestures and hand position.

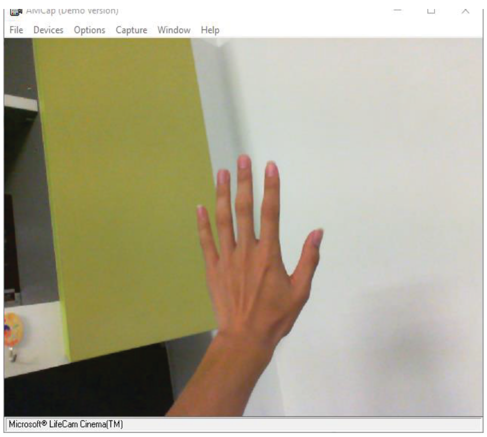

Figure 1.1: *Steps for Marker-less Augmented Reality Framework*

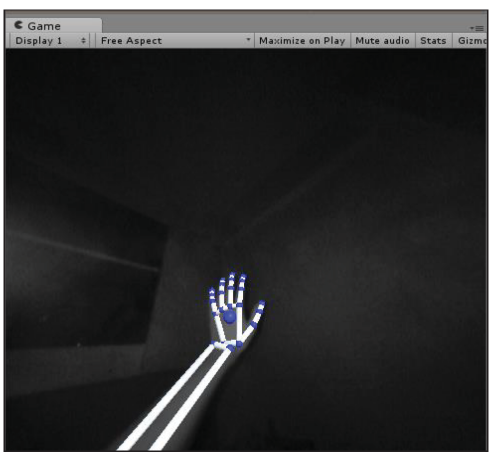

Figure 1.2: *Steps for Marker-less Augmented Reality Framework*

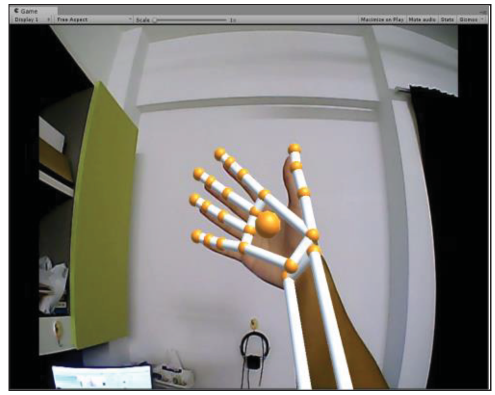

Figure 1.3: *Steps for Marker-less Augmented Reality Framework*

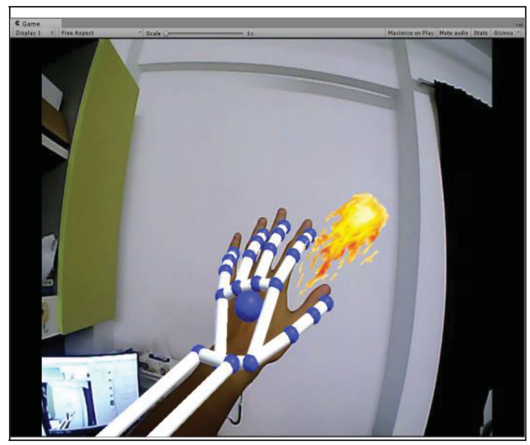

Figure 1.4: *Steps for Marker-less Augmented Reality Framework*

5 Implementation

A. Unity3D Leap Motion API

Leap Motion provides Unity3D assets to developers who are going to create Unity 3D project with Leap Motion device. In order to create a hand skeleton in Unity3D, controller class needs to be called to connect to Leap Motion Device. Then tracking data from the Leap Motion API will be used. The Leap Motion classes are defined in the Leap namespace. Controller object will be generated to connect to the Leap Motion device. The Controller object automatically establishes a connection to the Leap Motion service which then passes tracking data, in the form of Frame objects to your application. For each frame, a skeleton of hand can be drawn in Unity3D by using functions in Hand class when hands are detected in the frame.

B. Fingertip Tracking & Calibration

Since Leap Motion device uses IR (Infrared) camera to obtain hand model and gesture, there must be distortion that lead to the inaccuracy in scale and position of hand model. Thus, to undistort and calibrate the real hand and hand skeleton, mapping need to be done. To segment the hand region, an adaptive skin color-based method [Plopski et al. 2015] is used. Skin color model [Jones and Rehg 2002] provided by the statistical color models is used to determine pixels inside hand region. If the probability distribution $P(rgb)$ is larger than a threshold prefixed, it is inside the hand region.

$$\mathrm{P(rgb)} = \frac{\mathrm{C[rgb]}}{\mathrm{T_c}}$$

Where $c[rgb]$ gives the count in histogram bin and T_c is the total count that sums up all the bins in the histogram. By applying this methodology, the detected hand contour can be drawn.

Next, fingertips are going to be detected. Based on the Fingertip Tracking algorithm, fingertips are detected from the contour of a hand using a curvature-based algorithm. That is, the candidates of fingertips are the points with higher curvature values than a threshold by computing a dot product of ith point on the contour to the preceding and succession points.

$$K_l(P_i) = \frac{\overrightarrow{P_\iota P_{\iota-1}} \cdot \overrightarrow{P_\iota P_{\iota+1}}}{\left\|\overrightarrow{P_\iota P_{\iota-1}}\right\|\left\|\overrightarrow{P_\iota P_{\iota+1}}\right\|}$$

Where l is the displacement index and the best value are integers between 5 to 25 which is learnt from experience.

