DRIVER POSTURE RECOGNITION FOR 360-DEGREE HOLOGRAPHIC MEDIA BROWSING

Ying-Wei Chuang[a], *Shih-Wei Sun*[b,c], *and Pao-Chi Chang*[a]

[a] Department of Communication Engineering, National Central University, Taiwan
[b] Department of New Media Art, Taipei National University of the Arts, Taiwan
[c] Computer Center, Taipei National University of the Arts, Taiwan

ABSTRACT

To interactively display scene photos in an in-car environment, we propose a 360° holographic media browsing system, which can recognize a driver's posture to trigger relative visual information to be displayed on a head-up display (HUD) in a car. Recognizing a driver's posture is a challenging issue, due to the unstable lighting conditions and occlusion issues in a limited space in a car. In this paper, we propose a driver posture recognition scheme for triggering the 360° media to be displayed on a HUD. The contributions of this paper are trifold: (1) a real-time driver posture recognition algorithm analyzed from the geometric relationship among multiple joints in a spatial-temporal domain, (2) a user study for recognizing postures with different level is provided for further user interface design, and (3) a 360° holographic media is displayed on a HUD based on a unity game engine.

Index Terms— posture recognition, driver, holographic media, 360

1. INTRODUCTION

Sharing panorama/360° photos and videos to a social network is more and more popular in the recent years. For example, sharing 360° photos [1] and 360° videos from 360° video capturing devices, e.g., Richo Theta 360° video camera [2] or a built-in 360° camera on a mobile device [3], is a popular behavior. Browsing for 360° Facebook photos from a mobile device can be interacted from a touch panel or a digital compass sensor on a mobile device. However, assuming that the 360° media can be projected to a head-up-display (HUD) [4], a better user interaction should be designed for an in-car environment. Motivated by browsing social media on a HUD [5], in this paper, we propose to browse 360° social media on a HUD using different postures behaved by a driver without any hand touching on a panel for efficient system response. When a car is driven to a target geo-location, triggered by the GPS-based sensing algorithm, a 360° media provided by social members is pushed to a driver, and can be interactively browsed by the proposed postures.

In this paper, as shown in Fig. 1, we propose an interactive in-car 360° holographic media browsing system to allow a driver to browse 360° multimedia from a holographic display in the car based on the skeletons obtained from a Kinect camera mounted in the car. The contribution of this paper is trifold: (a) we proposed an interactive holographic 360° media browsing system, (b) the proposed poses behaved by a driver are designed to keep the eyesight in the front direction of the car for further driving, and (c) the proposed posture recognition method can be operated in real-time for further triggering the 360° multimedia shown on the holographic display device in the car.

2. INTERACTIVE 360° MEDIA BROWSING BY A DRIVER'S POSTURES

In the proposed prototype, the 360° media captured from 360° cameras and mobile devices with panorama mode are projected on a transparent acrylic sheet, as shown by the bottom left part of Fig. 1. The interactive visual content is displayed from a mobile device (above the driver's steering wheel), and the user's postures can be captured from a Kinect camera mounted in the car (the top left part of Fig. 1). When a driver is driving, the user's hands are naturally put on the steering wheel. The visual contents captured from corresponding location provided from the social members are projected on the acrylic sheet (the object in the yellow dashed-rectangle of Fig. 1), with a holographic displaying user experience.

3. PROPOSED DRIVER POSTURE RECOGNITION

To trigger the 360° media to be projected on the HUD in the car, we propose four different postures to be recognized for triggering the visual content on the HUD. In the designing phase of the proposed system, as shown in Fig. 2, a volunteer is invited to sit on the driver's seat in a car, trying to keep the eyesight straight to the frontal direction of a car, and behave multiple postures for designing the triggering postures for the 360° holographic media displayed on the head up dis-

978-1-5386-2761-7

Fig. 1: The proposed interactive in-car 360° holographic media browsing system.

