**Paper:**

Carlos Busso, Zhigang Deng, Serdar Yildirim, Murtaza Bulut, Chul Min Lee, Abe Kazemzadeh, Sungbok Lee, Ulrich Neumann, and Shrikanth Narayanan. "Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information." Proceedings of the 6th International Conference on Multimodal Interfaces 2004, State College, PA, USA, October 13-15, 2004. New York, N.Y.: Association for Computing Machinery, 2004. 205-211 . Print.

**Summary:**

Emotion is expressed in different approaches. The paper analyzed human emotion expressions from two different aspects: facial expression and speech. For the facial part, they put 102 markers on a face and reduced the 102 markers into a 5 main areas: forehead,

eyebrow, low eye, right cheek, and left cheek area. Subsequently, Principal Component Analysis (PCA) method is used to reduce the number of features per frame down to a 10-dimensional vector for each area and then use the K-nearest neighbor classifier to classify different emotions.

**Applicability:**

After discussing with our group, we think that the approach that use markers to collect data and reduce to different feature vectors is very interesting and we have never practiced the similar approach in class before. It would be interesting to learn and practice, even though we still need to do a lot of additional research about the details of how they did the PCA approach. The paper describes that they used the K-nearest neighbor classifier which is one of the topics that we discussed in class before, so it will help us get a good handle of how we are going to process data and get the final result. Since their data handling has some relation with our course material, we think it is safe to also apply some other method that we learned in class, such as SVM, neural network, which may generate even better result than the original approaches that are suggested by the paper.

**Issue:**

There are couple of issues about the paper we may need to consider before we apply their approaches. First of all, the paper is more focused on live feed data instead of static images. So we may need to modify their approach in some way so that their methodology is more applicable for our project. Secondly, in their approach, they have a motion detector which can generate 4 more feature vectors. The 4 more feature vectors may have some potential to improve the detection precision. In order to get a similar precision as the paper we may need to come up with some innovative way to help us improve the detection result.

**Paper:**

Wang H, Huang H, Hu Y, et al. Emotion detection via discriminative kernel method[C]//Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments. ACM, 2010: 7.

**Summary:**

This paper is about combining two traditional facial expression recognition methods together: one is using static data extracted from one single face image, the other is using motion dependent data obtained from dynamic face image sequences. To integrate these two types of facial expression data, a hybrid kernel is used, such that the advantages of both of them are exploited. This method is called Discriminative Kernel Facial Emotion Recognition. In addition, they used Linear Discriminant Analysis to transform the two types of original data into two more discriminative lower-dimensional, which makes it more efficient and effective.

**Applicability:**

One application we can refer to in this paper is using static facial expression features. There are 22 feature points on a human face, also called as “landmarks”, such as eyes, eye corners, eyebrows, mouth corners, nose tip, and etc. These features are used to construct facial expression features in the sequel. In general, facial deformations caused by an emotion can be measured in terms of either angles or distances between certain facial feature points. To support a size invariant representation of facial data, angle metrics are preferable because they save the effort for face normalization that is necessary for distance based features. Moreover, typical angles show a large coincidence between different persons whereas typical distances vary considerably between different persons.

**Issues:**

Several issues need to be clarified here. One of the main issues is that they used a combined feature to improve accuracy in this paper. We probably don’t need to do this from the beginning. We will deal with static features first. Another one is associated with the first one, that they used a hybrid kernel to enhanced utilization of facial expression features. For only static features, we will use one single SVM first. Even though the paper is about something different from what we are going to create, it’s still of great value in reference. Future refinements can be done for our final project according to the paper, if there is time so to speak.

Submitted by:

Yang, Henrik, Francis