International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 29 Issue 4 – MAY 2022.

ROBUST REPRESENTATION AND RECOGNITION OF FACIAL EMOTIONS USING EXTREME SPARSE LEARNING

Mr.B.Karthikeyan
Assistant Professor
Department of Electronic Communication and Engineering
Velammal College of Engineering and Technology
Viraganoor, Madurai Dt, India

A. Mohana Priya
Department of Electronic Communication and Engineering
Velammal College of Engineering and Technology
Viraganoor, Madurai Dt, India

S.Shali
Department of Electronic Communication and Engineering
Velammal College of Engineering and Technology
Viraganoor, Madurai Dt, India

Abstract - Recognition of natural emotions from human faces is an interesting topic with a wide range of potential applications, such as human-computer interaction, automated tutoring systems, image and video retrieval, smart environments, and driver warning systems. Traditionally, facial emotion recognition systems have been evaluated on laboratory controlled data, which is not representative of the environment faced in real-world applications. To robustly recognize the facial emotions in real-world natural situations, this paper proposes an approach called extreme sparse learning, which has the ability to jointly learn a dictionary (set of basis) and a nonlinear classification model. The proposed approach combines the discriminative power of extreme learning machine with the reconstruction property of sparse representation to enable accurate classification when presented with noisy signals and imperfect data recorded in natural settings. In addition, this paper presents a new local spatio-temporal descriptor that is distinctive and pose-invariant. The proposed framework is able to achieve the state-of-the-art recognition accuracy on both acted and spontaneous facial emotion databases.

Keywords: Facial Emotion, Spatio-Temporal, Sparse Learning, Recognition.

1. INTRODUCTION

Digital media archives are increasing to colossal proportions in the world today, which includes audio, video and images An Image refers as a picture produced on an electronic display. A digital image is a numeric representation of a two-dimensional image. Digital image processing refers to processing of digital images by using digital computers. Nowadays, most of the applications prefer digitalized version, to reduce memory space. Lot of application depends on digital images. One of the important application is medical image processing. This paper presents a new local spatio-temporal descriptor that is distinctive and pose-invariant. The proposed framework is able to achieve the state-of-the-art recognition accuracy on both acted and spontaneous facial emotion databases. This paper presents a new local spatio-temporal descriptor that is distinctive and pose-invariant. The proposed framework is able to achieve the state-of-the-art recognition accuracy on both acted and spontaneous facial emotion databases.

2. Problem Source

Facial emotion recognition systems have been evaluated on laboratory controlled data, which is not representative of the environment faced in real-world applications. To robustly recognize the facial emotions in real-world natural situations, this paper proposes an approach called extreme sparse learning, which has the ability to jointly learn a dictionary (set of basis) and a nonlinear classification model. 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.

3. Related Work

The survey has proposed by Jain [1] an image processing method to extract paper currency quantity. The extracted ROI may be wormed with Pattern Recognition and Neural Networks matching method. First they obtain the image by easy flat scanner on glue dpi with an exacting size, the pixels level is place to attain image. A few filters are useful to extract denomination assessment of note. They employ dissimilar pixel levels in different quantity notes.

The paper was presented by Mirza and Nanda [2] a technique for validating paper currency of India. The technique employs four characteristics of paper currency plus identification mark, security thread, latent image and watermark. The scheme may extract the hidden features i.e. latent image and watermark of the paper currency. The anticipated work is an attempt to propose an approach for the characteristic extraction of Indian paper currency.

The review was presented by Chakraborty et al. [3] a widespread review of study on a assortment of developments in existing years in classification of currency denomination. A number of techniques applied by a diversity of researchers are proposed briefly in organize to evaluate the condition of art. In this paper the author also focusing primarily on currency detection system including different steps involved in it like image attainment, feature extraction and categorization system uses different algorithm.

The paper was demonstrated by Reel et al. [4] of the heuristic analysis of characters and a number of serial numbers of Indian currency notes to recognition of currency notes. To distinguish a character from a given currency image, there is require to extract feature descriptors of such image. As an extraction technique

4. Disadvantages of Existing System

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.