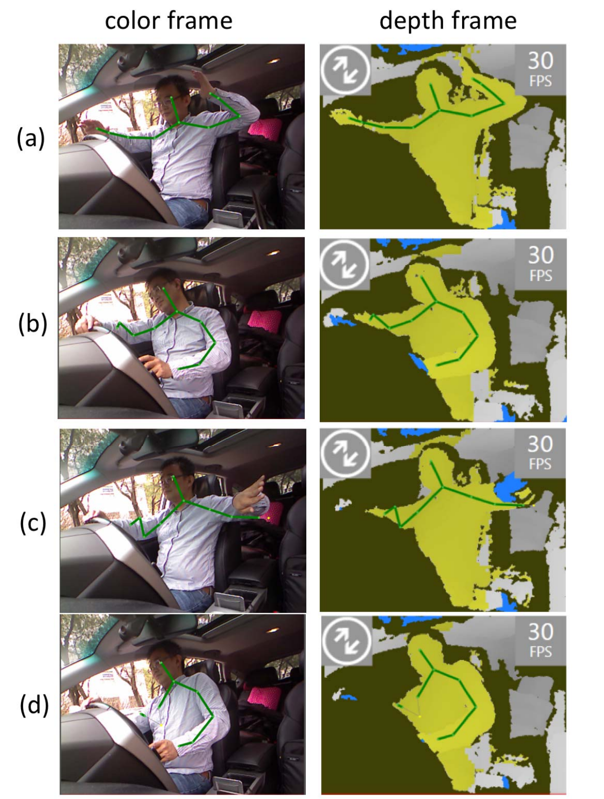


Fig. 2: A prototype is implemented in an in-car environment with a driver's poses: (a) head touching, (b) side head tilting, (c) hand waving, and (d) shoulder rotation.

play (HUD) in the car (the bottom part of Fig.1). As shown in the left part of Fig. 2, the user's skeletons (the green lines) are detected by Kinect sdk, from the corresponding depth frames (the right part of Fig. 2). According to the real car user experience, four driver's postures, i.e., head touching, side head tilting hand waving, and shoulder rotation, are decided as the target postures for triggering the relative holographic 360° media. Therefore, we propose four methods to recognize the above-mentioned postures.

3.1. Head Touching from a Hand

To detect the head touching event from a driver's hand, we propose to measure the Euclidean distance from a given head joint position $(x_t^{head}, y_t^{head}, z_t^{head})$ (in Fig. 3) in a 3D space to a detected right hand joint position $(x_t^{rh}, y_t^{rh}, z_t^{rh})$ in the same the frame measured at time t, when the calculated distance is less than given threshold touching h_t:

$$\sqrt{(x_t^{rh}-x_t^{head})^2, (y_t^{rh}-y_t^{head})^2, (z_t^{rh}-z_t^{head})^2} < h_t, \tag{1}$$

it means the distance from the right hand joint to the head joints is very close (even touched). Under this condition, a head touching event can be recognized to trigger the corresponding 360° multimedia contents to be displayed to a driver. The geometric relationship among the joints are depicted in Fig. 3.

3.2. Side Head Tilting

To detect a head tilting event, similar to Eq. (1), we propose to measure the Euclidean distance from a given right shoulder joint 3D position $(x_t^{rs}, y_t^{rs}, z_t^{rs})$ (in Fig. 4) to the head joint position $(x_t^{head}, y_t^{head}, z_t^{head})$ from the same frame:

