Report

Facial landmark and movement detection

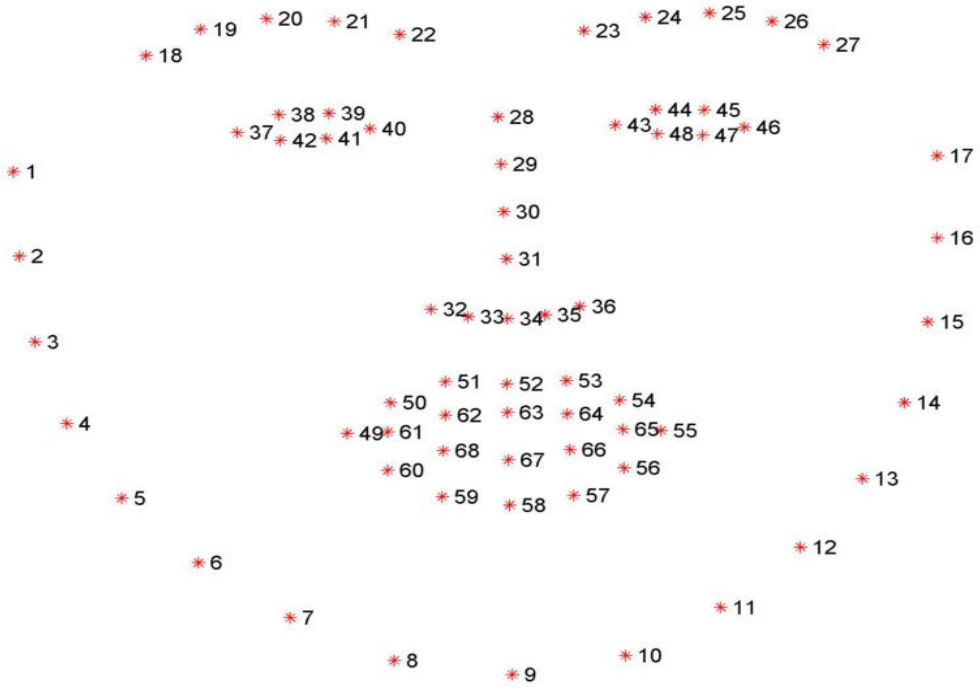
Justification

The deaf community is composed by the people who is hearing impaired. Because of this disability they cannot learn to speak, so the deaf uses another language called “sign language” which is a body language. Even though they created this language they still have problems communicating with the “listeners” because the don't know the language, which is difficult to learn. To reach a solution, many developers have created software and hardware that can interpret the signs and turn them into text or sound. The listeners are happy with theses creations, but the deaf disagree because this tools does not give a proper sign language interpretation. This software and hardware interpreters focus in the movement of hands to interpret the sign, but this language is a body language that also involves face and torso gestures as explain Lance Forshay, of the department of linguistics of the University of Washington, in his letter “An Open Letter to UW Office of News and Information on SignAloud Project in Media.” in which he also add the difference between to words an only be the raise of an eyebrow. Given this, the state of the art is incomplete, therefore this challenge in focus on created an application that can, in real time, find facial landmarks and determine certain movements focusing specifically on the eyebrows for sign language interpretation. [1]

Process

For achieve this first the computer must determine a face or faces in front of the camera, and then move to the eyebrows; for which the used of supervised learning was needed to determine all different kind of faces. For the machine learning algorithms the Dlib library was used. For face detection the Dlib library uses an HOG (histogram of oriented gradient) + SVM (support vector machine). This detector uses comes pre-trained, the training process involved a dataset of images, and an xml with the name of each of those pictures and the x and y coordinates, width, and height of a handmade box that covers the face area in the image. For landmark detection the Dlib library also comes with pre-trained shape predictor. For the shape predictor to work, first the face must be found, after the face detector acquires a group of faces from the image, the landmark detector for the mouth, nose, eyes, jaw, and eyebrows using regression trees and gradient boosting algorithm. The implementation of this landmark detector is based on Vahid Kazemi and Josephine Sullivan project, One Millisecond Face Alignment with an Ensemble of Regression Trees. For the training of the landmark detector the process is similar to the face detector. The dataset involves an xml with the name of the pictures, the box surround the face and the x and y coordinates of each point that involves the landmark, these are 68 points to label the face parts already mention, as it can be seen in the illustration 1. [2, 3, 4]