5. Methodology

MATLAB (matrix laboratory) is a multi-paradigm numerical computing environment and fourth-generation programming language. Developed by Math Works, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages, including C, C++, Java, and FORTRAN. Although MATLAB is intended primarily for numerical computing, an optional toolbox uses the MuPAD symbolic engine, allowing access to symbolic computing capabilities. An additional package, Simulink, adds graphical multi-domain simulation and Model-Based Design for dynamic and embedded systems. In 2004, MATLAB had around one million users across industry and academia. MATLAB users come from various backgrounds of engineering, science, and economics. MATLAB is widely used in academic and research institutions as well as industrial enterprises. MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. Typical uses include:

- Math and computation
- Algorithm development
- Modeling, simulation, and prototyping
- Data analysis, exploration, and visualization
- Scientific and engineering graphics
- Application development, including Graphical User Interface building

MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. This allows you to solve many technical computing problems, especially those with matrix and vector formulations, in a fraction of the time it would take to write a program in a scalar non interactive language such as C or FORTRAN.

6. Proposed Work

Our proposed dynamic descriptor is different from existing OF based representations in three aspects: (i) we propose a new set of spatio-temporal features to capture the dynamic information hidden in a flow field, (ii) the extracted features are encoded effectively to achieve pose-invariance, and (iii) only the statistics of the extracted features is retained as discriminative information for further processing.

Reading input Action Image

The objective of the present work is to develop a facial emotion recognition system that is capable of handling variations in facial pose, illumination, and partial occlusion.

Training Algorithm for Extreme Sparse Learning

A new classifier called Extreme Sparse Learning (ESL) is obtained by adding the ELM error term to the objective function of the conventional sparse representation to learn a dictionary that is both discriminative and reconstructive.

Optical flow correction for head movement Face Detection

The locations of 39 landmarks were extracted from an expressed facial image with arbitrary head pose. The coupled scaled Gaussian process regression model was then applied to normalize the facial pose. Although the model was trained based on only a few discrete head poses, the method has ability to deal with continuous head pose variations.

Fisher Discriminative Dictionary Learning (FDDL)

In this method, dictionary learning based on Fisher discrimination criterion is used to improve the classification performance. The method aims to learn a structured dictionary, where the criterion imposed on sparse coding causes the sparse coefficients to have small within-class scatter but large between-class variance.

Features Face Detection

Due to large pose variations, the face detection failure Rate is quite high.

Many video clips consist of more than one human subject, making it difficult to isolate the subject of interest.

There is a wide difference in the way that the same emotion is expressed by the various subjects. Some of emotions are very confusing and hard even for a human expert to identify correctly.

Features Nose tip to the midpoint Detection

While the nose tip is considered as the origin of the face coordinate system, the reference vector connecting the nose tip to the midpoint between the centres of the two eyes is considered as the positive y-axis

Centers of the two Eyes Detection

It consists of video sequences, where each sequence is labeled with one of the six basic emotions (joy, surprise, anger, fear, disgust, and sadness). Since the location of nose point is provided in the database, preprocessing involves only face alignment based on constant distance between the two eyes.

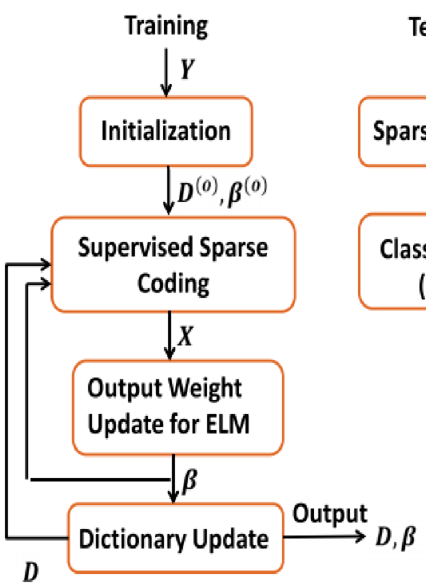
International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 29 Issue 4 – MAY 2022.

Features Uemo of Mouth region

For better illustration, we zoomed out the OF of the mouth region before and after correction In that the emotion-related OF (Uemo) is almost zero in the mouth region

Spatio-temporal descriptor construction feature extraction

A spatio-temporal descriptor is obtained by concatenating the spatio-temporal features extracted at each local region in the video. the construction of the spatio-temporal descriptor. The local regions are determined by dividing the volumetric data into M3D blocks (could be overlapping or non-overlapping) as shown in to preserve the geometric information of descriptors, each block is further divided into N 3D cells.



