Stress Detection Using Techniques From Computer Vision

Prithvi Krishna Gattamaneni and John Ryan

I. Problem Statement

Research has shown[1] that staying in a sitting position while working at a desk for long hours can contribute to health problems. Furthermore, a link has been established[2] between extended inactivity and anxiety. In response to this problem, doctors and employers alike have recommended hourly breaks from sitting and staring at the computer, and that employees take five minutes to walk around and stretch their muscles. The problem is that despite this effect being well-known, most workers do not take action to reduce their level of inactivity at the workplace, and the phenomena continues to affect white collar workers worldwide.

II. Has this problem been solved?

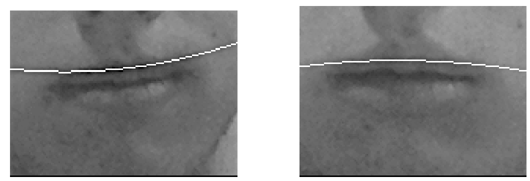
Nobody seems to be doing this out of the box right now. There is a lot of research going on in this field, but it does not seem like there is a readily available consumer product. While there are fitness products to monitor overall physical activity over the course of a day there does not seem to be a product to monitor a persons activity at a desk. Besides the recently popular trend of standing desks (which are inconvenient and expensive) there do not seem to be any software solutions to motivate users to get physical movement.

III. Why hasn't it been solved?

From past experience, many firms send their employees a mail once a year with some Best Practices regarding desk posture, and taking a walk every 30-60 mins. But they do no enforce this, nor do they have a way to enforce this. Many of the workstations today have webcams which are hardly ever used and there is no product that firms can readily deploy for their employees. Another motivation to solve this problem is that the average CPU usage for a desktop is only around 8% [3] and monitoring a persons facial expressions should not cause any overloads on the CPU.

IV. What's our magic weapon?

The magic of our solution to the problem is that it takes advantage of and combines two aspects of modern technology: the ubiquitousness of laptop webcams, and the development of simple, portable, and efficient computer vision software (such as OpenCV, FFT image convolution), and hardware such as GPU present in most modern computers.



(a) Normal (b) Tired

Fig. 1: By examining the curvature of the user's lips, the program can detect drooping, a sign of tiredness.

A. Background

Our approach to solving this problem is based in the principle that human emotion is largely conveyed through the eyes and the mouth. Some examples include a smile indicating happiness, raised eyebrows indicating surprise, a furrowed brow indicates anger, a frown indicates discontent, and more. With this in mind, we have the application begin by identifying the eyes and the mouth of a face in a video input, by using Haar wavelets for cascade classification (see section B).

Inspired by the concept of "eigenfaces" and the use of PCA for expression categorization present in the literature[4], we began by having our application convert the images of the facial features to high dimensional vectors based on its convolutions with several Morlet Wavelets. However, we ultimately discarded this approach; the expression of stress or tiredness which we want the application to detect is too subtle to be effectively distinguished from a normal face in a high dimensional vector space. The application of PCA in facial expression recognition is more effective when used to detect more strongly varying emotions, as in [5].

B. Our approach

The application begins by finding the eyes and the mouth in the image using cascade classification. This approach involves convolving the image with Haar wavelets known to distinguish facial features well, and choosing the two best guesses for the eyes, and the best guess for the mouth. We made this approach more accurate by specifying minimums sizes for the windows (since we know the user will be within a certain distance range from the webcam).

