Determining Emotions via Facial Expression Analysis Software

Richard W. Taggart, Michael Dressler, Poonam Kumar, Shahroze Khan, Jean F. Coppola

Pace University
rt86688p@pace.edu

Abstract—Facial expression recognition has been a highly researched topic in the field of biometrics for decades. It has been used with the intent of both identifying specific individuals, and in understanding human relations and communication. Grouping these expressions into emotions is a subset of this research, which is still new. The focus of this study is to assess the validity of iMotions as a facial expression recognition biometric system, and determine its use in association of those expressions with emotions. This may include understanding the difference between an emotional or reflexive response to external stimuli. A potential case study with this type of software is in medicine and geriatric care. Patients in these situations may not always be able to communicate their state of being with a care provider. Using a system such as iMotions may be able to assist these victims in the future. The purpose of this study is to evaluate iMotions biometrics platform as a system for collecting and analyzing facial expressions.

Index Terms: iMotions, Emotient FACET, facial expression, emotional response

Introduction

Facial recognition is a key biometric characteristic, which can be utilized to differentiate one being from another. People inherently use subtle differences in the facial structure to recognize another associate, colleague, or mutual acquaintance. These characteristics are considered to be specific to each individual, allowing them to be used as an identification mechanism. Facial characteristics are also utilized and manipulated by an individual to convey expressions [8]. Each of these expressions can be associated with a specific human emotional response. Determining a state of being, mood, or emotion through facial recognition is the topic of study in this paper.

Biometric Systems

Biometric systems have been a focus of study for decades, but are finally entering the consumer market. Most applications of biometrics fall into the category of security, from unlocking a mobile device with a fingerprint, to facial recognition software running on security camera video footage [12]. Facial recognition will collect key features from a video or image frame, and then compare those characteristics to a biometric database, while looking for similarities or possibly even a match [12].

Facial Features and Expressions

Humans have been studying themselves for a long time, and the description of facial features is no exception [8]. The measurement, collection, and systematic analysis of facial expression has been a focus of study since the initial publication by Paul Ekman and Wallace V. Friesen in 1976, almost half a century ago [7]. The specific method and deliberate analysis of such features are commonly known as the Facial Action Coding System (FACS), originally created by P. Ekman. There have been many studies performed regarding the collection of facial recognition metrics, and this space is well explored. Topics from understanding the correlation of quality of input, for example varying lighting conditions [9], to understanding which facial expressions represent reflexes to external stimuli compared to true human emotion [7].

Facial expressions are a gateway into the human mind, emotion, and identity. They are a way for us to relate with each other, share understanding, and compassion. They are also a way for us to express pain, sorrow, remorse, and lack of understanding [10]. These characteristics can be crucial to understand when working with patients, especially patients who are unable to communicate in other ways. These victims include post-stroke patients and those suffering from dementia or Alzheimer’s disease. For these patients, it can be helpful to identify a state of being, or emotional response to external stimuli, including pain, through their facial expressions.

Proceedings of Student-Faculty Research Day, CSIS, Pace University, May 6th, 2016

C2-1
iMotions

This study examines the use of facial recognition software in the field of identifying human emotion. This will require the use of a biometric research platform software system called “iMotions.” iMotions, Inc. was founded in 2005, and its headquarters are in Copenhagen, Denmark. They also have an office in Boston, Massachusetts. According to iMotions, the software is able to combine “eye tracking, facial expression analysis, EEG, GSR, EMG, ECG and Surveys” The platform is used for various types of academic and business-related research [1]. The validity of this software and its ability to associate specific facial characteristics with specific human emotions will be tested. This will also perform as a case study for the usability of such a system.

Galvanic Skin Response

Galvanic Skin Response (GSR) is another biophysical sensor, which determines human skin resistance under different psychological conditions. The GSR is an older term for electro dermal activity (EDA), which is simply the electrical conductance of the skin.

These sensors also detect an increase in physical attributes marking a state of being including: heart rate and sweat measurements. Sweat glands are controlled by the sympathetic nervous system. A change in the electrical resistance of the skin that is a physiochemical response to emotional stimulation increases the sympathetic nervous system activity. GSR is a method of measuring the electrical resistance of the skin, which varies with its moisture level. With other sensors, these devices can help determine wellness, and emotional responses to external stimuli.

