Emotional Recognition

Abstract
Human computer interaction has become an increasingly important field of computer application design. As society and individuals increasingly rely on computers for every aspect of their lives, new technologies and products can thrive when the time taken for the understanding needed to operate software packages and hardware devices is lessened. Creating intuitive interfaces for people, as well as reducing the number of errors received by the computer can save individuals and business a lot of time. Given the competiveness of software and hardware tools and packages it is important that developers consider not just what and how an application will conduct a function, but also the particular effect using it has on its audience as this can be the deciding factor in the success or uptake of these tools.

Emotional recognition is a frontier of applied computing that aims to bridge this gap between how a user feels and how a program runs. It improves user testing feedback by having feedback throughout the use of an interface, media application or from an image or voice. Many different techniques are currently being researched for application in a number of different fields which range from security, law and psychological conditions to computer games, social media and web design. [2], [3], [4]

With the recent popularity of mobile gaming for example, psychologists are often employed by large gaming companies to ensure a balance of reward and effort which makes games fun in a way that entices customers to keep playing. Currently simple game metrics such as duration of game and hours spent playing can be collected by the developers which can be then considered for further development. Emotional recognition technology could be of great benefit to this industry by offering detailed analysis of how certain game features have an effect on their customers. The aviation industry has employed aspects of gamification technology and emotional recognition in the form of giving their attends more work when it senses they are bored and less when they are optimally focused.

A more familiar application of emotional recognition takes place in the court house. When witnesses are testifying it is not only the logic of what they are saying that is given attention, but also the way it is delivered. Polygraph machines are a good example of machines that aim to give an incite into an individual's emotional state. By placing monitors on an individual and observing their change in response to questioning, deception and stress can be detected by the machine and interrogator. These techniques are considered for use in court, however the technology has not been updated for several decades despite advancements in technology that help remove the ambiguity of the machines output. A polygraph test also relies on a highly trained individual to interpret the results, which can also be influenced by witnesses or suspects who have above average control of their bodies natural responses such as breathing and heart rate. It has also been known that individuals with certain psychological conditions will regularly pass polygraph tests, as they do not respond in a way that can be detected by the machine.

Emotional recognition in a court room setting has the benefit of modern computing techniques such as machine learning to adjust for discrepancies in human nature. Video footage can also be analysed and an emotional state such as fear or anger can be determined by algorithms. This extends the understanding of an individual's state, not just when giving testimony, but also from CCTV or while remotely given evidence. The expertise of the algorithm can be implemented in two different courtrooms at the same time which is more efficient than traditional polygraph tests. Giving more forms of reliable evidence could reduce miscarriages of justice, especially in those which involve interpreting how someone is feeling.

Emotional Recognition

Emotional recognition can take advantage of a number of emerging technologies. As facial recognition software has become more accurate, it can be applied to detecting micro-expressions in general features in faces such as lips and eyebrows. One of the major difficulties with a computer vision approach to this problem is that currently, computers can only process images in two dimensions even though the world is in three. Such problems that stem from image recognition include a strict scope of possible input images as a computer can have a difficult time differentiating one face from another, detecting a face at an angle, or even interpreting non human objects as faces such as seen in the popularity of 'face swapping' algorithms. This difficulty is most readily addressed by matching the program specifications where possible with an expected image format which could include standards similar to a passport photo in terms of lighting, size of face, orientation and background.[3]

There are many algorithms currently used to determine an emotional state of an individual however most of these techniques are more statistically based then through means of computer science.
An example of a computer program written in a pseudo code style is given below for analysing a still image and returning an emotional state. This procedure makes use of applied computing concepts such as facial recognition, and searching algorithms as well as the possibility to be developed further with emerging technologies such as machine learning and wearable technology. [5]

```
1. Receive input file -> start scanning image for face
        If face found {
                Separate from background }
   Locate facial features(extract data)
2. If image (!inDatabase) {
        add to database}
3. Else {
        compare from database }
4. return emotion [1], [5]
```

