RECOGNITION OF FACIAL EMOTIONS STRUCTURES 
USING EXTREME LEARNING MACHINE ALGORITHM

Damodharam seenu1, B.Siva kesava reddy2, L.Balaji3
1(Dept of ECE, IFET College of engineering, Villupuram,India, damoseenu@gmail.com)
2(Dept of ECE, IFET College of engineering, Villupuram,India,sivakesavareddy4@gmail.com)
3(Asst prof, Dept of ECE, IFET College of engineering, Villupuram,India, balaji87.ram@gmail.com)

Abstract— This paper proposes an approach called Extreme Sparse Learning (ESL), which has the ability to jointly learn a dictionary (set of basis) and a non-linear classification model. The proposed approach combines the discriminative power of Extreme Learning Machine (ELM) with the reconstruction property of sparse representation to enable accurate classification when presented with noisy signals and imperfect data recorded in natural settings. Additionally, this work presents a new local spatio-temporal descriptor that is distinctive and pose-invariant. The proposed framework is able to achieve state-of-the-art recognition accuracy on both acted and spontaneous facial emotion databases.

Keywords— Extreme Sparse Learning, non-linear, spatio-temporal.

1. INTRODUCTION

- Facial emotion recognition in uncontrolled environments is a very challenging task due to large intra-class variations caused by factors such as illumination and pose changes, occlusion, and head movement.
- The accuracy of a facial emotion recognition system generally depends on two critical factors:
  1. extraction of facial features that are robust under intra-class variations (e.g. pose changes), emotions.
  2. design of a classifier that is capable of distinguishing different facial emotions based on noisy and imperfect data (e.g., illumination changes and occlusion).

Sparse representation is a powerful tool for reconstruction, representation, and compression of high-dimensional noisy data (such as images/videos and features derived from them) due to its ability to uncover important information about signals from the base elements or dictionary atoms. While the sparse representation approach has the ability to enhance noisy data using a dictionary learned from clean data, it is not sufficient because our end goal is to correctly recognize the facial emotion. In a sparse-representation-based classification task, the desired dictionary should have both representational ability and discriminative power.

2. RELATED WORK

The sparse representation of a signal over an over complete dictionary is achieved by optimizing an objective function that includes two terms: one that measures the signal reconstruction error and another that measures the sparsity. This objective function works well in applications where signals need to be reconstructed, like coding and denoising. On the other hand, discriminative methods, such as linear discriminative analysis (LDA), are better suited for classification tasks [1]. The textures are modelled with volume local binary patterns (VLBP), which are an extension of the LBP operator widely used in ordinary texture analysis, combining motion and appearance [9]. Facial expression reflects not only emotion but other mental activities such as those in cases of clinical approaches. This survey also deals with brief details of various approaches like optical flow method, local binary patterns, Pyramid of histogram of gradient (PHOG) and Local phase quantisation (LPQ) method, Facial action coding system (FACS) [10]. To learn the mappings that achieve pose normalization, we use a novel Gaussian Process Regression (GPR) model which we name Coupled Gaussian Process Regression (CGPR) model. Instead of learning single GPR model for all target pairs of poses at once, or learning one GPR model per target pair of poses independently of other pairs of poses, we propose CGPR model, which also models the couplings between the GPR models learned independently per target pairs of poses [11]. The emotion recognition problem from arbitrary view facial images, in this paper we propose a novel method based on the regional covariance matrix (RCM) representation of facial images [12]. Traditionally, facial emotion recognition systems have been evaluated on laboratory controlled data, which is not representative of the environment faced in real world applications. Although facial emotion recognition has been extensively studied in the past, most of the existing feature extraction approaches require frontal facial images and even small changes in facial pose may reduce their effectiveness. Only a few researchers have attempted to solve the facial pose challenge. To the best of our knowledge, none of the existing methods can learn a non-linear classifier in the context of simultaneous sparse coding and classifier training.

3. PROPOSED SYSTEM

The proposed approach combines the discriminative power of Extreme Learning Machine
(ELM) with the reconstruction property of sparse representation to enable accurate classification when presented with noisy signals and imperfect data recorded in natural settings. We propose a dictionary-based classification method called Extreme Sparse Learning (ESL) to recognize facial emotions in real-world natural situations. The proposed approach combines the discriminative power of Extreme Learning Machine (ELM) with the reconstruction capability of sparse representation. The proposed framework is able to achieve state-of-the-art recognition accuracy on both acted and spontaneous facial emotion databases.

Advantages

- Different facial regions detection.
- Reconstruct the original signals from noisy and imperfect samples.
- High accuracy.

Modules

1) Input Image
2) Sparse Representation
3) Feature extraction
4) Decision level

Module 1:
- Initially the keywords are extracted from the database. It is used to create a dictionary of keywords, which can be used in training. Image sensing can be defined as a technique that attempts to extract meaningful information from input image.
- It can be characterized as the process of analyzing images to extract information that is useful for different segmentation.

Module 2:
Sparse learning refers to a collection of methods to learning that seek a trade-off between some goodness-of-fit measure and sparsity of the result, the latter property allowing better interpretability.

Module 3:
- It is a pre-processing step to pattern recognition and machine learning problems. It is often decomposed into feature construction and feature selection.
- To describe the content of images, color and texture features are extracted in this work. Here they refer to color and texture descriptors of an image as its signature. To form a signature, color and texture segmentations are obtained for the input image.

Module 4:
- In our approach, the Extreme Sparse Learning to achieve good recognition accuracy on facial emotion databases.

4. IMPLEMENTATION

To systematically evaluated facial emotion recognition framework using different classes. Note that the proposed algorithms were implemented using MatlabR2012.
5. CONCLUSION

In this paper, we proposed a novel classification scheme called ESL, which is motivated by the recent advancements in the field of sparse representation and supervised dictionary learning. ESL incorporates reconstruction properties of sparse representation and discriminative power of a nonlinear ELM for robust classification. In addition, we proposed a novel OF-based spatio temporal descriptor for pose invariant facial emotion detection. We have performed extensive experiments on both acted and spontaneous emotion databases to evaluate the effectiveness of the proposed feature extraction and classification schemes under different scenarios. Our results clearly demonstrate the robustness of the proposed emotion recognition system, especially in challenging scenarios that involve illumination changes, occlusion, and pose variations.

REFERENCE