Literature Review

In 1976, Mr. Paul Ekman published the first well known work on systemically analyzing human expression. With his colleague, he built on prior work and devised a system known as FACS, to record and measure facial expressions [7]. This system has been used for nearly half a century in many works while collecting samples of the human face. He illustrated a mechanism to be used which can be used to measure specific features of the face, and will quantitatively identify certain facial expressions, compared to other ones [7].

Facial Recognition Technology

The first semi-automated, or “man-machine”, facial recognition technology was created in 1968 by Bledsoe, Wolf, and Bisson. The researchers took mugshot photographs and created data by hand measuring facial features of individuals such as the distance between pupils, width of mouth, the point of a widow’s peak, etc. The data was then entered into a computer database. Success was measured by having a computer compare the data in the database to new sets of entered data from different photographs and producing matches [16]. In 1971, Goldstein, Harmon, and Lesk developed a standardized list of twenty-one facial features that could be measured to create a linear embedding algorithm. Unfortunately, this method proved to be too time consuming because the facial features were still being measured by hand [17].

By 1991, Turk and Pentland discovered that they could use Kirby and Sirovich’s Eigenvectors algorithm, published in 1988, to detect faces in images. Eigenfaces uses relatively small amounts of data, called “Eigenvectors,” taken from two-dimensional images to find variations in faces. Turk and Pentland’s research was especially important at the time because of their ability to use ever expanding computing power to fully automate facial recognition. Their research played a major role in creating modern facial recognition software [18].

Since the early to mid-1990’s, the U.S. government has funded projects to produce more advanced facial recognition algorithms and software products. Additionally, the Defense Advanced Research Products Agency (DARPA) created the Face Recognition Technology Evaluation (FERET) to help speed along the advances being made in Facial Recognition Technology. FERET lasted for five years and led to the production of multiple commercial products by the early 2000’s. Finally, the government has tested vendor products over the years since the 2000’s to “assess the capabilities of commercially available recognition systems” and “to educate the biometrics community and the general public on how to properly present and analyze results [19].”

iMotions

Articles based on the use of, or based in-part on the use of, iMotions have been conducted on various topics such as: recognition of emotion on faces of different races [2], “technology-supported design research” [3], and a survey of the techniques used to measure “emotion-enhanced interaction” between people [4]. Additionally, studies using iMotions primarily for eye tracking are prevalent. These studies include a feasibility study of “inferring web page relevance” using eye tracking, [5] and measuring human emotions using eye tracking [6]. However, there are no known studies on the use of the iMotions platform within the geriatric community.
Galvanic Skin Response

Various experiments have been conducted and reported for GSR. The studies and experiments included the comparison of the GSR in the hypnotic and waking state, effect of music stimuli on response, driver’s galvanic response and the risk of accident, effects of alcohol on GSR [15]. In one of the experiment, The University of Akron examined the biochemical and physiological responses to music stimuli. Music and Biology students from the university were involved in this study. Their Galvanic responses were measured before and after listening to different musical selections in an anechoic chamber and during controlled silence. It was noted that GSR responses were significantly higher for students who were in music major than biology major. Music may ‘turn on’ or ‘turn off’ specific emotions which can increase the certain chemicals in the blood and can be responsive to the emotions while measuring the GSR [11]. There was a color experiment conducted by 22 undergraduate or graduate students using Galvanic skin response in the University of Wisconsin. A pair of electrodes was attached to their fingers and their Heart rates were monitored. Black, blue, and magenta colors were presented to them for a thirty-second-time interval. In between every color transition, they were asked to close their eyes for fifteen seconds. When the student opened their eyes, computer program determined the time at which any particular student responded to command prompts. Data was then analyzed by comparing mean alpha and beta wave amplitude of EEG, heart rate beats per minute) and mean Galvanic Skin Response values between measurements taken during stimulation by magenta and blue. Average GSR response was then calculated and concluded that blue was considerably lower than magenta values which is because of increased sympathetic response [13]. There was one other experiment where GSR activity of the drivers were measured and concluded that the level of GSR activity does not depend primarily on the nature of the road or conditions only but also on the variation in the GSR levels (emotional tension or anxiety etc.) also. Distribution of GSR per unit distance travelled was found to be similar to the distribution of accidents per unit total distance of vehicle travel [14]

