Hand's 3D Movement Detection with One Handheld Camera

Mingming Fan*, Liang Zhang, Yuanchun Shi
Dept. of Computer Science and Technology, Tsinghua University, Beijing, China

Abstract

This paper presents a scheme to create a real-time and reliable method for recognizing vision-based hand's 3D movement and to use the movement parameters for controlling 3D objects. The algorithm for 3D movement detection is totally based on analyzing feature points from the only camera in user's hand. As the algorithm is based on frames captured from one camera in untrained environment, it's difficult to distinguish similar movements on optical flow images, especially between shifting and rotating. A novel differentiation algorithm by voting from some weak classifiers is used. The algorithm provides a method of direct mapping user's hand movement to object control. We design an application of controlling a virtual 3D cube's movement and estimate the accuracy of the algorithm. And the experients' result presents that the 3D movement detection algorithm is efficient and robust enough for real-time interaction.

Keywords: Virtual Interaction, 3D movement, handheld camera, features points, Classifiers

CR Categories: I.3 [Computer Graphics]: ;— [H.5]: Information Interfaces and Presentation—(e.g., HCI)

1 Introduction

Using the human hand as an input-device is an interesting topic in Virtual/Augmented-Reality (VR/AR) research. There is a kind of image-based hand gesture recognition system[Schlattmann and Klein. 2007] , in the system more than one cameras look at user's hands, in the case, hand gestures must be operated in the camera view field. In order to overcome this limitation , we propose that if the camera is held in the hand, then the movement of the camera will reflect the hand gesture which no longer be restricted within the camera view field..

Since we use only one camera and the input is 2D image, there are two major challenges in our work. One is how to distinct the shift and rotate movements, because these two movements (e.g. shift by X axis, rotate by Y axis) are very similar to each other. To solve this problem, we create a boosting classification composed of three weak classifiers. The other challenge is how to clearly distinct the 6 DOF （degree of freedom） when the hand's movement swithes among them. To crack this hard nut, we have also figured out how to detect the movements of zoom and rotate by Z axis. In order to evaluate our algorithm, we design an 3D virtual control task. Users are requested to control the movement of the virtual cube by holding a camera in the hand. And the result shows that the cube's movement is almost the same as the hand, which implies algorithm is accurate and real-time.

*e-mail: fmmbupt@yahoo.com.cn

VRST'08, October 27-29, 2008, Bordeaux, France

2 Hand's 3D motion detecting algorithm

The hand's 3D movement consists of different parameters of the 6 DOF, but only the major one represents the user's purpose, the others may be caused by the hand's jitters, so our task is to recognize the major movement from the hand's component movement.

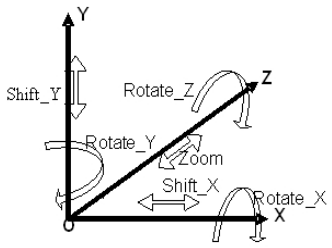

Figure1.the 6 DOF of 3D movement

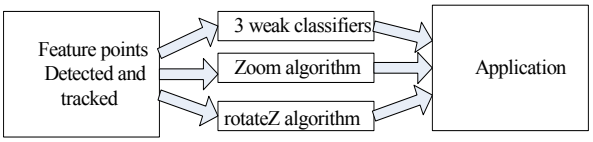

Figure2. The flowchart of the whole algorithm

2.1 Feature Points Detecting & Tracking

Corner-like points [Shi and Tomasi 1994] are easy to find on incoming frames and are relatively stable while being tracked[Bouguet 1999]. The 3D motion detecting algorithm is based on geometric manipulation of the points.

Figrue3. The green points in the image are corner points detected by the algorithm. We will track these points during the moving.

