Human Emotion Recognition Using Fused Physiological Signals

Summary

A new way to recognize human emotions in a real-time setting is presented. This is facilitated by the development of a new method for fusing emotion-based physiological data into one signal. These new signals retain the important variance found in different physiological data. This is facilitated by normalizing the signals and weighting the importance of each based on the overall variance. These weighted signals are then summed together into a new representation of the emotion.

Description

This research relates to the theme of Empowering a Soldier's Success using artificial intelligence and intelligent systems by monitoring the emotional state of soldiers. It allows for real-time feedback from them giving vital information that has the potential to increase the soldier lethality and save lives. The purpose is to recognize human emotion using physiological signals including heart rate, blood pressure, respiration rate, and skin conductivity (EDA). Using a new fusion-based method we can represent emotions as new signals that contains these modalities. The fused signals can be used as input to machine learning classifiers to recognize 10 different emotions including happiness, surprise, sadness, startled, skepticism, embarrassment, fear, pain, anger, and disgust.

We developed a new method to fuse each of the signals into a new physiological signal that retains the important information for each. We tested on 140 subjects each having sequences of 10 emotions, each with 8 different physiological signals. Using these signals, we want to retain relevant temporal information. To do this, we first normalized all training data to the same units of time (5000 frames). This allows us to make direct comparisons between each of the signals. Once the signals are normalized, we then fuse them into one signal that represents each emotion. This is done for each subject and emotion for a total of 1400 newly fused signals (140 subjects * 10 emotions) compared to the original 11,200 signals (140 subject * 10 emotions * 8 signals). To fuse the signals, we keep the signals (out of the original 8), that display the most important information. We hypothesized that signals with low variance will not significantly contribute to an overall emotion recognition rate. To study this, we calculated the variance for each signal, which is then normalized between [0,1] where 0 is the lowest variance and 1 is the highest. This normalized variance is then used to weight each signals importance. Each signal frame is multiplied by the weight, and then each signal is summed together to create our new fused signal. Again, this is done for each subject and each emotion. Our fusion technique retains the most variance for all the signals. Given the 1400 signals we next hypothesized that each emotion can be represented by one signal (compared to 140). To do this, we took the average value for each emotion across time to create 10 fused signals. Each of these signals can accurately represent the displayed emotion, allowing for accurate emotion recognition.

To test this, we hypothesized that the signals vs. time can be used as features to recognize emotion (i.e. each frame of time in the signal can be used as 1 feature). Based on this hypothesis, we conducted the following experiments. First, we used 10-fold cross validation to randomly split the data into training and testing sets where 90% of the data was used for training and 10% for testing. We then trained random forest, SVM, and naïve Bayes classifiers. We achieved a correct classification rate of 86%, 88%, 86%

respectively. For our next experiment, we then split each signal into training and testing data where the first 2500 frames were used for training a random forest and the last 2500 were used for testing. This gave us 25000 frames for training and testing (2500* 10 emotions). Using this experimental design, we correctly recognized 97.8% of the instances across 10 emotions. These encouraging results show the potential to predict emotion as the data was split based on time. We are also interested in how accurate pain can be recognized compared to the other emotions. To test this, we split the data into 2 classes of pain and no pain. We conducted the same 10-fold cross-validation experiments using random forest, SVM, and naïve Bayes classifiers. Using this design, we correctly identified 100% of all pain emotions compared to no pain for all three classifiers. These experiments show our fused signals can accurately represent the differences found in each emotion, especially when recognizing pain vs. non-pain.

This research has practical applications in predicting soldiers' emotional state, most importantly pain. Using a wearable sensor, real-time feedback can be retrieved from the soldier. This feedback can include stress, if the soldiers have been wounded, and the general emotional state of how they are handling the current mission. This information can be useful to adjust the mission as needed or retrieve them from the battlefield. This data is also useful in training scenarios. It would allow the instructors to gain insight into the state of the solider and see if the training is eliciting the desired response.

A new way to recognize human emotions in a real-time setting is conveyed through this research. This is facilitated by a new method of fusing physiological data into one signal by using the signal variance to weight each and sum them together.

Biography

Shaun Canavan received his PhD in Computer Science from Binghamton University. He is an Assistant Professor in the Computer Science and Engineering Department at the University of South Florida. His research is on Affective Computing, HCI, Biometrics and VR/AR. He won an AWS Machine Learning Research Award for investigating multimodal human emotion analysis. He has 18 publications in top conferences and journals, as well as has a patent for his invention on estimating hand pointing direction. He is serving on multiple technical committees for conferences and journals and is demo chair for Face and Gesture 2019.

Diego Fabiano is currently an undergraduate student in Computer Science in the Computer Science and Engineering Department at the University of South Florida. He is currently enrolled in an accelerated program that combines undergraduate and graduate work. He will graduate with a master's in computer science in Spring 2019 with a focus on human emotion recognition.