7. RESULTS AND DISCUSSION

1. Extreme Sparse Learning Processing

It is obtained by adding the ELM error term to the objective function of the conventional sparse representation to learn a dictionary that is both discriminative and reconstructive. This combined objective function (containing both linear and non-linear terms) is solved using a novel approach called Class Specific Matching Pursuit (CSMP). A kernel extension of the above framework called Kernel ESL (KESL) has also been developed.

2. Spatio – temporal Feature Extraction

The local regions are determined by dividing the volumetric data into M3D blocks (could be overlapping or non-overlapping). To preserve the geometric information of descriptors, each block is further divided into N 3D cells.

3. (ELM) Classification Facial Emotions

ELM is considered to be a state-of-the-art classification technique, especially for multi-class classification problems. ELM requires fewer optimization constraints in comparison to Support Vector Machines (SVM), which results in simple implementation, fast learning, and better generalization performance.

4. Emotional Count

The emotional count is used to detect the robust face representation and recognition of facial emotions.

5. Result Accuracy

It Shows the details about the sensitivity, specificity and recognition accuracy of the image details.

6. Time Calculation

This shows the computation time and number of features to recognize the facial image.

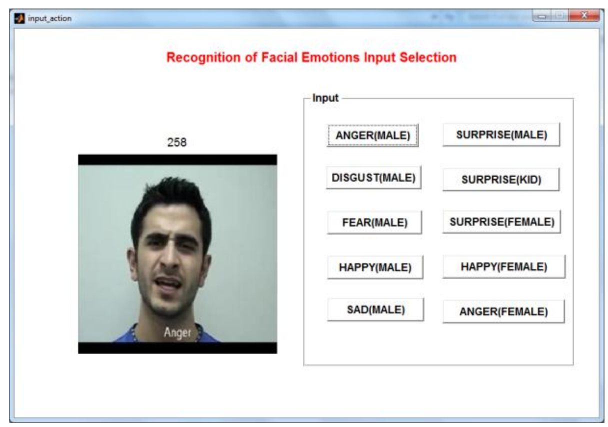


Fig. 2 Image Loaded

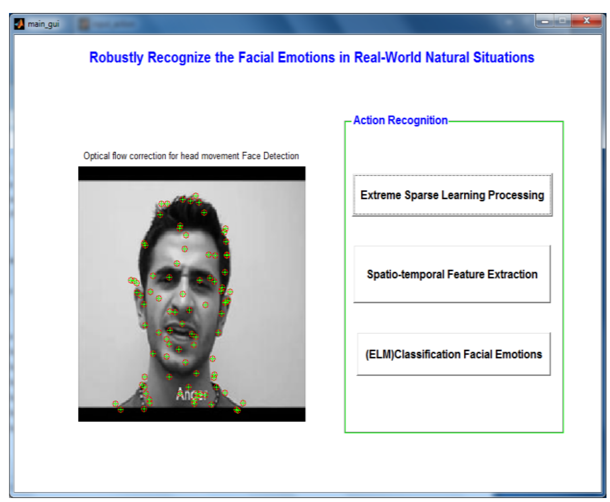


Fig3. Extreme Parse Learning Processing

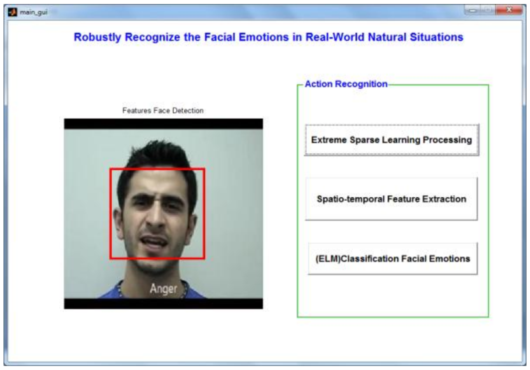


Fig4. Face Detection

Fig5. Eye and Mouth Distance Calc

ELM (Classification Facial Emotions)