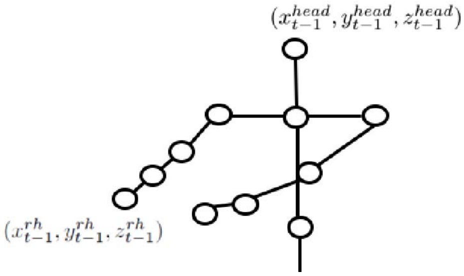
$$\sqrt{(x_t^{rs}-x_t^{head})^2, (y_t^{rs}-y_t^{head})^2, (z_t^{rs}-z_t^{head})^2} < h_s, \tag{2}$$

to detect the head tilting event to the right side when the detected distance is less than a side head tilting threshold h_s. In addition, by measuring the distance from the left shoulder joint position $(x_t^{ls}, y_t^{ls}, z_t^{ls})$ to the head joint position $(x_t^{head}, y_t^{head}, z_t^{head})$:

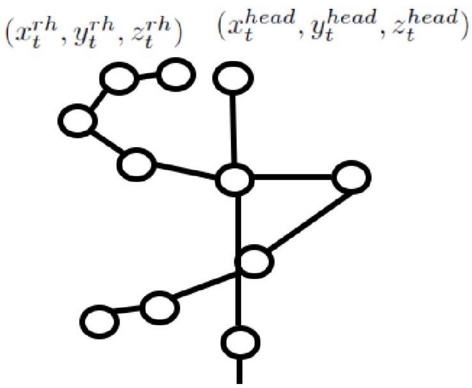
$$\sqrt{(x_t^{ls}-x_t^{head})^2, (y_t^{ls}-y_t^{head})^2, (z_t^{ls}-z_t^{head})^2} < h_s, \tag{3}$$

a head tilting event to the left side can be recognized. As shown in Fig. 4, when the geometric position of the head

978-1-5386-2761-7

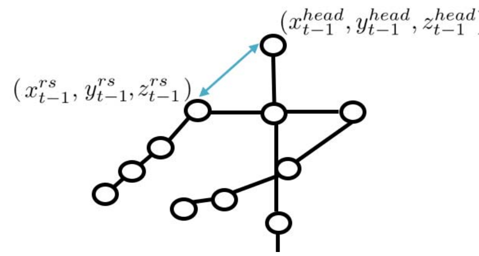


(a) The geometric relationship of the joints at time $t-1$.

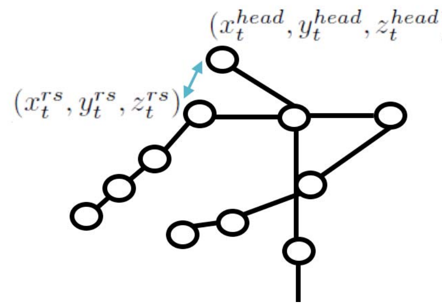


(b) The geometric relationship of the joints at time t.

Fig. 3: The posture of head touching.



(a) The geometric relationship of the joints at time $t-1$.



(b) The geometric relationship of the joints at time t.

Fig. 4: The posture of side head tilting.

joint and shoulder joints are moved very close, a side head tilting posture should be recognized.

3.3. Hand Waving

To detect a hand waving event, the motion vector is calculated from the right hand joint position $(x_t^{rh}, y_t^{rh}, z_t^{rh})$ (in Fig. 5) measured at time t to the same joint $(x_{t-1}^{rh}, y_{t-1}^{rh}, z_{t-1}^{rh})$ measured at the previous time instance $t-1$ by calculating:

$$\overrightarrow{\mu_t^{rh}} = (x_t^{rh} - x_{t-1}^{rh}, y_t^{rh} - y_{t-1}^{rh}, z_t^{rh} - z_{t-1}^{rh}), \quad (4)$$

as a motion vector. In addition, at frame t, a basis vector is defined from the 3D position of the left shoulder $(x_t^{ls}, y_t^{ls}, z_t^{ls})$ to the 3D position of the right shoulder $(x_t^{rs}, y_t^{rs}, z_t^{rs})$ at the current time instance t as:

$$\overrightarrow{\mu_t^{b}} = (x_t^{rs} - x_t^{ls}, y_t^{rs} - y_t^{ls}, z_t^{rs} - z_t^{ls}). \quad (5)$$