Project Requirements

This study covers the validity of iMotions software, as well as the GSR with experimentation and discovery. First, tests will regulate the validity of the iMotions software system as a human emotion recognition system. The primary objective achieved through data collection based on still frame images and videos of older adults in conjunction with emotion invoking videos, as well as music. The samples collected with be run through the iMotions software toolset. The biometric collection system will have a set of data which can be analyzed and interpreted based on the input, and expected behavior of the system. Further testing of this system will include the ability to handle varying inputs including diverse test subjects. Some examples include: varying skin tones, altering features such as wrinkles, age demographics, and facial accessories such as eye glasses.

The possibility of inducing or changing a specific mood or state of mind via music or some other sensory input will also be investigated. Recognition of these types of emotional state changes with this software through experimentation will also be explored. Due to the nature of this project, a test collection system and method for inducing emotional responses will need to be put in place. This will include collecting videos from test subjects, corroborating various feelings, which they experienced through the test, and analyzing the iMotions system output which could then be correlated with the original raw data. GSR can be used in conjunction with other biometric collections to determine many things. Potential uses of a GSR system in addition to facial expression recognition will be investigated. There may be a follow-on case study pending the findings in this work.

Methodology

The research methodology of this project requires gathering of various sample video footage of 30 participants including male, female, participants of different age, color, physical features etc. to record their facial expressions and reactions which will be later analyzed by using iMotions software. In this research, we will also use GSR plug and play module of iMotions.

GSR Module of iMotions is a plug & play integration with GSR devices that delivers real time sensor output (emotional reactions) in the user interface and in synchronized raw data export. iMotions also provides an open API to allow us to integrate the own GSR sensors to forward data into the software. We can also integrate it with other best in class sensors like EEG, EMG, ECG. GSR solution will allow us to analyze the different emotions of people (i.e. stressed, sad, happy, etc.) through iMotions.

Test Subject Participation and Data Collection

We will collect sample videos from a test subject group of roughly 30 participants. If the scope of this study permits, that group will be expanded to 100. This group will be separated into two sets of individuals: the control group, and the experimental group. Each of these groups will be directed to either a mobile app or a website which
will provide a consent form, an initial demographic survey, and then the test area. Each participant will be exposed to a 5-10 min set of videos, which are expected to induce some set of reactions. While these videos are playing, the application will record their face, which will later be used to determine specific facial expressions. This video and the original stimuli will be stored and later analyzed. A screenshot of this application is depicted in Figure 1. The sample set will consist of approximately equal number of men to women. We would like to make this set of people as diverse as possible: glasses vs. no glasses, facial hair, etc. More research is required to determine the minimum set of participants required to ensure a fair and reasonable study may be conducted.

Data Processing

Once the video footage is collected the videos will be analyzed in iMotions software. This software will determine the level confidence it has that a given expression is associated with a specific facial expression. We will be able to analyze the set of videos in iMotions in bulk. This will allow us to separate the analysis portion of the study from the data collection. For each stimulus, there is an expected expression, a measured expression with some level of confidence, and the true expression felt by the subject. We will compare the expected and measured expressions based on the peak, trough, and average confidence level that iMotions produces for each video segment. We will also try to evaluate the level of correlation between measured and true expressions. A prototype of this type of analysis is displayed in Figure 2.

Figure 1: Screenshot of Data Collection Interface

For the control group, the stimuli provided will intentionally not attempt to elicit strong responses of any kind. This will allow us to collect an unbiased set of footage and collection metrics for a baseline. This baseline will then be compared against the experimental group measurements. The experimental group will be shown a set of stimuli, which will be expected to resonate strongly with the test subject in all areas under test. The following emotional responses are expected to be collected: Joy, Anger, Surprise, Fear, Contempt, Disgust, and Sadness. In each of these areas, for both groups, we will collect the peaks and troughs of each measurement. We will also collect the average levels of these emotions. The input stimuli will be separated into segments, which can be individually analyzed. Over each of these segments the peak, trough, and average confidence level for each emotion will be recorded.