Once the mouth image is identified, it is convolved with a Morlet wavelet to detect vertical gradient. For each column of the result of the convolution, the points of highest response are recorded (these should be points between the lips), and a quadratic of best fit is calculated. In other words, the curvature of the lips is calculated - negative curvature is

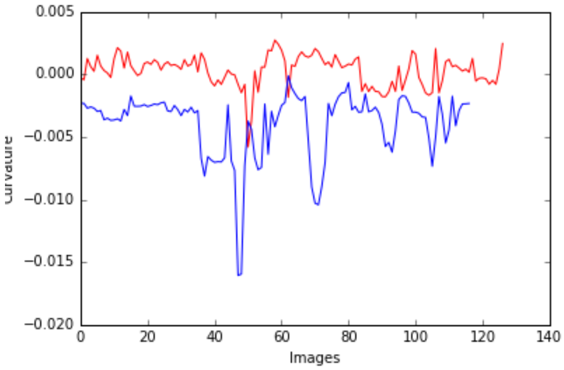


Fig. 2: Curvatures in the training images (blue is normal, red is stressed). The threshold -0.0015 results in both precision and recall exceeding 95%.

a smile, and positive curvature is a frown (or a drooping mouth). Figure 1 displays the calculated parabolas for two images.

The reader may wonder why a smile is negative curvature. This is a result of the origin being in the top left corner of an image, and, in our implementation, the rows corresponding to different positive y-coordinates. Thus, a conceptual overlaid Cartesian grid would be flipped vertically.

Once the curvature of the lips is found, it is compared to a learned threshold (see Figure 2). To account for error, we do not output that the user should go have a walk unless they are found to be stressed by the application in at least eight of the previous ten frames. For each stressed frame, the user earns a stress point, and stress points are reset every ten frames. If the user ever has 8 points, they are advised to take a break.

As of right now, the eyes are found, but not processed. Future work should include detecting drooping of the eyes or eyebrows in a similar fashion. The authors advise that PCA is probably not a good approach for this, for the reasons described in the beginning of this section.

V. Does it work? Does it scale?

This application has been tested only on the authors, but has proven to be reasonably successful. The scalability of the product is boosted by the ubiquity of web cams on workers' computers, and is as easy to distribute as any software package. Furthermore, more learning can be done to improve the functionality of the classification - the stress points system can be tweaked, and the thresholds improved, or more complicated polynomials used for curvature. Opportunities for learning include deployment, as well as various facial expression databases, such as the CMU Extended Cohn-Kanade database. The code is available at https://github.com/JohnPaulRyan24/Stress-Test and a video demonstration can be found at https://www.youtube.com/watch?v=WGSRE9fRaE0.

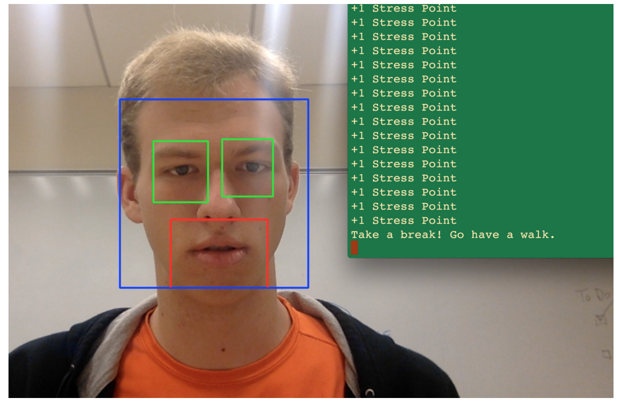


Fig. 3: Stress detected in the video demonstration.

References

[1] Hamilton, M.T., Healy, G.N., Dunstan, D.W., Zderic, T.W., and Owen, N. 2008. Too little exercise and too much sitting: inactivity physiology and the need for new recommendations on sedentary behavior. Curr. Cardiovasc. Risk Rep. 2(4): 292298. doi:10.1007/s12170-008-0054-8.

[2] http://www.ncbi.nlm.nih.gov/pubmed/23637169 (get citation)

[3] M. Kazandjieva, et al., Measuring and analyzing the energy use of enterprise computing systems, Sustain. Comput.: Inform. Syst. (2013)

[4] A principal component analysis of facial expressions. / Calder, AJ; Burton, AM; Miller, P; Young, AW; Akamatsu, S. In: Vision Research, Vol. 41, No. 9, 01.04.2001, p. 1179-1208.

[5] Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. PN Belhumeur, JP Hespanha, DJ Kriegman Pattern Analysis and Machine Intelligence, IEEE Transactions on 19 (7), 711-720