

Face and Smile Detection Using Symmetry Measures and Line
Detection

Amy Wibowo

May 1, 2008

1 Objective

The goal of this project is to implement a system to detect the mouth from a head shot of a single person and determine whether the person is smiling, frowning, or neither.

There are many possible applications for smile detection. A social robot that can recognize facial expressions as cues of emotional state can adapt its behavior towards a person accordingly. Expression detection can help image search engines with queries that have emotion keywords like “gloomy”. A smile detector could be used to control facial expressions of avatars for online games or communities.

2 Methodology

A symmetry measure was developed by Reisfeld, Wolfson, and Yeshurun in 1990 [1] as a way for finding points of interest in photos. Inspiration for the measure came from psychophysical evidence that the human visual system has a very strong sense of symmetry and quickly notices symmetric objects. The symmetry measure places high importance on contrast so that unchanging surfaces, which are highly symmetric but have no contrast, are not identified as interesting points. The symmetry measure can be used for detecting facial features, because the eyes, nose, and mouth are all symmetric. It has been used as a first step of face identification by the developers of the measure and others [2] [3].

If we denote the gradient at each point p_k as ∇p_k , it is useful to compute a modified magnitude and phase measure of this gradient (r_k, θ_k) , where $r_k = \log(1 + \|\nabla p_k\|)$ and $\theta_k = \arctan(\frac{\partial}{\partial y} p_k / \frac{\partial}{\partial x} p_k)$.

In order to measure the symmetry about a point p , we want to look at pairs of points (p_i, p_j) around p_k that have p_k as a midpoint. When we care most about the symmetry in the direction that makes an angle ψ with the horizon, we only need to look at the pairs of points in the set $\Gamma(p, \psi)$ that satisfy:

$$\Gamma(p, \psi) = \left\{ (i, j) \mid \frac{p_i + p_j}{2} = p, \frac{\theta_i + \theta_j}{2} = \psi \right\}$$

For each point p in an image, in the direction ψ we compute the symmetry measure $S_\sigma(p, \psi)$ over the set of pair of points in $\Gamma(p, \psi)$:

$$S_\sigma(p, \psi) = \sum_{(i,j) \in \Gamma(p,\psi)} D_\sigma(i, j) P(i, j) r_i r_j$$

ψ can also be thought of as one of a number of discrete angle bins. Each component of $S_\sigma(p, \psi)$ is discussed below.

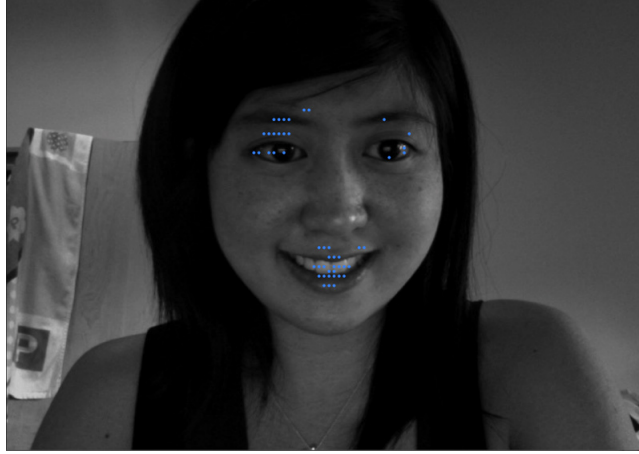


Figure 1: Points in the image found to have a high symmetry measure.

$D_\sigma(i, j) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\|p_i - p_j\|}{2\sigma}}$: This operand is a distance weight function that decreases as the distance between two points increases. This gives the symmetry measure a local nature. σ is a parameter than can be adjusted as necessary.

$P(i, j) = (1 - \cos(\theta_i + \theta_j - 2\alpha_{ij}))(1 - \cos(\theta_i - \theta_j))$: α_{ij} is the counterclockwise angle that the line through points p_i and p_j makes with the horizontal. This operand is at it's maximum when the gradients at p_i and p_j are oriented directly towards each other or away from each other. The second half of the operand discounts the case that occurs on a straight edge, which is not considered interesting.