2.2 3D Motion detecting algorithm

Shift in x or y axis is calculated by the average change in corner points' coordinates from two neighboring frames, which will

exclude the influence of zooming and rotation by z on coordinates(see figure1). Zooming rate can be calculated by the change of the distance between the corner points and their center.(see Figure3) The degree of rotation by z can be detected from the change of slopes of the lines between corner points and the center point.(see Figure3)

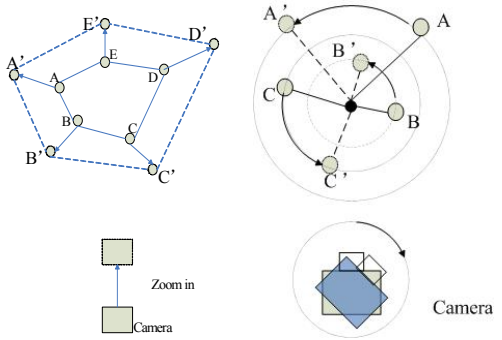

Figure3. (a) shows points radiate or shrink from the center point in zooming. (b) shows points in rotation by z axis go in a clockwise or anticlockwise direction

Since the shift x/y axis and rotation y/x axis cause the same changing directions for all the points(see Figure1), it's quite difficult to distinct them. We create some novel weak classifiers. The final result will be voted by the three weak classifiers.

Weak classifier 1: Classified by the changing of the shape of the object

When the hand shifts by x or y axis, objects viewed by the camera also shift with the same shape, but when the hand rotates by x or y axis, the shape of objects viewed by the camera will change (Figure4:a and b).

Weak Classifier 2: Classified by the differences between two classes of feature points

Suppose the points do not update during the movement, then according to the different changes of the corner points' coordinates, we divide points into two groups. The center of groups are A and B. The distance between A and B are almost the same while shifting, but quite different while rotating.(Figure4).

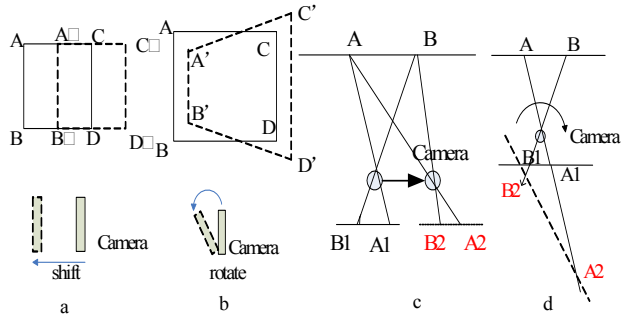

Figure4.The shape of the observed points doesn't change while shifting, but changes while rotating.(Figure a and b) $|A_1B_1|-|A_2B_2|$ is nearly zero while shifting ,but changing all the time while rotating(see Figure c and d).

Weak Classifier 3: Classification by the update speed of feature points in the view field

Based on the experimental testing, the system update frequency (while a lot of feature points move out of the view field during the camera movement till the number of the feature points in the view field becomes less than 3, the system needs to re-detect new feature points, i.e., the system needs to be updated.) is slower under a shift gesture than a rotation gesture.

3 A virtual control task

We design an application in which users hold the camera in hand to control the virtual cube's 3D movement. We did the experiments in different environments and evaluated the accuracy of the algorithm.

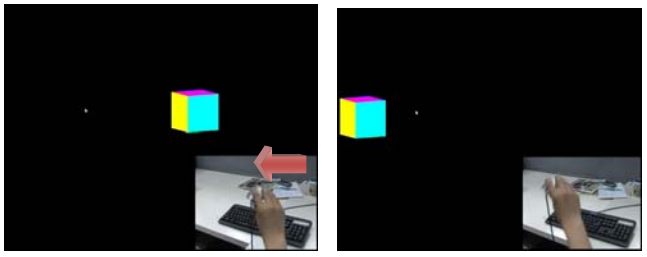

Figure5. The first picture shows that a user waves the hand in middle air, our algorithm detects the motion and directly manipulates it into the cube's movement. The following two images show the cube shifts when the hand shifts. The read arrow shows the direction of the hand's motion.

4 Conclusion

From the interaction perspective, the paper provides a basic 3D user interface for controlling the application. Our algorithm can be used in VR games and the large display interaction. The virtual control task demonstrates that the algorithm is real-time and accurate. We have also tested the algorithm in different backgrounds. The result of the evaluations shows that our algorithm works well in different background except blankness, single color or fast moving background.

References

BOUGUST, J.V. 1999. Pyramidal implementation of the Lucas Kanade Feature Tracker Description of the algorithm. Intel Corporation Microprocessor Research Labs.

SCHLATTMANN, M. and KLEIN, R. 2007. Simultaneous 4 gestures 6 DOF real-time two-hand tracking without any markers.In *Proceedings of ACM VRST 2007*.

SHI, J. and TOMASI, C.1994. Good features to trac. In *Proceedings of IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recognize.*.