Neck gesture recognition by using constancy of head turning

Hiroki Taguchi , Hisataka Suzuki , Yannick Littfass , Akihiko Shirai
ShiraiLab, Kanagawa Institute of Technology
Email: neckui@shirai.la

ABSTRACT

This article proposes a motion recognition method for Oculus Rift, which is currently the most popular head mounted display available, by using neck motion as non-verbal interaction. The method can define three gestures, namely, “agree,” “disagree,” and “make a question,” by determining the constancy of the head turning angle.

Related Work

Hyun proposed external gesture recognition software by using the upper-body kinematics information of game players. However, given the occlusion of the HMD user, the accuracy of the detected neck angle differs between two Kinect devices.

Sakai proposed the AccuMotion method, which can separate natural movements by using the kinematics of the accumulated motion of a user by utilizing the inner product of vectors.

Constancy of neck and head movement

The neck is a complex structure comprising seven bones, which are difficult to define as matrices. However neck gestures are often used to tell “Yes,” “No,” and to ask questions in the major linguistic area. The neck and head movement Environment range has constancy between each user . A value is used to obtain the constant property of neck movable angle when users equip Rift in several subjects.

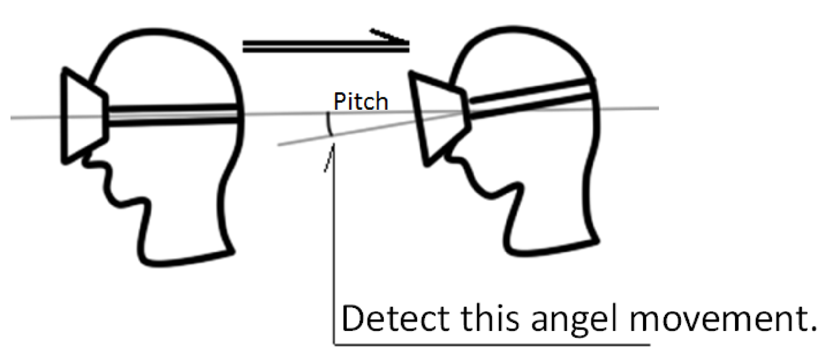

Figure 1: What is “Constancy of neck and head movement.”

Environment

We used Rift and its sensor fusion, which contains magnetometers, gyroscopes, and accelerometers. By combining the values obtained by the sensor fusion, we can determine the rotation angle of the user's head in three dimensions in accordance with the right-handed coordinate system. The rotation for each axis is given in the form of pitch-yaw-roll (see Figure 2). In the X axis, the rotation of the upper direction is positive pitch. In the Y axis, the rotation of the left direction is positive yaw. The rotation of the Z axis is positive and tilts to the left in the XY plane roll. These values are obtained from OVRDevice.GetPredictedOrientation in Oculus SDK as rotation of yaw-pitch-roll. We converted the values from Quaternion type to Vector3 type.

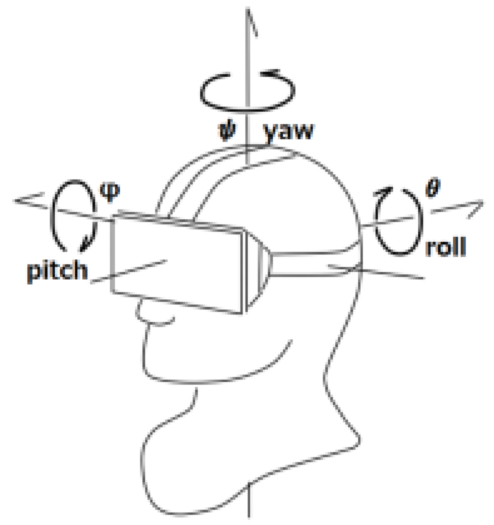

Figure 2: Coordinates in this experiment.

Experiment protocol

Subjects wore Rift and recorded six operations. They watched “center-bottom,” “top-center,” “left-center,” “center-right,” “tilted left,” and “tilted right” scebes for 10 times each. The scene contained spherical textures, which are taken in an actual room and one computer-generated cube. The graphics are generated by Unity3D. Subjects must imagine the center and ground plane from visual feedback and somatesthesia.

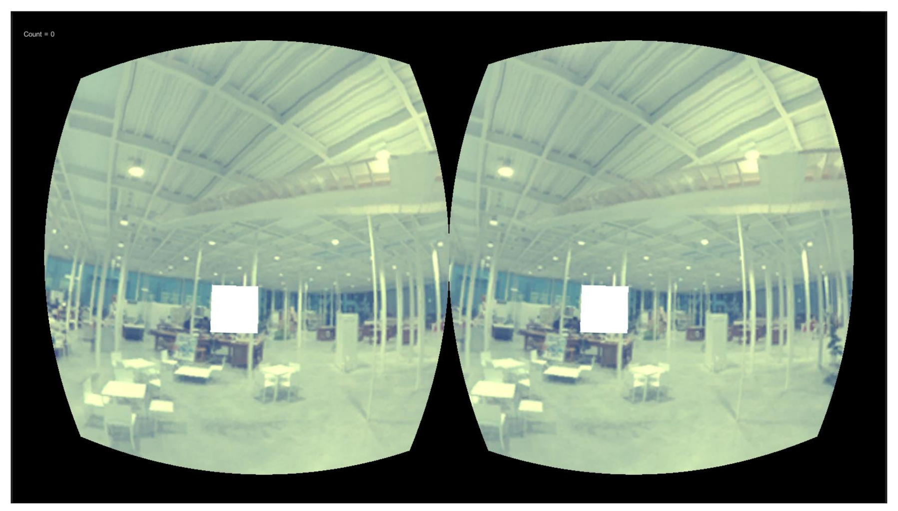

Figure 3: Screen of the subject’s viewing.

