Hand Valley Point Detection

BTP Endsem Report

Submitted By

Balaji K V

BTech (Dual Degree)
in
Computer Science and Engineering

Roll No. 17CS02001

Under the guidance of

Dr. Debi Prosad Dogra

School of Electrical Sciences
Indian Institute of Technology Bhubaneswar

2020

Abstract

In this report, I present a methodology to detect four valley points in a bare human hand using computer vision algorithms. This is required to develop a vision-based framework and system for human-computer interaction in an Augmented Reality environment. The application being designed enables users to hold and interact with virtual 3D objects augmented on to their hands. It uses an efficient colour segmentation algorithm which also works in real-time scenarios. The dominant features(valley points) are detected and tracked to estimate the camera pose relative to the user's hand.

Index Terms - Augmented Reality, Camera Pose Estimation, Hand Segmentation, Convexity Defects

1. Introduction

Augmented Reality (AR) provides a new paradigm for human–computer interaction (HCI) in which virtual and real objects can be fused in the user interface. Recently developed tracking and interaction methods in AR allow users to work with and examine the real physical world, while controlling augmented objects in the system in a more feasible fashion. In general, an AR system requires some indication of where exactly the virtual objects should be augmented. This has been conventionally accomplished by AR markers such as ARTag. However, in recent marker-less AR is considered to be ideal since it does not require the forethought of adding markers to a scene.

To register virtual objects accurately in an AR system, the relative position of the camera with the scene is required. Six degrees of freedom (6DOF) that defines the position and orientation of the camera relative to the scene have to be tracked and virtual objects are then augmented in the specific position in the AR environment. The interactions with hands can be roughly classified into two groups—device-assisted hand interaction and bare- hand interaction. In bare-hand interaction, the natural features on the hands, such as the skin color, shape, and so on, are detected and tracked by processing the images captured by the cameras. To be an effective interaction method, bare-hand interaction, which is based on computer vision, should be able to work successfully under uncontrolled light conditions and background.

In this report, I present a technique to detect hand valley points using skin colour segmentation that are further used to reconstruct camera's pose relative to the user's hand.

The rest of the report is organised as follows: section-2 reviews related works, section-3 describes the problem statement, section-4 gives the proposed methodology, section-5 list downs the results, section-6 gives future work plan and section-7 provides the conclusion.

2. Related works

The use of hand gestures provides an attractive alternative to cumbersome HCI devices.Although glove-based devices can measure the hand movements with high accuracy and speed, glove-based gesture interfaces require the user to wear cumbersome devices that are connected to the computer via cables. This is more costly, limits the user's motion, and hinders the natural interaction feel.

Computer-vision-based methods have been considered to be more promising for natural HCI in [1]. The studies of vision-based hand detection using normal cameras, such as the web cameras, have also been explored. Hand interaction based on normal cameras would be more appealing to everyday users. To extract the image regions corresponding to the human skin, either background image subtraction [2] or color segmentation [3] is typically used.

In the HandyAR system [4], the human hand is used as a distinctive marker to render virtual objects on the palm. Seo et al. in [3] presented an AR system on mobile devices in which the virtual objects are augmented onto the palm. The virtual objects would react according to the opening and closing movements of the hand, which can be detected through tracking the fingertips.

