Emotisense

Wendy Kopf, Tomasz Ksepka, Costin Vicoveanu
Current Progress

- Face Detection - Histograms of Oriented Gradients with SVM in DLIB
- Facial Feature Extraction - Face Alignment with Ensemble of Regression Trees in DLIB
- Face Detection and Feature Extraction on webcam and video running on TkInter GUI
- Gabor Kernel Generation in OpenCV
- Neural Net Classifier (SKNN) trained on feature vectors of JAFFE
- New QT GUI
Face Detection

Image divided into clusters of pixels (cells)
For each cell, a histogram of gradient directions is created
The Histogram of Oriented Gradients (HOG) Descriptor is created by concatenating the histograms across the image.
Face Detection

Human Face (or any object) is recognized using the computed HOG descriptor and a trained SVM.

(a) a test image; (b) gradient image of the test image; (c) Orientation and magnitude of Gradient in each cell; (d) HoG of cells; (e) the average gradient image over the training example; (f) the weights of positive SVM in the block; (g) the HoG descriptor weighted by the positive SVM weights.
Facial Feature Extraction

68 Facial Feature Points to locate
Placement decision made based on difference in pixel intensities
Check points similar to point locations from database of expressions, shape achieved from fiducial points

Figure 2. Landmark estimates at different levels of the cascade initialized with the mean shape centered at the output of a basic Viola & Jones[17] face detector. After the first level of the cascade, the error is already greatly reduced.
## Classification

### Assemble training data

**Face databases**

- [https://facedetection.com/datasets/](https://facedetection.com/datasets/)
  - Large 2D datasets
  - Unclassified
- **JAFFE** - [http://www.kasrl.org/jaffe.html](http://www.kasrl.org/jaffe.html)
  - Japanese Female Facial Expression Database
  - 213 images
  - 6 basic facial expressions + 1 neutral expression
Classification

- Train neural network on feature vectors for images in dataset
- Feature vector for an image obtained by convolving image with output of Gabor filters at extracted fiducial points
  - A set of Gabor kernels with 5 spatial frequencies and 8 orientations will be used to create a set of 40 Gabor filters for scale and rotation invariance
  - 40 Gabor filters x 68 fiducial points = 2720 - length feature vectors
    - Much larger than used before
- Neural network will output a closeness result for each emotion
  - Results will be weighted and the face will be classified with a final emotion
Classification

Example of 3 Gabor kernels (different sizes and rotation) and result of convolution with input image
Classification - Current Working Version

- Learning framework composed of Sklearn and sknn
  - Parses JAFFE dataset rating to get training output
  - Generates feature vectors for images as training input
- The input is scaled/normalized [0.0, 1.0]
  - Fed to a neural net through a pipeline
Classification- Current working version

- Currently have 68 points to detect on face
- Also have 40 filters working on the images
- From these, we have a working solution that currently has a 30 minute training cycle.
Final Plans

- Add face recognition for a group of input images
  - Focus group and/or recurring analysis
- (Maybe) add general sound recognition
Live Demo

Join us for a live demo of face detection, feature extraction, and some classification!

Thank you!