Moreover, an inner product operation from μ_t^{rh} to μ_t^b:

$$v = < \overrightarrow{\mu_t^{rh}}, \overrightarrow{\mu_t^{b}} >, \quad (6)$$

is used to measure the hand waving valid motion, as shown in Fig. 5(b). When the magnitude of the valid motion vector is larger than a threshold, i.e.,

$$v > h_v, \quad (7)$$

a hand waving posture can be recognized. In addition, under this condition, the sign of the horizontal direction of the the motion vector at the can be used to judge the right hand waving:

$$sgn(v), |v| > h_v. \quad (8)$$

The plus sign of Eq. (8) represents hand waving posture left direction, and the minus sign of Eq. (8) represents a hand waving posture to the right direction.

3.4. Shoulder Rotation

To keep the driver's eye sight to be concentrated at the front direction of a car, we propose to recognize a shoulder rotation posture for triggering 360° multimedia. Given a 3D position $(x_t^{cs}, y_t^{cs}, z_t^{cs})$ (in Fig. 6 (a)) of a center shoulder joint, a vector from the position of the head joint $(x_t^{head}, y_t^{head}, z_t^{head})$ to the center shoulder joint is:

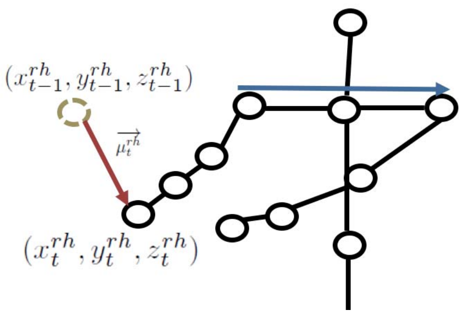
$$\overrightarrow{\mu_t^{hc}} = (x_t^{head} - x_t^{cs}, y_t^{head} - y_t^{cs}, z_t^{head} - z_t^{cs}), \quad (9)$$

, and another vector from the 3D position of the left shoulder joint $(x_t^{ls}, y_t^{ls}, z_t^{ls})$ to the center shoulder joint as:

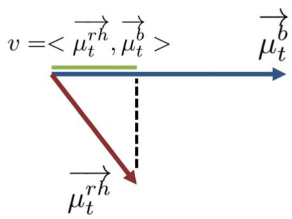
$$\overrightarrow{\mu_t^{lc}} = (x_t^{ls} - x_t^{cs}, y_t^{ls} - y_t^{cs}, z_t^{ls} - z_t^{cs}), \quad (10)$$

with a outer product operation for obtaining the normal vector as:

978-1-5386-2761-7

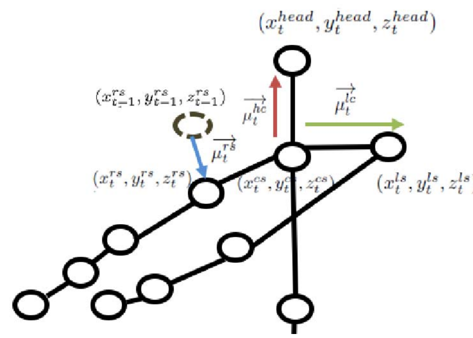


(a) The joints at different positions.

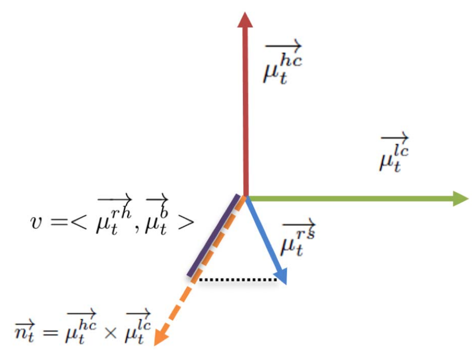


(b) The vector projection relationship between the vectors.

Fig. 5: The posture of hand Waving.



(a) The joints at different positions.



(b) The vector projection relationship between the vectors.

