BEX3012 Project Report Detecting Facial Expressions in Professional Tennis Matches

Stephanie Kobakian

28 March 2017

Contents

1 Introduction

Facial Recognition is beginning to be explored in sports environments, this presents quality issues.

helpful for applications

2 Method

2.1 Materials

2.1.1 Pre Processing Images

A subset of faces was created specifically for emotion recognition purposes. To present only a single face to each software required cropping the region of the image based on Google's bounding box result[1]. Each new image created contained the area within the face bounding boxes found in the previous study, as well as a small border to frame the face. This resulted in images of differing sizes to be passed to the APIs. These 1319 new images were hosted on Google Drive to allow for URL access from the API to the individual images.

[1] coordinates derived from prior study

Figure 1: Here are six examples of Faces in the image set after extracting faces from the full broadcast video stills.

2.1.2 Software APIs

We will consider three Emotion Recognition software.

Table 1: This details the capabilities we considered important in recognising emotions in images of faces.

Attribute	Google	Microsoft	Skybiometry
Batch Processing	10 per second	20 per minute	100 per hour
Emotion Output	?	Numeric Proportions, (0.0-1.0)	Confidence Value, (0-100)
Number of Emotions	4	7	7
Cost and Access	Account and Payment	Account	Account and Payment
API Access	REST	REST	REST

As can be seen above, a noticeable difference between the three is the amount of times the API can be called. Skybiometry had the largest imposition on Bath Processing as it only allowed 100 API calls to be processed per hour. Microsoft also had a limit imposed, but this allowed for much more to be processed with the possibility of 1200 images to be processed within an hour, after accounting for the wait time between each group of 20. Google Vision's API batch processing limit had a minimal effect.

The range of outputs from the software is displayed as Google provides likelihoods of an emotion occurring on a particular face. Microsoft and Skybiometry provide outputs on the same emotions[2]. However the values that they provide differ as Microsoft provide Proportions whereas Skybiometry results in a Confidence of the emotion occurring in the specified face.

2.1.3 Processing API results

As the results provided by each of the software differed, processing was needed to create comparable sets.

2.1.3.1 Microsoft

Table 2: The Microsoft API provided 8 numerical values, one for each emotion.

anger	contempt	disgust	fear	happiness	neutral	sadness	surprise
0.001504	0.000631	0.002371	0.003267	3e-05	0.716105	0.004738	0.271353

2.1.3.2 Skybiometry

[2]Based on Paul Ekmans emotion theories

Table 3: The Skybiometry API provided 7 numerical Confidence values, one for each emotion.

neutral	anger	disgust	fear	happiness	sadness	surprise
0	21	41	0	0	5	76

2.1.3.3 Google

Table 4: The Google Vision API provided four likelihood possibilities, one for each emotion.

joy	sorrow	anger	surprise
VERY_UNLIKELY	UNLIKELY	VERY_UNLIKELY	UNLIKELY

These differing outputs needed to be arranged in a comparable manner. We began by considering the Microsoft output. Noticing that the sums of all emotion values was approximately one we considered that dividing the individual Skybiometry emotion confidences by the sum of the confidences would give the likelihood of each emotion being the dominant emotion in a face.

This calculation was simple to perform and resulted in numeric values comparable to Microsoft's values.

Skybiometry Transformed:

Table 5: The Skybiometry API provided 7 numerical Confidence values, one for each emotion.

neutral	anger	disgust	fear	happiness	sadness	surprise
0	0.146853	0.286713	0	0	0.034965	0.531469

We then needed to transform the Google output to also be comparable. This was the most challenging due to it's unusual form.

A numeric value was assigned to each possible likelihood value. This would allow for numeric comparisons to be made. The values were assigned according to the table below:

Table 6:

Likelihood	Value
VERY_UNLIKELY	10
UNLIKELY	30
POSSIBLE	50
LIKELY	70
VERY_LIKELY	90

As both Microsoft and Skybiometry included a neutral category we considered how we may incorporate this for the Google results.

Google Transformed

Our approach considered that when a face was "very unlikely" to be any of the four emotion categories it could be deemed "neutral". However when one emotion was stronger than other it could still be possible that the face only had a small lean toward this emotion. Therefore the Confidence value for neutral was derived as:

$$neutralConf_i = 100 - max(joyConf_i, sorrowConf_i, angerConf_i, surpriseConf_i)$$

The process to derive the numeric values followed the process undertaken for Skybiometry values. Where the individual emotion confidences were divided by the sum of the emotion confidence values for each face.

This resulted in congruent emotion data to be compared

Table 7: The Google Vision API provided four likelihoods, one for each emotion.

joy	sorrow	anger	surprise	neutral
0.0666667	0.2	0.0666667	0.2	0.4666667