

Emotion Recogniser

Tianqi (Carl) Liu, *z5019791*

Han Zhang Zeng, *z1234567*

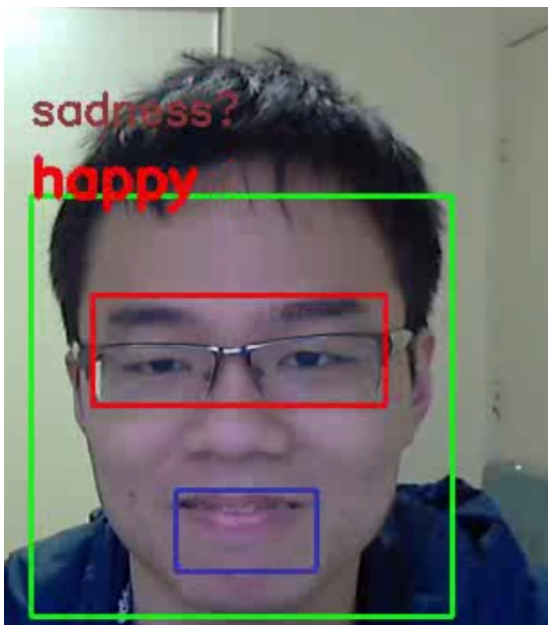
Abstract—This is a report on the emotion recognizer project for comp9517 at UNSW. the report includes our goal, reference survey, problem decomposition, detailed specification, design, test and justification, group contribution and conclusion.

I. INTRODUCTION

The goal of the project is to able to recognize human emotion on real time live stream to a degree of accuracy, at the start of the project we set the benchmark to around 70% because for most current object recognition algorithm an accuracy of 70% can be achieved (we will discuss this later in the test and design parts). There are 4 emotions that we are concerned with:

- 1) *neutral*
- 2) *happy*
- 3) *angry*
- 4) *sadness*

other: when the recognizer is not sure we consider the emotion to be some other emotion



For the final product, as shown in the photo, The face will be framed telling us which components are used for emotion recognition. There are 2 emotions displayed, the larger text on top of the face frame (Green) is the best

estimation of the current emotion, the darker red text is the 2nd best predication.

II. PROBLEM DECOMPOSITION

The project consists of two major components:

- 1) **Face Extraction**
- 2) **Emotion Recognition**

A. Face Extraction

In order to perform the emotion recognition algorithm, we need to extract the face from the image. We decided to use the fisherface recognizer which requires the input image to be of fixed size (detailed explanation will be in survey section) therefore it is essential that we first extract the face.

We chose the Haar Cascade algorithm for facial extraction because it is a really fast algorithm while maintaining a relatively high accuracy. Speed is very important for our application because during live stream, we want the algorithm to spend as little time as possible on the computation or else the stream will lag and the result will be unfavorable. Accuracy is not as important as speed in this context. When it comes to live stream, at every $\frac{1}{60}$ second, a new frame will be captured and computed, it is barely noticeable that one or two frames were not captured during a single second.

B. Emotion Recognition

The emotion recognition phase can be further classified into 2 parts.

- 1) **Training**
- 2) **Prediction**

In order to perform real time recognition we need fast algorithms, the best way to achieve this is to use a trained recognizer. The most common recognizers for the face are Haar Cascade Recognizer, Fisherface and Eigenface (details of each recognizer will be in survey sections)

C. Training

For the training process, we require large amount of data. We found a frontal face database from Cohn Kanade which contains 6000 images on 7 emotions.

- 1) *neutral*
- 2) *contempt*
- 3) *disgust*
- 4) *fear*
- 5) *happy*
- 6) *sadness*
- 7) *surprise*



sample images

We wrote a python code which classifies the emotions to folders named as the emotion, so in the happy folder contains all images that are happy



folder layout

We train the recognizer by extracting the images from each of the folder, detailed explanation will be in algorithm design section.

D. Prediction

For the prediction phase, we extract the face using Haar Cascade then pass it into the recognizer, the recognizer will output a specific value that corresponds to one of the emotions. The emotion will be displayed on the screen above the head frame.