We will also use video from already existing facial recognition repositories, which can be fed through iMotions software. Depending on the quality and intent of these videos, it may be helpful for us to determine the effectiveness of the software in question.

Data and Trend Analysis

Each of these can be plotted separately on a per individual basis. This study will also associate all of the data collected and plot the trends, which are observed over the entire sample set. The focus of this study is limited to the following questions: Does gender have an
effect on the ability of this software to recognize emotions? What about other traits such as age, race, ethnicity, physical features or deformities? Do glasses or facial hair affect the ability to determine the effectiveness of the software?

Throughout the experiment, the subject will be expected to associate with a specific emotion, and elicit a response accordingly. The main focus of these experiments is to determine the level of confidence iMotions has detecting each of these expected responses. We will compare the metrics gained from the control group against those collected from the experimental group. Approximately 4 times during the experiment, the external stimuli will pause, and the subject will be asked which emotion they most closely resonate with. This may help us gauge the level of confidence that we have in this software.

Participant waiver: we will provide an agreement with the test subjects to keep the video footage collect anonymously and to not associate their name with any of the other information collected. Any information collected will be used solely as a generalization to provide potentially correlations in demographics.

iMotions provides API’s to the platform which can be used to import and export data. This study will investigate the potential use of these API’s in order to automate parts of these processes. These API’s may also be used by other applications to gain a confidence level in a subject based on prior input from that subject. If the scope of this study permits, the control and experimental groups will be switched. This will allow the study to collect a baseline for each participant as well as effectively double the sample size of the group under experimentation.

There are several competing solutions available, which provide facial recognition solutions. Some of these platforms claim to associate a person’s mood or state of mind with their facial expressions. The scope of this study may be limited to iMotions, but these platforms are a good area to investigate in this field.

**Experiment Performed**

To start the experiment with iMotions, the Emotient FACET module has been enabled. One of our team members was exposed to 2-3 minutes of video, which induced some set of reactions on his face. We also captured teammate’s live face recording in Motions software while he was watching the test video. We then, created a new study in iMotions and added him as a Respondent. It created a graph, which consisted of different reactions on his face e.g. fear, joy, anger, surprise, confusion etc. The video of the team member was later analyzed based on his reactions on the graph. Data on the graph was also exported into an excel sheet by using data export feature of iMotions. Exported data contained the detailed analysis of the study, e.g. Study Name, Name of the respondent, Age, Stimulus shown to the team member while his face was getting recorded, Reactions on his face etc.

To conduct the GSR experiment, we first connected and configured the Shimmer device in iMotions in order to collect GSR with optical heart rate. Once connected, a team member strapped the three nodes to their fingers and then listened to five 30-second samples of different types of music. The types of music ranged from classical symphonies to rock. Because the GSR can only be collected physically in front of an iMotions machine, data is available immediately. The data was created in graph form, similar to the ones associated with the facial recognition module, depicting two charts; GSR and PPG (heart rate).

**Results**

The following findings have been discovered throughout the duration of this study. Each of these findings is backed by the experiments conducted from the research in this project.

**Finding 1: Remote Access to iMotions**

iMotions is not designed for remote use. As mentioned previously, the iMotions software is an excellent platform for collecting data from multiple biometrics sources, as long as the experiment is taking place locally. Part of the focus of this study is to evaluate the ability for this tool to process and analyze data from external sources. A web-based test bench application was created to collect videos from remote participants. The test bench would record videos and then save them on a remote disk, or cloud drive. This data web based data collection framework is available online as an open source application [20]. Unfortunately, due to lack of Pace IT resources, the study was limited to the participation of the members associated with this project. Each member recorded their face while watching a predetermined stimuli set of 5 YouTube videos of 40 seconds. These recordings were stored in a Microsoft One Drive folder, and later accessed by the team for analysis. The iMotions team has an API, which they can enable for Pace University. However, this study was not able to take advantage of the 2-week trial period with the iMotions API. This API will potentially enable researchers to automate parts of the data collection and export process.
Finding 2: iMotions Video Processing