Fig. 6: Shoulder rotation

$$\overrightarrow{n_t} = \overrightarrow{\mu_t^{hc}} \times \overrightarrow{\mu_t^{lc}}. \tag{11}$$

The motion of the right shoulder is defined as:

$$\overrightarrow{\mu_t^{rs}} = (x_t^{rs} - x_{t-1}^{rs}, y_t^{rs} - y_{t-1}^{rs}, z_t^{rs} - z_{t-1}^{rs}), \tag{12}$$

and the projection from the motion to the normal vector is calculated by:

$$\mu_t^{sn} = <\overrightarrow{\mu_t^{rs}}, \overrightarrow{n_t}>, \tag{13}$$

to measure the valid amplitude of the shoulder rotation posture, as shown in Fig. 6(b). A larger value of the norm of Eq. (13), i.e., a larger μ_t^{sn} value represents a more rotation posture of a right shoulder.

4. EXPERIMENTAL RESULTS

4.1. The Testing Environment

Head touching, side head tilting, hand waving, and shoulder rotation postures are designed as the targets to be recognized in this paper. There were 10 users invited to test the recognition accuracy of the proposed method. 5 of the users are males and 5 of them are females, ranging from 22 years old to 24 years old. The postures are behaved by the users in a lab environment, as shown in Fig. 7. The proposed method is implemented in Unity 4.6.1, and the threshold settings of the propose method are empirically determined as: $h_t = 0.05$, $h_s = 0.09$, $h_v = 0.1$, and $u_t^{sn} > 0.05$.

4.2. Implementation for 360 ° Media Browsing

Based on the proposed pose recognition method, the 360 ° photos are integrated into a Unity environment, to be displayed in a virtual reality environment. We implemented a sphere object in the 3D environment, and the photos captured by a 360 ° camera, i.e., Ricoh theta s, are imported as the texture of the sphere. In addition, the displayed visual contents are rendered by a notebook computer, and the mobile device is connected to the notebook computer via a wireless local network. The operated real-time rendering visual contents are transmitted to the mobile devices, displayed on the built-in acrylic sheet in the car as a holographic projection at the head up display (HUD). The examples of the proposed 360° photos in Unity are shown by Fig. 8(a) and Fig. 9(a). When a horizontal motion of a user is recognized, as shown by Fig. 8(b)-(d), the visual content is displayed from right to left. On the other hand, when a diagonal poses, e.g. side head tilting, part of the content are correspondingly displayed by the motions of a user, as shown by Fig. 9(b)-(d).

4.3. Posture Recognition Results

The accurate recognition results (counted times from all users) are shown as the green bars in Fig. 10 and Fig. 11. The users are invited to behave the postures with three lev-

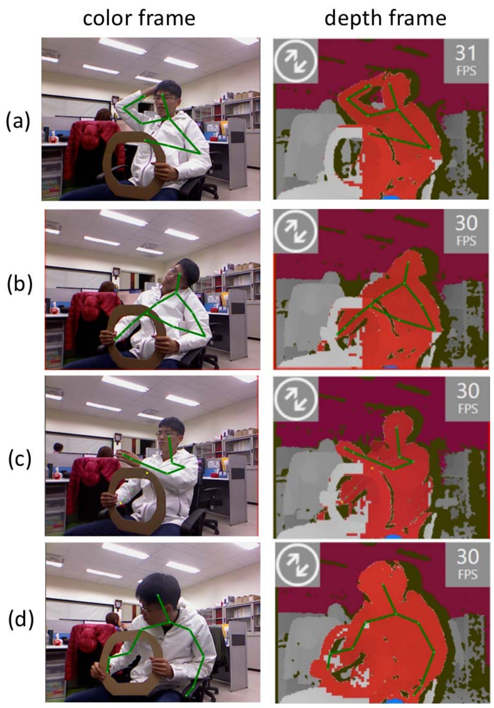


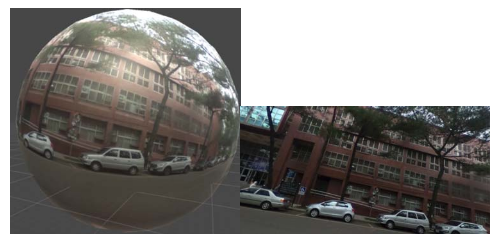
Fig. 7: A prototype is implemented in a lab environment with a driver's poses: (a) head touching, (b) side head tilting, (c) hand waving, and (d) shoulder rotation.

