Determining Mood from Facial Expression

Matthew Wang (mmwang@stanford.edu), Spencer Yee (spencery@stanford.edu)
Stanford University

OVERVIEW

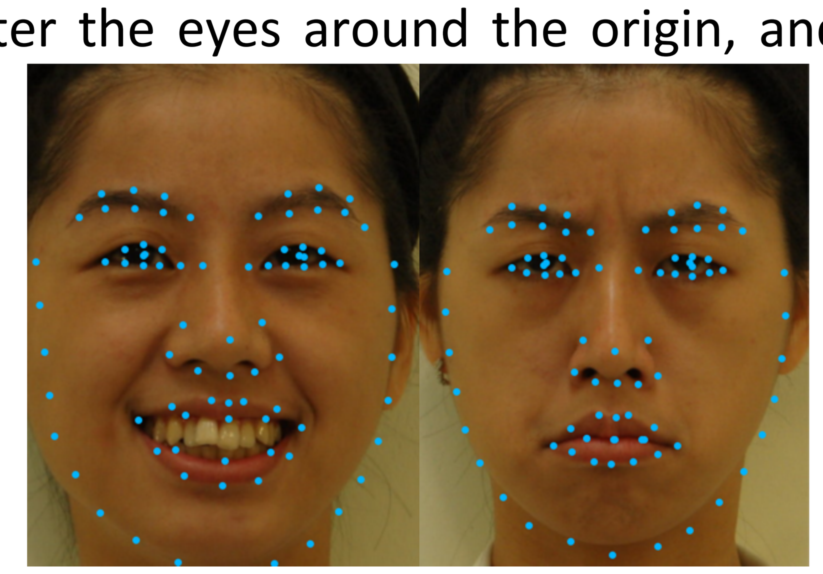
Facial expressions play an extremely important role in human communication. As society continues to make greater use of human-machine interactions, it is important for machines to be able to interpret facial expressions in order to improve their authenticity. Using faces from three different databases, we trained a multiclass Support Vector Machine to classify faces into eight different emotions: anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise. We found that, due to a wide variation in facial features among different people, it is difficult to identify emotions accurately.

DATA

Our data consists of 1166 frontal images of people's faces from three databases, with each image labeled with one of the eight emotions. The TFEID [1], CK+ [2], and JAFFE [3] databases consist primarily of Taiwanese, Caucasian, and Japanese subjects, respectively. The TFEID and JAFFE images are both cropped with the faces centered.

FEATURES

We have a total of 204 features: 167 features given by the public Face++ face detection API [4], and 37 features derived from this data. We used Face++ to get our original 167 features, which consist of a smiling metric and the x- and y- coordinates of 83 facial landmarks, which are shown in the picture below. Since each image can have differently sized faces at arbitrary locations within the image, we normalized each image by translating the face to center the eyes around the origin, and then scaled each image to fix the distance between the center of the eyes to a constant. The remaining 37 features are angles between certain landmarks that we decided varied among emotions.



Happiness Anger

MODELS

Softmax Regression

We used the MATLAB built-in function with a feature set of the 37 angles we selected, because our entire feature set was too high of a dimension to efficiently be calculated with softmax regression. The parameters of the model are those that maximize the log-likelihood function:

$$\sum_{i=1}^{m} \log \prod_{l=1}^{k} \left(\frac{e^{\theta_l^T x^{(i)}}}{\sum_{j=1}^{k} e^{\theta_j^T x^{(i)}}} \right)^{1\{y^{(i)}=l\}}$$

Multiclass Support Vector Machine

Using the LIBSVM library [5] on the entire feature set, we implemented C-Support Vector Classification with a radial basis kernel function. We experimented with different values for the parameters γ (in the kernel function) and C (the regularization parameter) and chose the values $\gamma = 0.0005$ and $C = 2.5$ for the kernel equation:

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}, \gamma > 0$$

RESULTS

We used all 1166 of our images to run tests. In the first table, precision and recall are calculated as the average of the precision and recall for each emotion. The second table shows the precision and recall for each emotion for the multiclass SVM 5-fold cross validation test.

Test	Softmax Regression			Multiclass SVM		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
Training	78.30%	76.20%	76.44%	98.89%	98.76%	98.88%
2-fold CV	66.21%	63.33%	63.35%	80.70%	79.43%	78.70%
5-fold CV	68.52%	66.34%	66.27%	84.56%	83.82%	83.09%

Emotion	Anger	Contempt	Disgust	Fear	Happiness	Neutral	Sadness	Surprise
Precision	78.43%	82.83%	90.12%	81.45%	94.15%	71.52%	80.81%	91.26%
Recall	82.19%	75.93%	86.39%	75.94%	97.79%	83.7%	68.97%	93.82%

The SVM model clearly fits our data better than the softmax model did. Our SVM classifier is most capable of recognizing happiness and surprise, and is least capable of recognizing sadness, which is often misinterpreted as neutral. This accounts for much of the disparity between precision and recall for neutral and sadness.

DISCUSSION

Since emotion tends to be very subjective in itself, it is probably difficult to achieve a significantly higher accuracy than our SVM does. Upon testing our algorithm on other people, we noticed that different people express the same emotion in different ways, and that it is practically impossible to capture all of these ways in our model. Some people are also much less expressive than others, and are often not sure how to express a certain emotion. Furthermore, the databases tend to be homogenous within themselves, both in facial similarity (largely due to race) and in the way certain emotions are expressed. As a result, training on one database and testing on another tends to yield unsatisfactory results. As expected, happiness and surprise were more easily expressed and identified than the other emotions. Certain characteristics, such as a smile or a wide open mouth, are unique to certain emotions, which causes them to be classified with a much higher accuracy.

FUTURE

As we refine our algorithm, it is important that we obtain access to a much larger and much more diverse database to make our model more robust to different people. We would also identify features that are more indicative of emotion. Finally, we would streamline our face detection and identification process so that it is less redundant and identifies emotion more quickly.

REFERENCE

[1] Li-Fen Chen and Yu-Shiuan Yen. (2007). Taiwanese Facial Expression Image Database. Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan.

[2] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101.

[3] Michael J. Lyons, Shigeru Akemastu, Miyuki Kamachi, Jiro Gyoba. Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998).

[4] http://www.faceplusplus.com

[5] Chih-Chung Chang and Chih-Jen Lin. (2001). LIBSVM – A Library for Support Vector Machines. National Taiwan University, Taipei, Taiwan. http://www.csie.ntu.edu.tw/~cjlin/libsvm/