iMotions accepts two formats for video processing: .mp4 and .wav. Both types of videos were processed in this experiment, and each yielded reasonable results within the tool. It is worth noting that video processing portion of the iMotions software consumes significant CPU and memory resources on a machine. Through this testing it is recommended that a laptop with iMotions 6 installed have at least 16GB RAM, and a reasonably fast processor like the Intel i5 or i7 cores (3.0+ GHz). To import the video, a new study must be created. Next a stimulus must be added, and it must be a “Face Recording.” There should be no other stimuli added to this study. Now iMotions is configured to enable import of external video. This step can only happen once per study, which is inconvenient if the videos are not all taken at a single time. Once the videos are selected, iMotions can post-process the face recordings. In this manner, it behaves the same as if the videos were recorded live through the platform. When the videos are processed, a baseline measurement must be taken. The FACET baseline must be at least six seconds long and should consist of video of the participant in a neutral facial position. iMotions can use this baseline to adjust the relative emotient thresholds, which yields a more accurate result. It has a confidence level associated with the following emotient FACETs: Joy, Anger, Surprise, Fear, Contempt, Disgust, and Sadness. It also detects emotient valance (Positive, Neutral, Negative) and will associate a video segment with a relative confidence level. Figure 4 shows an example of the video processing and associated emotient confidence levels. These visual queues are helpful while watching the participant’s video. Significant events in the video can be manually recorded and checked in this manner.

Finding 3: Data Collection, Export, and Analysis

The iMotions platform is intended for data collection from several domains of biometrics. It is not intended to provide higher analytical capabilities. In order to post-process the videos and gain valuable insights from the data collected another program must be used. Examples of statistical packages include R, SPSS, Excel, and libraries available for Python. iMotions will allow a user to export the processed data into a large CSV file. This file can be read in by many applications. A python script was written to post-process this data and massage it for graphical representation. This is also available online with the iMotionsDataCollector [20]. The script will produce several JSON files from the original CSV file, which are sent to the iMotionsDataCollector web application to be rendered. A simple graph depicting the relative responses of two participants can be observed in Figure 3. Based on the graphical renderings of this data, it was observed that the tool rarely assigns an emotient FACET with more than 0.5 confidence level, which is rated on a (-1, 1) rating scale. It should also be noted that the current rendering system does not perform well with large volumes of data. This should be taken into consideration in future projects. Through manual observation of the processed video, it was evident that the only times a confidence level exceeded 0.5 was when there were clear positive facial features, such as exposed teeth in a grin. From this data, it can be determined that a more powerful external stimuli needs to be applied, in order for the participant to engage with the experiment. Furthermore, it was observed that the participant’s state of mind prior to the experiment might alter results. For example, if someone is angry, they appear to be less affected by videos intended to illicit joy when compared to other participants.
Finding 4: GSR

The data collected from the GSR module was reasonably accurate. The test subject, exposed to different types of music and their heart rate responses were picked up instantly, where as the skin conductance response was not as efficient due to some possible factors such as room temperature, body temperature, and limited exposure of music. However, the heart rate data provided promising results. For example, when the participant was exposed to classical music, heart rate remained leveled and did not fluctuate from the base line. However, once the music was switched to a sample of hard rock, the participant’s heart rate increased noticeably and continually fluctuated. The response time was little to nothing, and any fluctuation was shown by the live graphs.

Impactful Uses of Biometrics

Although biometric recognition has been in existence for hundreds of years, it has become more available to the general public within the last 20 years due to the obvious advancements in technology. Currently, iMotions and other similar software applications have been used in various industries to better provide feedback from not just the researchers, but the users themselves. Two of the biggest trades where facial recognition and GSR, as well as other biometric modules, have been applied are in the Healthcare and Marketing industries. In healthcare, it has been used to improve care for patients both physically and internally. In marketing research, companies have been able to gain data in order to better understand the needs and wants of their consumers.