els, which are strong, mid, and weak levels. For example, a head touching from a hand (Fig. 10(a)), the accurate recognized postures are with strong posture level, i.e. the most left green bar. When users behave the posture as the weak level, most of the postures are determined as the unknown postures (the right most blue bar). Similar situations (green bars decreases from strong to weak) happens for the other postures. We should notice that, the weak hand waving postures are recognized as the wrong results (the red bar in Fig. 10(c)) and unknown results (the red bar in Fig. 10(c)), because of the too small motion from a user's hand. In addition, when a user's hand is moved to the left direction, the false recognition rate is increased (the red bars in Fig. 11(c)), even with strong level posture. The main reason is the occlusion situation happened to left side skeletons of a user, and the false recognition results happened in most of the cases.

(a) 360 ° image in unity. (b) browsing the right part.

(c) Browsing the mid part. (d) Browsing the right part.

Fig. 8: A captured scene on the road.



(a) 360 ° image in unity. (b) Browsing the left part.

(c) Browsing the mid part. (d) Browsing the right part.

Fig. 9: A captured scene in front of a building.

5. CONCLUSION

In conclusion, according to the user study and the experimental results, the proposed method can efficiently recognize a driver's posture in an in-car environment. The contributions of this paper include: (1) we analyzed the geometric relationship among multiple skeleton joints in a spatial-temporal space of a user to recognize postures, (2) to design the further natural user interface, a user study with different levels are tested, and (3) we projected a 360° image on the HUD that allows a driver can browse the real scene photos in a car. In the future, the proposed prototype will be integrated into a real in-car environment to develop further applications in a car.

6. REFERENCES

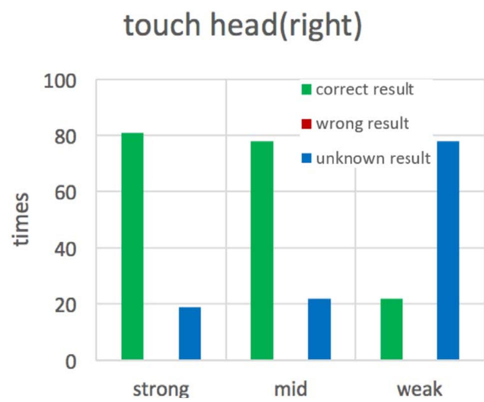
[1] "https://www.facebook.com/facebook360/," .

[2] "https://theta360.com/en/gallery/," .

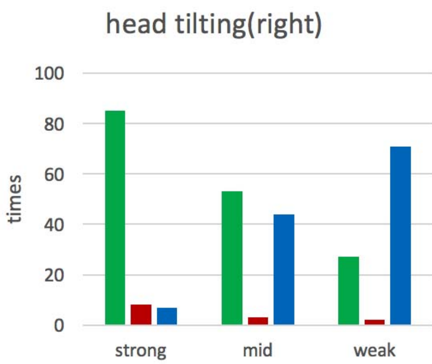
[3] Liu et al., "A 360-degree 4k x 2k panoramic video recording over smart-phones," *IEEE ICME*, 2016.

[4] Okumura, "Novel 3d head-up-display for automobile," *ACM SIGGRAPH Asia*, 2015.

