COQ 302
Eyelid Gestures for People with Motor Impairments.

Kushagra Rode (2020CS10354)

Scribing presentation of Shreyansh Singh and Atharv Dabli

November 2022

1 Abstract

This paper discusses the prospect of using eyelid gestures to control smartphones. This is particularly very useful for people with motor impairments. Currently there are some technologies to assist people with motor impairments like eye-trackers, brain-computer interfaces, and mechanical devices (e.g., joysticks, trackballs), but these devices are intrusive, expensive for common people and limited in their functionality, but the smartphone is easily accessible to most people and it can perform a much diverse range of actions.

Currently, there are some existing technologies that use eyelid interaction but it mostly focuses on eyeball movement and blinks. This research paper discusses the use different eyelid states which could be easily identified.

2 Eyelid Gesture Design

Humans have two eyelids and considering them being open or closed we get four eyelid states. Namely, both eyes open, right eye closed, left eye closed and both eyes closed (We can also have a half-closed eyelid state, however, this is harder to detect and sustain by humans.) Furthermore, we can also divide the duration of the eyelid closed as short duration and long duration. Here short duration means time is taken to intentionally blink the eye (more than the time taken in a spontaneous blink) and long duration means closing the eye and then sustaining it for some time. Combining them, we can actually have infinite eyelid gestures. Both eyes being open is used as a gesture delimiter since it is the state our eyes are in when we are awake.

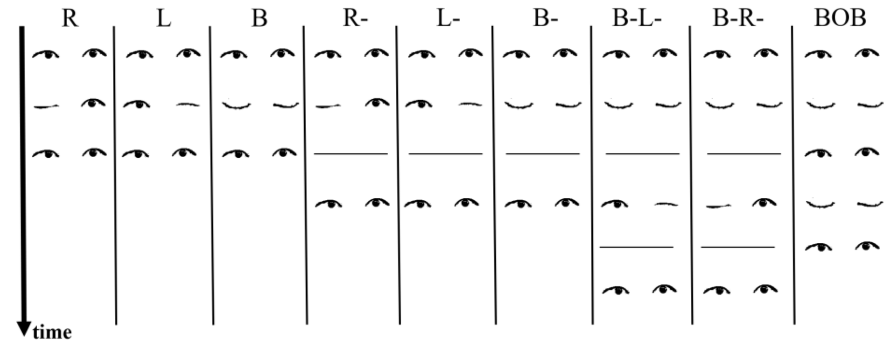


Figure 1: Eyelid Gestures

The algorithm detects nine of these eyelid gestures. These eyelid gestures are R, L, B, R-, L-, B-, B-L-, B-R-, BOB. Here R means right eye closed, L means left eye closed, B means both eyes closed, and O means both eyes open. - represents that it is to be done for a long duration (for eg. B-L- means close both eyes for a long duration, then close left eye for long duration).

3 Recognition algorithm

The front camera of the smartphone takes continuous images (at a rate of 30 frames per second) and these images are fed to Google Mobile Vision API, which generates a stream of probability pairs of each eye being open (P_L, P_R). These probability pairs are then classified into both eyes open or any eye closed using an SVM classifier. This separates the probability pairs for a given gesture. Now, to classify a gesture into long and short duration, another SVM classifier is used. For short duration, 50 sample pairs of (P_L, P_R) and for long duration, 100 samples of (P_L, P_R) are generated. Then an SVM classifies short-duration gestures into R, L, B, BOB and long-duration are classified into R-, L-, B-, B-L-, B-R- by another SVM classifier. All SVM classifiers have used the Radial Basis Function kernel. Default parameters of scikit-learn library were used to implement the SVM classifier.

4 Testing of Algorithm

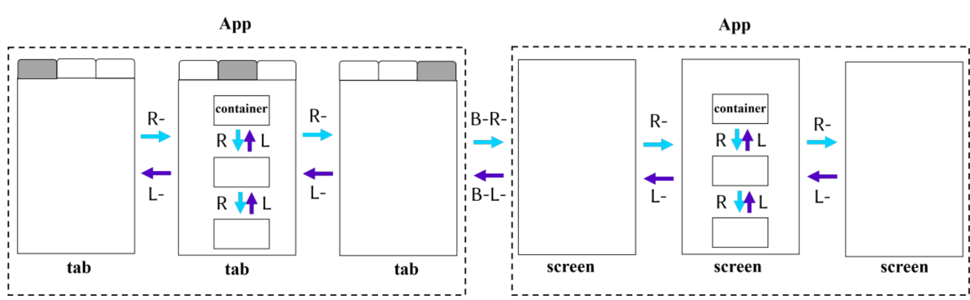
The algorithm was trained and tested on 12 healthy individuals and four people with severe motor impairments. There were people of both genders and some very glasses; one wore contact lenses, and while testing rooms and position of the smartphone were changed. This was done to introduce variation in the data collected. Healthy individuals kept their smartphones in their hands while both sitting and standing. People with motor impairments, they sat on their wheelchairs, and their phone was kept at their eye level using a tripod stand.

First, the app presented them with different eyelid states randomly. On collecting data for all four eyelid states, the algorithm was trained on this data. Now, data is collected for nine eyelid gestures. The evaluation app generated a target gesture to be performed. Each gesture is presented ten times, and gestures are presented randomly.

For eyelid state evaluation, the overall accuracy is 0.92, and for eyelid gestures used, dependent and used independent evaluation is done. In user-dependent evaluation, five samples are used to train the data, and rest five were used for testing the algorithm. The overall accuracy, in this case, is 0.76. In user-independent evaluation, the data of the rest users were used to train the algorithm, and the data of that user is used to test the algorithm. The overall accuracy, in this case, is 0.70.

5 Interacting with mobile apps

In a smartphone, there are three types of navigation. First is switching between the apps, since it is a rare task to tough eyelid gesture (B-L-, B-R-) assigned to it. Then there is navigation between the screens of an app, the eyelid gesture assigned to it is R-, L- and finally, for navigation between the containers of the app, which is the most common task, is assigned R, L. Along with the navigation, BOB is used for selecting an item.



6 Future Works

The half-closed eyelid state could also be used to make different eyelid gestures, which will increase the state space. Currently, short and long duration are only used; along with these two long-long duration could also be used. The user who tested it, suggested that the duration classifier should be trained for each user. The eyelid gestures should be modifiable according to the needs of the user. These eyelid gestures could also be used to control other electronic devices like TVs, ACs and tablets.