With so many possible applications, the healthcare industry is one of biometrics’ most promising opportunities and is expected to generate revenues of $3.5 billion annually by 2024, according to a report by technology research group Tractica. Besides its potential economic gains, the main practices for biometric applications in the healthcare industry have been for the patients of medical institutions. For example, the experience of pain is inherently subjective and, many times, biased. One patient’s rating of pain on a 10-point scale may be a 5 where as another patient going through similar tests may rate their pain as a 10, especially patients with dementia, who often are not able to verbally describe their pain. Because of this, many clinicians have trouble accurately assessing the pain. But, with facial recognition software like iMotions, a clinician can measure pain based on video images of their patients, and then have the data print results of how intense the pain is and where it may be coming from, resulting in better general assessment of patients. Besides pinpointing pain, facial recognition technology can also be used to spot genetic disorders in patients. Between 30-40% of genetic disorders, including Down’s syndrome, involve some kind of change to the face or skull. According to researchers, by collecting facial photos of previous patients with particular genetic disorders and feeding them through advanced computer vision technology, the software can, in a way, “learn” what facial changes are linked to which disorders. This will again allow doctors to better assess their patients. By combining multiple biometric methods, (e.g. facial coding, GSR, EEG, eye-tracking) one can also diagnose many psychiatric disorders as well as neurological disorders.

The other big industry, next to healthcare, that can benefit from biometric data is the Market Research industry, specifically neuromarketing market research. This form of marketing help’s researchers to observe consumers’ cognitive and emotional responses toward certain stimuli such as pictures of products, TV ads, and online advertisements. The main purpose of advertising is to create awareness and to influence the buying behavior of consumers, hopefully resulting in consumers to remember those specific products. This is where eye tracking, facial coding, and GSR come in to play. By using these modules, researchers can collect data which can depict how quickly consumers notice the product, how long they look at the product, and how interactive they were with the given stimuli. This helps to better gauge companies during their product testing phase and allows them to release products which their consumers actually want or need. To show how important accurate advertising has become, Google was recently granted a patent for “Pay-Per-Gaze”, an eye-sensor which will be built into their forthcoming Google Glass computer. This gaze tracking system will track user’s interests in products based on measurements of pupil dilations while looking at certain online ads/commercials. Historically, advertisers have captured responses from customers after their purchase, with surveys or other feedback questionnaires to determine if a specific ad had any impact on the purchase. However, the time, and technology, has come where marketers will be able to provide real-time analysis through targeted content directly to consumers based on how they are feeling.

Conclusion and Recommendations

This research paper has demonstrated that Emotient’s FACET technology module of iMotions software can easily recognize the human emotional reactions on faces of people. The experiment consisted of analyzing the emotions data from the facial study based on the team members as respondents. It was very interesting to find that this module of iMotions software can pick up even a micro expressions or even little flickers of emotions from respondents face and can easily tell how that person is
feeling (i.e. joy, sad, fear, anger, surprise etc) at any particular time when he/she was being recorded.

Additionally, this study also conducted an experiment using wireless GSR devices from Shimmer that is integrated with iMotions software and concluded that it can monitor the emotions of the person under experiment by measuring his/her skin conductivity through reusable electrodes attached to two fingers of the hand. While doing the experiment, we have noticed that the iMotions platform was not built for remote use. Successful studies are possible when completed remotely, however the amount of time needed to complete remote studies can substantially shorten the amount of time for analyzing results. Even with the computer itself being correctly set up for remote users, these users can’t easily remedy software or hardware issues that may arise. This may be true for users who are on campus as well, but they should be able to find solutions to problems in a much quicker manner. Also, it is not possible for the users conducting studies to collect GSR data from remote respondents since the shimmer device that collects the data from the respondents must be connected to the computer that is running iMotions. Therefore, it is suggested that future studies at Pace be completed by students and respondents who are on campus and can easily access the iMotions computer on a regular basis.

Currently, it can take several minutes for the application to process just two or three face recordings, each of less than 300 MB. This will likely be a hindrance to future studies that intend to process a larger sample size, such as thirty or one hundred videos, which are each just three minutes in length. Videos for individual studies conducted with iMotions must all be processed at once, since the application will not allow additional videos to be added to a study at a later time. A dedicated desktop or more powerful laptop with significant computing power and memory is recommended for running the iMotions software. It is recommended that that the computer where iMotions software is running should have at least an Intel i7 processor, have at least one terabyte of storage space, thirty-two gigabytes of random access memory, and a high end graphics card. Additional storage space in the form of an external RAID unit could also be used to store data. It is also advisable to store the videos of the respondents in the data space provided by university to conduct the study through iMotions because the external storage is very difficult to use. iMotions is an exceptionally powerful platform for biometric data collection, and should be evaluated in further studies at Pace University.

References