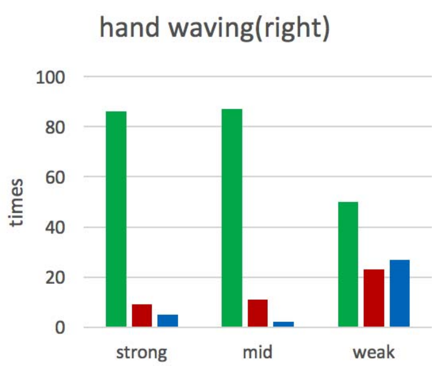
[5] Chang et al., "A social media based real scene navigation system with a holographic projection on a hud," *ACM UbiComp*, 2015.



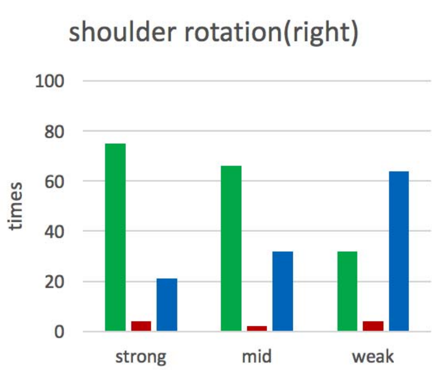
(a) head touching from a hand.



(b) Side head tilting.

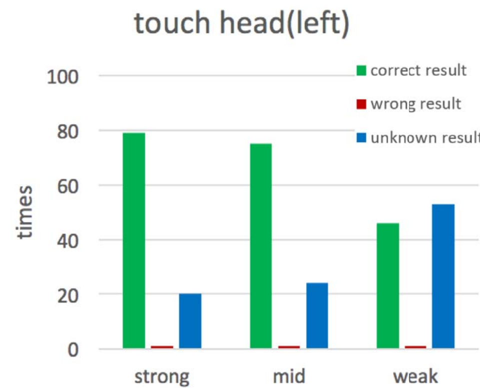


(c) Hand waving.

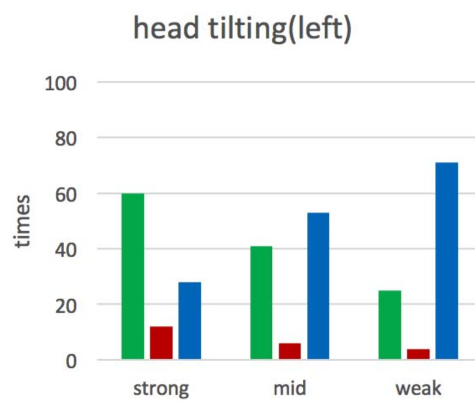


(d) Shoulder rotation.

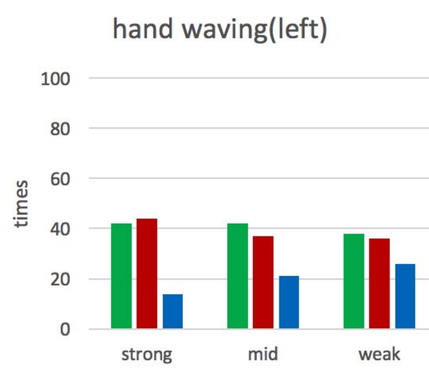
Fig. 10: Pose to the right direction.



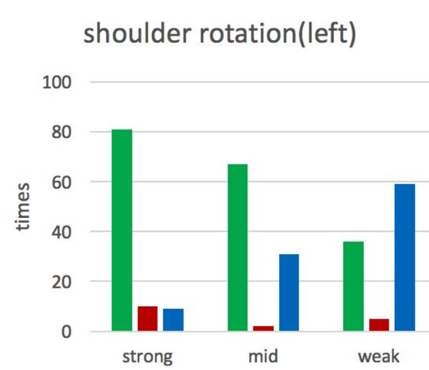
(a) head touching from a hand.



(b) Side head tilting.



(c) Hand waving.



(d) Shoulder rotation.

Fig. 11: Pose to the left direction.