Gesture recognition in Virtual Reality

What does gesture recognition offer for virtual reality entertainment and how is it implemented?

Stephan Zuidam

Amsterdam University of Applied Sciences

Virtual Reality Minor 2017

Introduction

Gesture recognition (henceforth GR) is a technology that tracks user movement and can recognize certain pre-defined movements, gestures, and respond to that. Gesture recognition is widely used on mobile applications where the user's swipes are tracked for different fingers: swiping with two fingers translates to action A or drawing a circle triggers action B.
Basic implementations of GR in virtual reality entertainment can be seen in the game Left-Hand Path[1] by Strange Company. The player, equipped with a staff in their right hand, have to *draw* figures with their left hand to cast spells. Different gestures cast different spells. Another example is Oculus First Contact[2] by Oculus. The player can interact with a small flying robot and wave at the robot. The game recognizes the user making a waving motion and the robot responds to this by waving back at the player.
Being able to recognize certain movement by the user in Virtual Reality opens up a whole lot of possibilities for user experience as well as bridging the gap between the real and virtual world.

Relevance

Virtual Reality aims to immerse the user into a different world. Currently the user is *restricted* to use the head mounted display (HMD) and the controllers to experience a virtual reality. The controllers are needed to manipulate the world and this restriction takes away from the immersive experience. Manipulating the world with non-physical devices, like voice, lip reading, expressions and recognition of gestures eliminate this restriction. GR doesn't necessarily have to be for finger, arm and hand movements only. It can be applied to the user as a whole, full body. However, with the current market available technology, the only device that tracks the user is the HMD (head) and the controller (hands).

[1] http://store.steampowered.com/app/488760/LeftHand_Path/
[2] https://www.oculus.com/experiences/rift/1217155751659625/

Implementation

GR, as the name suggests, has to recognize certain gestures and the shapes they produce. Therefor a library of possible gestures has to exist. The user produces a figure by moving the controllers, the system filters out the movement and creates an image, the image is compared to the library of known gestures and returns if it's an existing gesture or not. This process can be broken down into four different steps:

Gesture collection

A library of pre-determined gestures has been created with possible deviations from the shape to get a general shape for the correct image.

Data acquisition

The user moves the controller and thus the sensors tracking the controller. The system collects the path the controller traveled and the orientation and specific points.

Image processing

The gesture the user made is filtered out and normalized to create a capture of the moment, an image that can be compared to any of the existing gestures.

Recognition

After the image processing stage, the image is compared to the images in the gesture library and determines which gesture looks most like the produced gesture. If there is a math the system can respond to this. Otherwise nothing happens or the user is notified the gesture is not recognized.

Future

The described implementation is limited to single handed motions and gestures. If the user were to cross their arms to create a *stop sign* the system would not recognize it as a proper gesture. When implementing GR the first thing is to establish the value of GR in the project and what type of gestures the user is expected to make and use.
One other problem is '*when* do you track the user's motions?' For complicated gestures the user generally has to hold a button and then perform the gesture. This is done to make sure that random movements when the user does not intend to perform a gesture is recognized as a known gesture. Plus, constantly tracking the user's motions and comparing these to the library of gestures is a heavy performance and could result in a dip in the run of the virtual experience.

Conclusion

The challenge in creation GR for VR experiences depends on the precision the gestures have to be recognized. 3D images have to be compared to pre-determined images which restricts the user to a limited set of actions.
However, GR is an immersive solution to the problem of having intrusive interfaces. And it has the potential to completely rid of controllers making the experience only more immersive.

Sources

- "A BASIC GESTURE AND MOTION FORMAT FOR VIRTUAL REALITY MULTISENSORY APPLICATIONS" by Annie Luciani, Matthieu Evrard, Damien Couroussé, Nicolas Castagné, Claude Cadoz, Jean-Loup Florens
  https://arxiv.org/ftp/arxiv/papers/1005/1005.4564.pdf

- "Gesture Recognition and Its Importance in VR" by Appreal-VR
  https://appreal-vr.com/blog/gesture-recognition-in-virtual-reality/