Results

Figure 4 shows the average and deviation of the pitch-yaw-roll angle when four subjects watch the center. Roll is stable between subjects, but pitch and yaw are spread within +/-10° between subjects. Pitch can be explained as the difference of the center of gravity of a fixed position or the head. Yaw is the summation of the errors of the magnetometers.

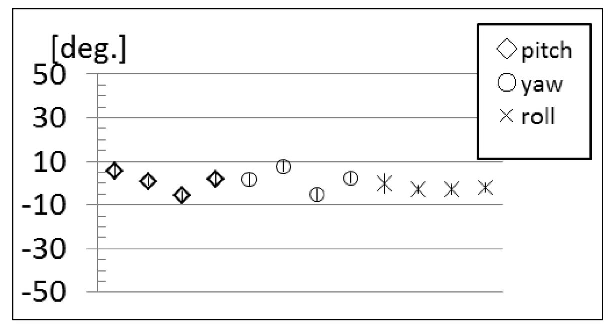

Figure 4: Average and deviation of the center position.

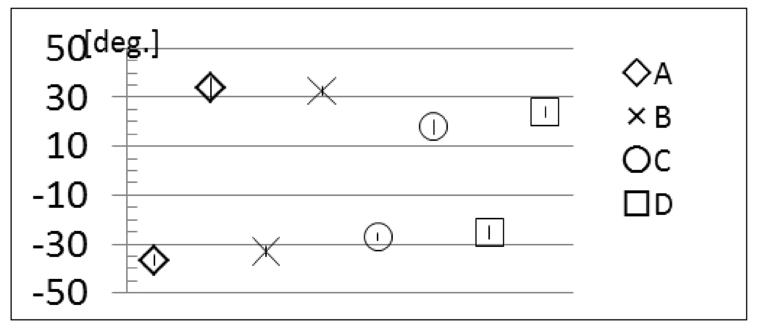

Figure 5: Average and deviation of roll when subject is tilting.

Figure 5 shows the average and deviation of the roll angle of the neck when four subjects tilt their necks. Deviations are small. The range is placed within +/-20° to 30° with symmetry for all subjects. This result indicates that 30° ± 7° is appropriate for the detection of the action “question.”

Figure 6 shows the average and deviation of yaw angles when four subjects swing their heads from side to side. Similar to Figure 4, unsymmetrical variation occurred in Figure 6 This variation is caused by the drift error of the magnetometer or the contact condition of Rift for each subject. Results show that the range 40° +/- 9° is appropriate to detect the “No” action. However, the problem of setting the center value remains.

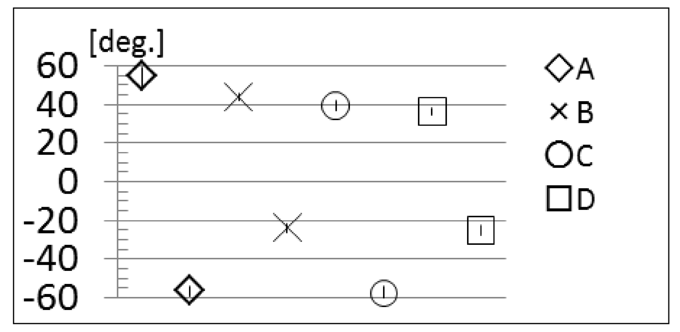

Figure 6: Average and deviation of yaw when a subject swings his or her head.

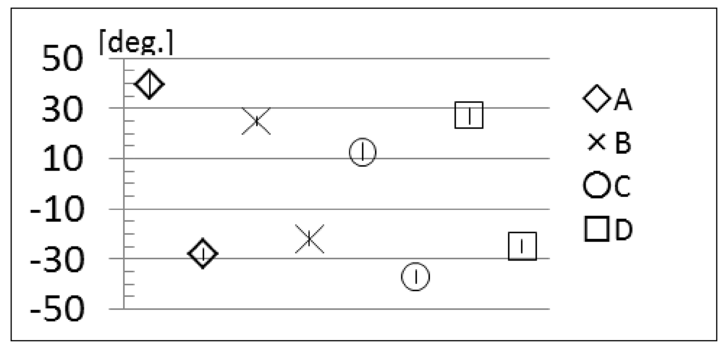

Figure 7: Average and deviation of pitch when a subject is nodding.

In the pitch for nodding, two subjects showed symmetry of +/-30° . However, half of the subjects showed an asymmetrical range (i.e., +40° to -30° and +10° to -35°). For the pitch direction of the jaws, around 18° of inter-subject variability (30%) exists. For the parietal direction, 27° of inter-subject variability (45%) exists. Therefore, inter-subject variability cannot be ignored. The physical limitations of the jaws, which have movable angles, are different. The range 30° ± 5° is appropriate to define “agree” or “back-channel feedback” detections. However, the motion contexts and frequency of nodding should still be considered.

CONCLUSION

We obtained experimental values from four subjects to define the “question,” “no,” and “nodding” actions.Constancy of head turning can be useful for neck gesture recognition to detect the “question” and “no” actions. Acceleration should be obtained to separate "agree" or "back-channel feedback" actions from other actions.