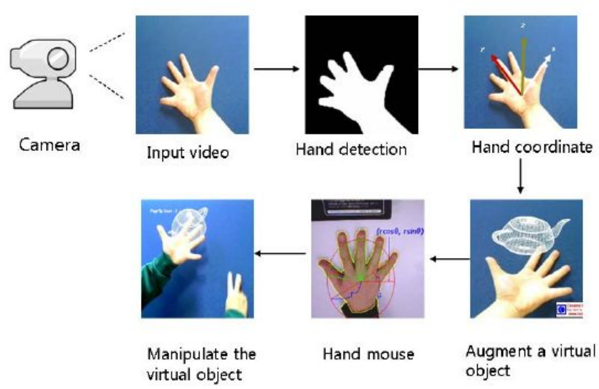

Figure 1. Steps for developing a marker-less AR System

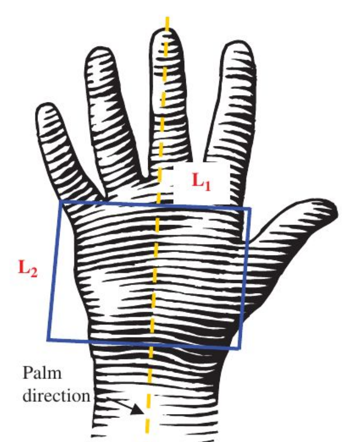

Figure 2. Virtual Square defining coordinate system

The four points are changing during the hand movements as the palm direction is varying with the movements such that the coordinate system created on the palm would not be as stable as the method proposed in this report. The four convexity defect points on the palm are detected and tracked to estimate the camera pose. For the same user, the 3D coordinates of the four convexity defect points are fixed when a coordinate system is established on the palm. As compared with the fingertips, the four convexity defect points are relatively static during the hand movements. Therefore, the virtual objects rendered would be relatively stable.

3. PROBLEM STATEMENT

Hand valley point detection involves detection of four valley points between the fingers which relatively remain static while the hand is moved in different directions. This stability results in a robust coordinate system which helps in augmenting virtual objects without less failures.

Input: A frame in a video.

Output: Coordinates of valley points in the hand detected in the video frame.

4. PROPOSED METHODOLOGY

4.1. Hand Segmentation

For hand region detection we utilize YCbCr skin color model which is proven to detect skin region effectively to segmenting hand region. Skin color models have been widely used for hand and face detection because the use of color information can simplify the task of hand localization in complex environments. Mostly the primary components of R,G,B are used for skin segmentation.

We use HSV and YCbCr masks to filter out the image area that is in range of skin colour parameters.It is known that the chrominance components of the skin color are independent of the luminance component. Moreover, in this work, by disregarding the luminance component, robustness of skin detection can be obtained in the case of variations in lighting conditions.

Skin colour range is as follows:

(a) HSV Values - (0, 40, 0) to (25, 255, 255)
(b) YCbCr Values - (0, 138, 67) to (255, 173, 133)

The output produces a binary image with pixel value as 1 for skin region and 0 for non-skin region.

4.2 Hand Contour

As the hand segmentation algorithm results in multiple contours of different sizes, the largest contour (of the hand) is filtered by sorting the contours by their areas using openCV functionalities.

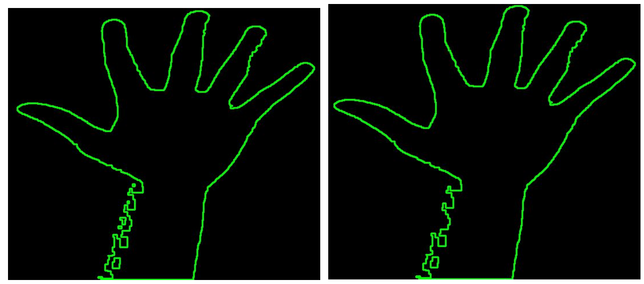

Figure 3. Filtering largest contour

4.3. Convex Hull Plot

The convex hull of a set of points is the smallest convex polygon that contains all the points in the set. Convex hull polygon is plotted around the hand contour for detecting convexity defect points in further analysis.

4.4. Hand Valley Points

The valley points in the hand are nothing but the convexity defect points of the current convex hull and contour system. These points are those present on the hand contour and farthest from the edge of the convex polygon formed between the points of contour that touch the polygon.

The coordinates of the detected points are extracted and corresponding coordinates in openGL frame are calculated for camera pose estimation work which is carried out in openGL framework. This is done by just subtracting the y-coordinates from the height of the image. This is done because the origin is at top-left corner in openCV whereas it is at bottom-left corner in openGL.