3 Implementation

For many possible applications of a smile detecting system, webcamra input seems the most useful choice. Upon capture of a still image by button press, the RGB information of the image is converted to grayscale with the weights .3 for red, .59 for green, and .11 for blue. Then gradient ∇p_k is computed for each pixel of the image by convolving the image with the standard first order approximation filter for the gradient. The symmetry measure for each pixel is computed using the gradient and the algorithm described in the previous section. In order to capture symmetry along the horizontal direction, ψ is set to include the angle bins between $\frac{\pi}{16}$ and $-\frac{\pi}{16}$ and $\frac{15\pi}{16}$ to $\frac{17\pi}{8}$. Pixels with a symmetry measure above a certain threshold are flagged as being potential facial features.

Groups of pixels representing potential facial features tend to clump around the eyes, mouth, and symmetric objects in the picture background. Figure 1 shows points that have been flagged as having a high symmetry measure. In order to have just one central point per facial feature, points within a certain distance



Figure 2: Potential facial features.

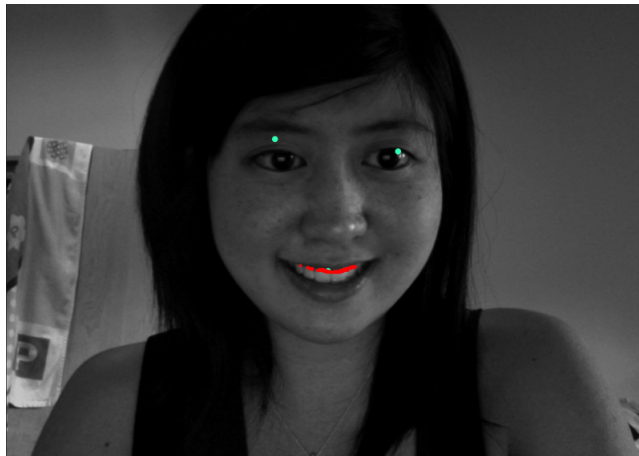


Figure 3: Line detection of mouth indicates a smile.

of each other are considered part of the same feature, and the centroid for each potential feature is computed. Figure 1 shows these computed centroids.

Noses, though symmetric, were not detected by this implementation of the symmetry measure, perhaps because of the lower contrast in a nose, compared to the eyes and mouth. A test is applied to groups of three potential features at a time to see if they fit the criteria for being two eyes and a mouth. Eyes will be two nearby points with similar y coordinates, and the mouth will be another nearby point whose y coordinate is a parametrized amount lower than that of the eyes and whose x coordinate is similar to that of the midpoint between the eyes.

When a mouth is identified, $\frac{\partial}{\partial y}$ is examined within a rectangle around the mouth point. The points for which $\frac{\partial}{\partial y}$ is higher than a certain threshold are assumed to represent the mouth. A mouth is classified as

smiling if the points with the most extreme x values are closer to the mouth points with the highest y value, frowning if closer to the lowest y value, and neutral otherwise. Figure 3 shows a detected mouth to be smiling.

4 Results and Further Thoughts

Various parameters needed to be adjusted during testing, including the thresholds for the symmetry measure and the threshold for mouth line detection. Eight images were used for testing, featuring three different subjects, two male and one female. One subject wore glasses. Out of the eight test images, in seven cases the eyes and mouth were correctly identified, and of those, five had the expression correctly identified.

Each incorrect identification could benefit from additional morphological processing to eliminate island regions in the mouth line detection, similar to what was done by Li and Kobatake [3]. Furthermore, other facial cues like eyebrow angle could be taken into account for more reliable expression identification. Also the implementation takes several seconds to process an image, making it not yet practical for streaming video emotion detection.

References

- [1] Reisfeld, Daniel; Wolfson, Haim and Yeshurunt, Yehezkel *Detection of Interest Points Using Symmetry* 1990: Proceedings, Third International Conference on Computer Vision.
- [2] Reisfeld, Daniel and Yeshurunt, Yehezkel *Robust Detection of Facial Features by Generalized Symmetry* 1992: Conference A: Computer Vision and Applications, Proceedings., 11th IAPR International Conference on Pattern Recognition, Vol.I.
- [3] Li, Yuanzhong and Kobatake, Hidefumi *Extraction Of Facial Sketch Image Based On Morphological Processing* 1997: International Conference on Image Processing (ICIP'97) - Volume 3 p. 316.