1. Automatic face detection can be performed by a number of algorithms such as Viola-Jones. Object detection is applied on the similar features of human faces, such as the distance and size of features in relation to each other such as eyes and nose. To make sure the face is sufficiently separated from the background, the colour space of human skin tones (which fall within a very narrow range of total possible colours) will be programmed into the algorithm. This makes it easier for the computer to determine where the face stops and the background begins for more advanced image processing.
By cropping the image to focus on lips, processing time is reduced by focusing on a small part of an image, as well as a region of ones face which varies the most in relation to differing emotional states. Extracting this feature can be performed a number of ways. To reduce the number of needed software components for this algorithm, it would be more efficient to re-use the colour filtering techniques previously used to separate the background from the face on the face itself to focus on lips.

2 and 3. By entering the image into a database, this can aid the computer in detecting and deciding which emotional state is a best fit for the received image. This stage in the algorithm would benefit from machine learning techniques such as a genetic or evolutionary algorithm as well as training data overseen by a human operator or self reported feelings from an individual.

4. Philosophers may argue that there is an infinite range of Human emotion. Psychologists have broadly categorised these feelings into one of 7 or 8 states. Statistical methods are applied against the given image and several already classified in the database and returns the state associated with the largest percentage of similarity.

In stage 1 of this algorithm, an error may occur if there is more than one face or only part of a face shown. This can be mitigated by having input photos given under optimal conditions however for use in a courtroom setting as described above, it could only be used in testimony and is unlikely to be gathered as standard evidence. Another way to ensure correct results would be to have the computer directed to what is a face or which faces to perform the rest of the algorithm. This is not entirely favourable as it wold slow down the entire operation.
Stage 2 suffers from a similar problem from a legal perspective as many CCTV cameras, especially those filming at night, are in black and white and would not be able to take advantage of colour analysis and would be more likely to throw errors in image analysis such as converting to a binary image to save database space or to segment features. To improve this section of the algorithm, a more sophisticated method of objected detection or a recursive version of the one employed could be used in when a black and white image is given.

Stage 3 can take advantage of the research that has already been undertaken in searching large data sets and shows promising results when combined with artificial intelligence techniques such as genetic algorithms. Despite the robustness at this stage some human oversight will undoubtedly prove necessary to ensure that the variables involved in genetic algorithms are producing desirable results. Over time the general algorithm could identify key features of certain emotions depending on the resolution and overall quality of the input image.
Section 4 also requires human intervention in the form of training the machine to produce desirable results. It also relies on limiting the broad range of human emotion down to a category of approximately seven differing states. It also requires each system to have access to a large database with may limit some applications of this technology. By returning a general confidence range of each emotional category, a baseline confidence should be returned until the algorithm has enough samples in the database to draw an accurate conclusion. [1], [5], [7]

Conclusion

The algorithm above can be applied to other computing areas such as MRI scans without much modification or by replacing the visual components with audio pattern matching. A system for sensing human emotions would ideally take advantage of different input methods to produce more reliable results and could be applied to a range of many different computer interfaces that interact with humans. In the last 3 years there has been a vast increase in the number of wearable technology products that prove the commerciality of monitoring aspects of out lives such as physical activity, sleep and even hydration. As of yet, there is not a widely available product that solely focuses on monitoring and maintaining emotional states, although the benefits of such technologies could easily be applied for uses in improving mental health or the understanding of improving user interaction with software. Web admins routinely collect detailed information on how users interact with their site based on metrics such as time spent, having data accessible on how users emotional state changes with use oft their website, social media or even video platform would be beneficial in generating more meaningful interactions for all parties involved. [5], [6]

[1] Emotional Detection Algorithm, N Mahajan, H Mahajan
[2] “Can a computer correctly identify emotion?”, R Burks
[3] “Emotional Recognition APIs”, B Doerfield
[4] “Computer intervention to teach people with Autisim to recognize emotions”
[5] “Using non-invasive wearable computers to recognize human emotions”, C Lisetti
[6] “Decoding spontaneous emotional states in the Human Brain”, P Kragel, K LaBar
[7] “Genetic algorithm and neural network for face emotion recognition” K Muthuharuppan