5. RESULTS

Figure 4. Skin colour Segmentation

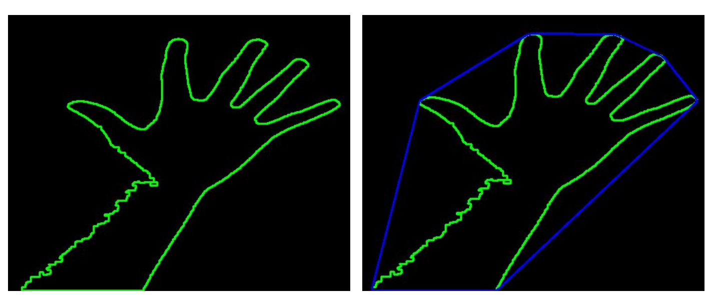

Figure 5. Convex Hull plot

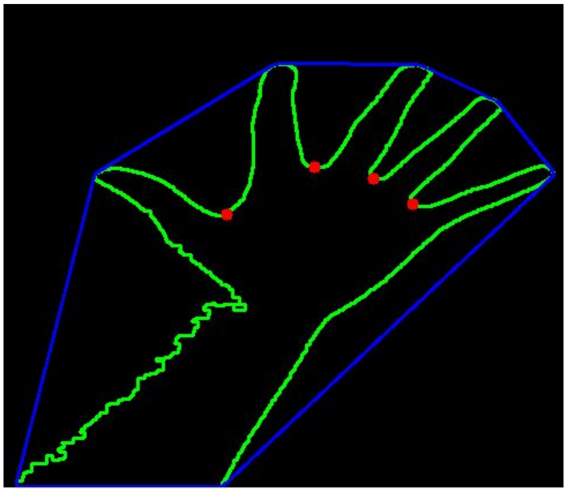

Figure 6. Hand Valley Points (Red Circles)

Valley points are detected as shown above in each video frame and their coordinates are to be used to create a coordinate plane on to which virtual 3D objects are to be augmented.

The skin colour segmentation algorithm used detects hand only in predefined skin colour ranges. It misses out if skin colour out of range and also detects other objects which are in skin colour.

6. FUTURE WORKS

6.1. Better Hand Segmentation

We have used a parametric model to recognise skin coloured pixels in the image. It is difficult to achieve a high recognition rate with explicitly defined skin models and the parametric methods in an uncontrolled environment. The key idea of the nonparametric methods is to estimate the skin color distribution from the training data without deriving an explicit model of the skin color. one such method uses the CAMSHIFT(Continuously Adaptive Mean Shift) algorithm that is computationally efficient and can deal with the challenges of irregular object motions due to variations in perspective and image noise. It can overcome problems such as changes in hand shape during motion and variation of skin colour across the hand.

6.2. Camera Pose Estimation

Once a robust framework is designed for detecting the hand features (valley points), we can move on to camera pose estimation where we reconstruct the camera pose relative to the hand to augment virtual 3D objects.

7.CONCLUSION

In this report, a skin colour segmentation method is proposed to detect hand valley points which are used as hand features for camera pose estimation in an AR application. The hand region is segmented and tracked based on skin colour distribution. With the features detected from the hand region, the user can interact with 2D objects using a bare hand. The camera pose can be estimated based on the tracking feature positions in order to augment virtual objects on the palm of the user's bare hand 3D object interaction.

REFERENCES

[1] Erol, A., Bebis, G., Nicolescu, M., Boyle, R. D., & Twombly, X. (2007). Vision-based hand pose estimation: A review. Computer Vision and Image Understanding, 108(1–2), 52–73.

[2] von Hardenberg, C., & Bérard, F. (2001). Bare-hand human–computer interaction. Proceedings of the ACM Workshop on Perceptive User Interface, 113–120.

[3] Seo, B.-K., Choi, J., Han, J.-H., Park, H., & Park, J. (2008). One-handed interaction with augmented virtual objects on mobile devices. Proceedings of VRCAI 2008 [CD-ROM].

[4] Lee, T., & Höllerer, T. (2007). Handy AR: Markerless inspection of augmented reality objects using fingertip tracking. Proceedings of IEEE International Symposium on Wearable Computers (ISWC), 83–90.

[5] Y. Shen, S. K. Ong & A. Y. C. Nee (2011) Vision-Based Hand Interaction in Augmented Reality Environment, International Journal of Human–Computer Interaction, 27:6, 523-544, DOI: 10.1080/10447318.2011.555297

[6] B. Lee and J. Chun, "Interactive Manipulation of Augmented Objects in Marker-Less AR Using Vision-Based Hand Interaction," *2010 Seventh International Conference on Information Technology: New Generations*, Las Vegas, NV, 2010, pp. 398-403, doi: 10.1109/ITNG.2010.36.

[7] Ng, Kah & Tan, Guat & Wong, Ya-Ping. (2013). Vision-Based Hand Detection for Registration of Virtual Objects in Augmented Reality. International Journal of Future Computer and Communication. 423-427. 10.7763/IJFCC.2013.V2.198.