After the fingertips being detected, mapping between the real hand fingertip and hand skeleton fingertip will be done to rearrange the position and scale of the hand skeleton. After mapping, the two hands will be able to overlay each other. Figure 2 shows the expected output for the mapping.

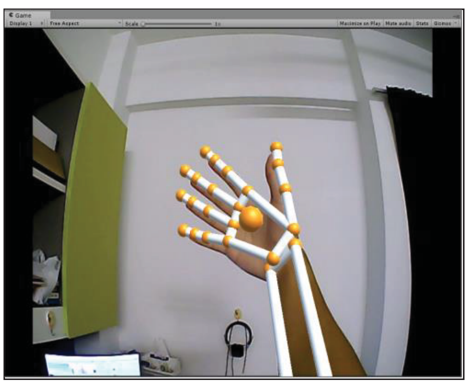

Figure 2: *Expected Output for Calibration*

6 Expected Result & Future Works

The proposed Marker-less Augmented Reality Framework will be able to support simple gestures such as waving, pointing, zooming in, zooming out, clenching and so on. Based on the gestures of the hand, users will be able to interact with the augmented object. As shown in Figure 3, if player opens his/her hand and points outwards, a fire ball will shoot in the direction of the palm. In Figure 4, users will be able to interact with virtual bubbles generated and burst the bubble when point at them. For UI design, a navigation menu will be displayed based on the position of the hand and user can use the other hand to select the elements he/she wants.

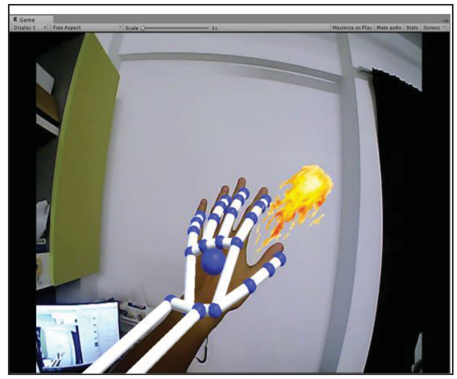

Figure 3: *Fireball*

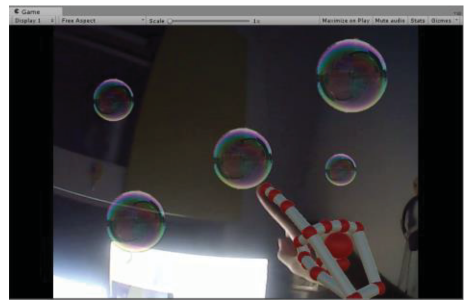

Figure 4: *Interaction with Virtual Bubbles*

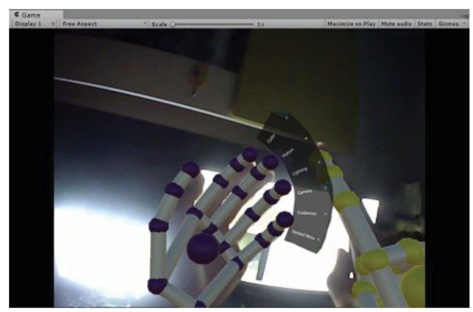

Figure 5: *Virtual UI Design*

The author of this paper will continue his work on the calibration and interaction by enhancing the mapping algorithm. More gestures will be added and applications will be developed using this framework. Meanwhile, HMD (Head Mounted Display) application will also be developed using EPSON Moverio BT-200 together with the Marker-less Augmented Reality Framework.

7 Conclusion

In this paper, the author presents a Marker-less Augmented Reality Framework which is composed by a Leap Motion device, a 150° wide range camera and a display screen. This framework will overcome the shortage of 2D hand detection and provide intuitive interaction and accurate gesture detection. Unlike the Traditional 3D hand detection, which returns distorted images and difficult to undistort. The framework will do the calibration and, at the same time, provide a cheap solution for Augmented Reality.

Acknowledgements

This project is supported by the National Research Foundation, Prime Minister's Office, Singapore under its IDM Features Funding Initiative, Nanyang Technological University (NTU) and Multi-Platform Game Innovation Centre (MAGIC).

References

Sonka, M., Hlavac, V. and Boyle, R., 2014. *Image processing, analysis, and machine vision*. Cengage Learning.

Bimber, O. and Raskar, R., 2005. *Spatial augmented reality: merging real and virtual worlds*. CRC press.

Billinghurst, M. and Kato, H., 1999, March. Collaborative mixed reality. In *Proc. Int'l Symp. Mixed Reality* (pp. 261-284).

Guna, J., Jakus, G., Pogačnik, M., Tomažič, S. and Sodnik, J., 2014. An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. *Sensors*, *14*(2), pp.3702-3720.

Weichert, F., Bachmann, D., Rudak, B. and Fisseler, D., 2013. Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, *13*(5), pp.6380-6393.

Lee, T. and Hollerer, T., 2007, October. Handy AR: Markerless inspection of augmented reality objects using fingertip tracking. In *2007 11th IEEE International Symposium on Wearable Computers* (pp. 83-90). IEEE.

Plopski, A., Itoh, Y., Nitschke, C., Kiyokawa, K., Klinker, G. and Takemura, H., 2015. Corneal-imaging calibration for optical see-through head-mounted displays. *IEEE transactions on visualization and computer graphics*, *21*(4), pp.481-490.

Jones, M.J. and Rehg, J.M., 2002. Statistical color models with application to skin detection. *International Journal of Computer Vision*, *46*(1), pp.81-96.