

The Smile Detector

The aim of this study is to identify human smiles in **real-time**. Unlike identifying arousal, which is relatively easy (can be achieved by measuring skin conductivity, blood volume pressure, etc.), identifying valence is a much harder task. Nevertheless, identifying and measuring valence can be worthwhile. It can be used in various applications, such as affective software to assess user's state, in distance learning systems, where the teacher cannot see each student reaction, and so forth. Our proposed smile detector, combined with other detectors (such as eyebrow gestures and head nod/shake detectors) can provide a valence measure, or even boost us towards a real emotion identifier.

Background

The smile detector is based upon Ashish Kapoor's system [4]. The current system uses IBM Blue Eyes camera (<http://www.almaden.ibm.com/cs/blueeyes>) for real-time detection of pupils, eyes and eyebrows [4], and also head nods and shakes [3]. The pupils are detected using an algorithm similar to Morimoto [2]. Eyes and eyebrows are identified using Eigen-points [1].

Technical stages

1. Framing the mouth area -

The mouth center is located under the eyes, at a certain distance below the lines that connects the pupils. First, this distance is calculated in respect to the distance between the pupils. The line connecting the pupils and the center of the lips is marked in yellow (see Figure 1). Second, we take into consideration head tilts when framing the mouth. In this case, the lips frame will not be lined below the eyes, but positioned according to the line connecting the pupils. See Figure 2 for a tilted face.

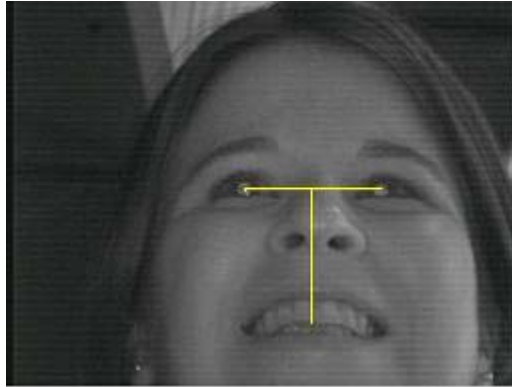


Figure 1: Locating the mouth center

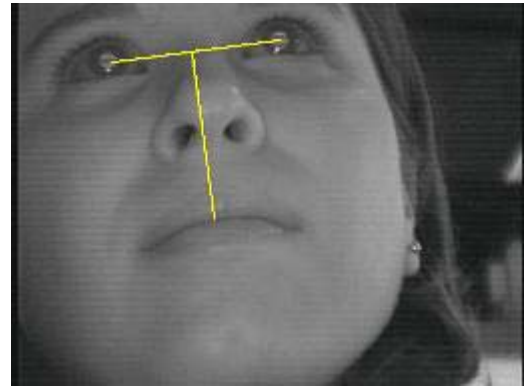


Figure 2: tilted face

3.

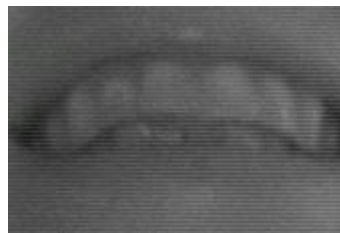
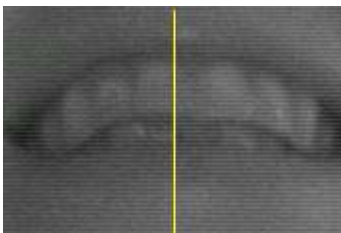
The size of the mouth frame is calculated relatively to the distance between the eyes. This is in order to take consideration the: 1) face size, 2) distance from the camera, 3) face looking away. For instance, the face in Figure 2 is closer to the camera, and therefore, the mouth box should be larger to fit the mouth. Another example (see Figure 3), shows the face slightly facing the side.



Figure 3: Face looking away

4. Identifying smile using a side cut of the mouth -

We analyze the wave form vertically to identify the pattern of the smile. Figure 4 presents a smile and its correlating waveform graph. The smile pattern is different than a closed mouth. Figure 5 shows a typical graphs of a smile (left) and no smile (right). A non smiling mouth typically has one minimum (black) where the lips close, a smiling mouth has two minimums, one for each lip area (the shade created where the lip entering the mouth).



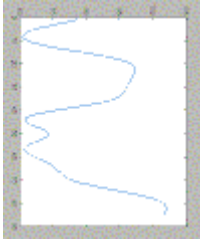


Figure 4: Mouth area with its wave form

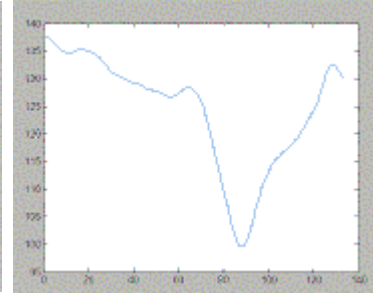
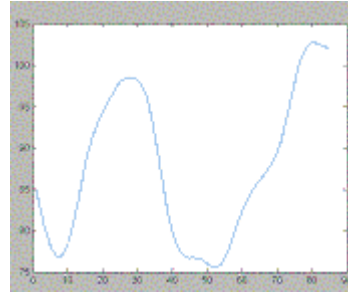


Figure 5: Smile vs. no smile

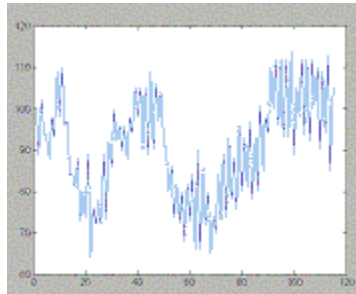


Figure 6: Noisy data

The data is quite noisy as presented in Figure 6. A low pass filter is therefore used. The strength of the filter is determined by the size of the mouth area.

Technologies

Current implementation uses a c program to capture images and locate the eyes. The files are downloaded to a Matlab program for the research and analysis. The next version will implement the whole system in c.

Evaluation

The algorithm was tested on 7 subjects. 1 had a beard and 2 were too far from the camera. An analysis of the "clean" data included 4 subjects with 106 images, from which 40 were smiles. There were 4 errors (3 missed smiles, and 1 false identification). This is 96% accuracy, however, half of the subjects and the images belongs to the training set. A more comprehensive evaluation is still required.

Problems and future work

- Resolution

Small faces, combined with a certain distance from the camera, which conclude in a mouth area matrix smaller than 100x150 pixels is insufficient for classification. Using a higher resolution camera is required to overcome this problem.

- Further training

Further training of the system is required to adapt its parameters, such as, mouth location, mouth box size, level of filter, etc.

- Beards

The current system does not support classification of bearded people. A further development of an algorithm for this purpose is required.

- Comprehensive Study

A comprehensive study with a large population is needed to evaluate the reliability of the system.

- Implementation in c

The next step is an implementation of the system in C for real time support.

References

1. M. Covell, Eigen-points. Proceedings of International Conference Image Processing, September 1996.
2. C. Morimoto, D. Koons, A. Amir, and M. Flickner, Pupil Detection and Tracking using Multiple Light Sources, Technical report, IBM Almaden Research Center, 1998.
3. A. Kapoor and R.W. Picard, A Real-Time Head Nod and Shake Detector, Workshop on Perceptive User Interfaces, Orlando FL, 2001.
4. A. Kapoor and R.W. Picard, Real-Time, Fully Automatic Upper Facial Feature Tracking, To Appear in the Proceedings of The 5th International Conference on Automatic Face and Gesture Recognition, 2002.