Illustration 1

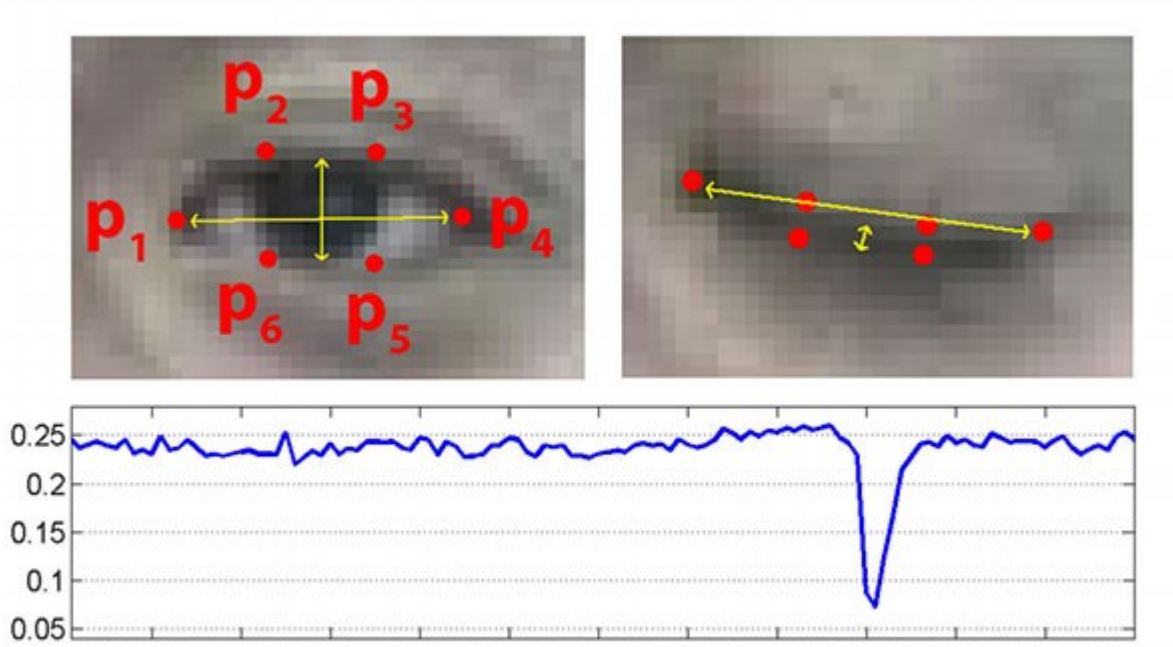


To actually determine the position of the eyebrow, whether is raised, normal or lowered, the algorithm was based on Tereza Soukupova and Jan Cech paper Real-Time Eye Blink Detection using Facial Landmarks. To determine blink the use a formula to calculate if the eyes is closed or opened called eye aspect ratio which involves using the points of the landmarks and divide the distance from the two vertical lines by twice the horizontal section of the eyes. This formula gives a real number from 0 to 0.3. When the eye is opened the is at its highest and when is closed it reaches 0. There is a graphic explanation of this process in illustrations 2, and 3. [5]

Illustration 2

$$\mathrm{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

Illustration 3



Based on this formula the position of the eyebrows was calculate by dividing the distance between the inner edge point of each eyebrow and the top of the nose by the distance between the center point of each eyebrow and the top of the nose. This formula on gave results from 0.38 to 0.51 being 0.38 when the eyebrows where lowered and the 0.51 when they where raised

Results

After completing the code and testing the following results where reached:

- The numerical value **range** may vary according the person's face **form** in the case of the eyebrows since some persons gave higher or lower eyebrows which will cause to formula to have a different range and therefore return false positives
- Shape predictor not too accurate to determine the eyebrows when moved to quickly

- Face position is too important

    a. 2D camera is not aware of depth perception

        since it is a normal camera, there is no depth information so the value calculated for the eyebrows in the formula will vary just by moving the face up and down will not moving the eyebrows.

- The background light affects the detection of the face and it's landmarks for proper you must be in front of a wall with low reflection or make sure the light source is not behind you.

- Can easily detect body language expression but not micro expression

    The result value from the formula varies with an plus/minus 0.2 error when the user is not moving the face, therefore it will not suit for micro expressions of the face which are extremely tiny.

References

1 University of Washington, department of Linguistics. Lance Forshay. Abril 2016. An Open Letter to UW Office of News and Information on SignAloud Project in Media. Online resource, retrived on march 5 of 2018 https://catalyst.uw.edu/workspace/file/download/5c706e255e89d3eea9cc0a7d6ac1fb3e89c44cfdc54774630e7637ef0eb47f94?inline=1

2 One Millisecond Face Alignment with an Ensemble of Regression Trees by Vahid Kazemi and Josephine Sullivan, CVPR 2014. Online resource, retrived on may 1 of 2018 https://pdfs.semanticscholar.org/d78b/6a5b0dcaa81b1faea5fb0000045a62513567.pdf

3 Dlib face detector Online resource, retrived on may 1 of 2018
- a http://dlib.net/ml.html
- b http://dlib.net/face_detector.py.html
- c http://dlib.net/train_object_detector.py.html

4 Dlib face landmark detector Online resource, retrived on may 1 of 2018
   a http://dlib.net/ml.html
   b http://dlib.net/face_landmark_detection.py.html
   c http://dlib.net/train_shape_predictor.py.html

5 Real-Time Eye Blink Detection using Facial Landmarks by Tereza Soukupova and Jan Cech. Online resource, retrived on may 1 of 2018 https://dspace.cvut.cz/bitstream/handle/10467/64839/F3-DP-2016-Soukupova-Tereza-SOUKUPOVA_DP_2016.pdf