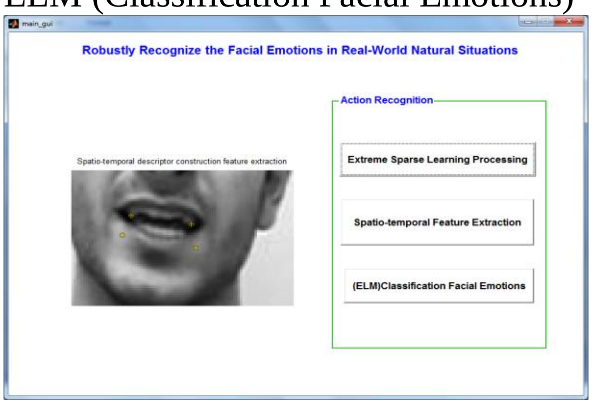


Fig6. ELM Classification Facial Emotion

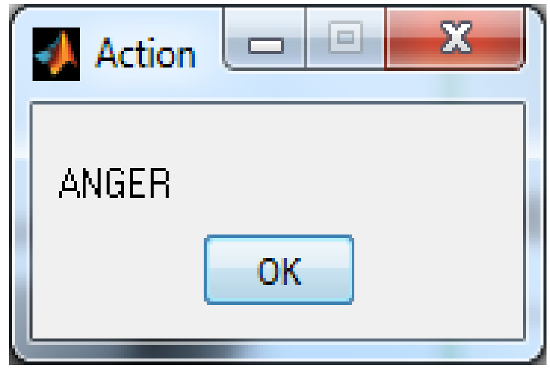


Fig7. Emotion Result

8. CONCLUSION AND FUTURE SCOPE

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 scenario

Furthermore, there is still a large room for improvement in the recognition accuracy when dealing with natural or spontaneous emotions. Possible ways to improve the proposed emotion recognition framework include: (i) combining the proposed spatio-temporal descriptor with static (appearance) based features to deal with failure in motion feature (e.g., optical flow) extraction, (ii) use of motion exaggeration techniques to improve the recognition accuracy for subtle facial emotions, and (iii) enhancing the OF correction model to remove the effect of facial muscle movement caused due to the person speaking

9. REFERENCES

[1] M. Liu, R. Wang, Z. Huang, S. Shan, and X. Chen, “Partial least squares regression on Grassmannianmanifold for emotion recognition,” in Proc. Int. Conf. Multimodal Interaction, 2013, pp. 525–530.

[2] M. Day, “Emotion recognition with boosted tree classifiers,” in Proc. 15th ACM Int. Conf. Multimodal Interaction, 2013, pp. 531–534.[39] T. R. Almaev, A. Yüce, A. Ghitulescu, and M. F. Valstar, “Distribution-based iterative pairwise classification of

emotions in the wild using LGBP-TOP," in Proc. 15th ACM Int. Conf. Multimodal Interaction, 2013, pp. 535–542.

[3] S. E. Kahou et al., "Combining modality specific deep neural networks for emotion recognition in video," in Proc. 15th ACM Int. Conf. Multimodal Interaction, 2013, pp. 543–550.

[4] T. Gehrig and H. K. Ekenel, "Why is facial expression analysis in the wild challenging?" in Proc. Emotion Recognit. Wild Challenge Workshop, 2013, pp. 9–16.

[5] K. Huang and S. Aviyente, "Sparse representation for signal classification," inProc. Adv. NIPS, 2006, pp. 609–616.

[6] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 2, pp. 210–227, Feb. 2009.

[7] D. H. Kim, S. U. Jung, and M. J. Chung, "Extension of cascaded simple feature based face detection to facial expression recognition," Pattern Recognit. Lett., vol. 29, no. 11, pp. 1621–1631, 2008.

[8] W. Gu, C. Xiang, Y. V. Venkatesh, D. Huang, and H. Lin, "Facial expression recognition using radial encoding of local Gabor features and classifier synthesis," Pattern Recognit., vol. 45, no. 1, pp. 80–91, 2012.

[9] T. Wehrle, S. Kaiser, S. Schmidt, and K. R. Scherer, "Studying the dynamics of emotional expression using synthesized facial muscle movements,"J. Personality Soc. Psychol., vol. 78, no. 1, pp. 105–119, 2000.