Optik 158 (2018) 1016–1025

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

Illumination invariant facial expression recognition using selected merged binary patterns for real world images

Asim Munir[a], Ayyaz Hussain[a,d,*], Sajid Ali Khan[b,c], Muhammad Nadeem[a], Sadia Arshid[a]

[a] Department of Computer Science & Software Engineering, International Islamic University Islamabad, Pakistan
[b] Department of Software Engineering, Foundation University Islamabad, Pakistan
[c] Department of Computer Science, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Islamabad, Pakistan
[d] College of Computing & Informatics, Saudi Electronic University, Riyadh, Kingdom of Saudi Arabia

ARTICLE INFO

Article history:
Received 22 December 2016
Received in revised form 22 May 2017
Accepted 2 January 2018

Keywords:
Facial expression recognition
Contrast enhancement
Local binary patterns
Local gradient coding
Merged binary pattern coding
Principal component analysis

ABSTRACT

Automatic facial expression recognition has always been a challenging task to understand human behavior from real world images. Certain type of issues are associated with such images that include poor illumination, different orientations and varying pose. The proposed technique first applies Fast Fourier Transform and Contrast Limited Adaptive Histogram Equalization (FFT + CLAHE) method to compensate the poor illumination. Then merged binary pattern code (MBPC) is generated for every pixel. Two bits per neighbourhood are produced to form a 16-bit code per pixel. This code merges local features to improve the effectiveness of facial expression recognition system. MBPC descriptor captures changes along fine edges and prominent pattern around eyes, eye brows, mouth, bulges and wrinkles of the face. The results of proposed technique are compared with different variants of LBP and LGC based techniques for both holistic and zoned images. Static Facial Expression in Wild (SFEW) dataset is selected for experimentation. Results clearly indicate that the suggested MBPC based technique surpasses other techniques with 96.5% and 67.2% accuracy for holistic and division based approach respectively. Moreover, results indicate that the performance of holistic approach is much higher than division based approach.

1. Introduction

A rapid knowledge about human behavior can be effectively gained by the facial expression. It is non-verbal means of communication among individuals and exhibits human emotions. Facial Expressions contain emotional signals and inner feelings of people about a subject that sometimes become complex to convey through the use of words and the way these words are vocalized. This states the significance of the human facial expression in the real-world situations where people express their own thoughts and feelings when communication with each other. Therefore, the challenging task of Facial Expression Recognition has attained academic value and has become the focus of research during recent years in the field of computer vision. Human emotions understanding by the means of computer with certain level of acceptable accuracy

* Corresponding author at: Department of Computer Science & Software Engineering, International Islamic University Islamabad, Pakistan.
E-mail addresses: asim@iiu.edu.pk (A. Munir), ayyaz.hussain@iiu.edu.pk (A. Hussain), sajidalibn@gmail.com (S.A. Khan), nadeem@iiu.edu.pk (M. Nadeem), sadia.arshid@iiu.edu.pk (S. Arshid).

https://doi.org/10.1016/j.ijleo.2018.01.003
0030-4026/

is a difficult task because these are produced by highly flexible deformations of the face elements. There is diversity in demonstration of same expression by different people in real world scenarios.

Facial expression recognition is easy to handle in controlled environment, however hard to handle in uncontrolled environment. Varying lighting conditions in uncontrolled environment is one of the factors which can cause unpredictable illumination effects on the face image. Normally, Histogram Equalization (HE) technique is utilized in pre-processing step to normalize the illumination effects and enhance the image contrast. HE adopts only a unique grayscale mapping process that does not consider regions with varying distributions of grayscales change.

Comprehending expressions require the extraction of facial features. Flexibility of these facial features restrict the accurate recognition of expression that is introduced even by a slight variation in light or pose resulting in local intensity variation. This paper attempts to identify the mental state through recognition of expressions in real world considering the challenges in the fields of biometric, security and human behavior identification through psychological study. Furthermore, robot vision, facial animation and virtual reality also require analysis of facial expression.

The Facial Expression Recognition is an extraction and classification process applied on the image or video to classify facial expression into angry, disgust, fear, happy, neutral, sad or surprise emotion by the use of computer. A typical Facial expression recognition system consists of expression image preprocessing, face detection and face region segmentation, expression feature extraction and expression classification.

There exist a variety of approaches for automatic facial expression recognition and most of these use synthetic dataset where the emotions are intentionally expression under controlled environments JAFFE [1], CK [2]. Ekman [3] defines that such dataset contain discretionary expression. In real world environment facial expressions are unconstrained with varying pose, different age, occlusions, non-uniform illumination and low resolution images thus making the task of emotion recognition real challenging as presented in SFEW database [4]. Such data is collected from sources like World Wide Web and TV broadcasts to represent the facial expression that have more importance in real world.

The approaches for expression recognition can be divided into two broad categories i.e. Geometric-based techniques and Appearance-based techniques according to definition of Ekman for expression recognition. In first approach, locations of landmarks are estimated to identify the landmark in the image and then the exact location of facial features' geometry and tracing of facial landmarks is achieved through the model [5,6]. A comparison of geometric based features is done with Gabor appearance based model [7] and results show that Gabor Wavelets being appearance based model generates better results.

The Gabor wavelet transform can detect the multi-scale, multi-direction of the texture changes, but it is time-consuming, and the characteristic dimension is huge, and the selection of dimensionality reduction methods also affects the facial expression recognition accuracy [8]. In another technique based on appearance model, more importance is given to pixel values rather than points [9] using Local Binary Patterns (LBP) [10]. LBP has been used with its variations for texture classification, image retrieval and face image analysis that can extract local features swiftly. But the traditional LBP operator applied to facial expression recognition is not perfect, it is by comparing the gray value of the center pixel and neighboring pixels to obtain the local texture feature, which does not accurately describe the texture. Uniform patterns are also suggest to reduce the dimensionality of the LBP features [11].Rotation invariant LBP is also implied that produces uniform pattern that helps in the reduction of rotations effect [12]. Another latest trend is to divides the face image into zones and applies LBP to obtain micropatterns from these zones [13]. Another method to obtain the micro expressions is to obtain LBP features from the vertical and horizontal projections that are used to record spatiotemporal information [14]. The robustness of LBP can be further increased by using Compound Local Binary Patterns (CLBP) that considers magnitude difference along with sign difference in the coding scheme to overcome the limitations associated with inconsistent LBP pattern [15]. Local Gradient Coding scheme also discovers the local features from face images better than of the traditional LBP algorithm, LBP uniform pattern, and Gabor filtering and improves accuracy of facial expression system for synthetic dataset [16].

In order to conserve local texture variations along with gradient difference, a novel method is proposed. Once face image is extracted after it is detected, this method first improves the contrast by converting the input image into frequency domain using Discrete Fourier transform (DFT) [17,18]. Then enhancement of the image is achieved using contrast limited adaptive histogram equalization (CLAHE) [19] in frequency domain. Prior application of this method makes face images suitable for feature extraction by compensating complex background and poor illumination that mainly occur due to use of accessories and/or uneven light reflections. MBPC extracts 16-bit code for a pixel using 2-bit code per neighbourhood. The least bit of MBPC holds local texture difference whereas higher bit specifies magnitude change of the gradient around neighbourhood of a pixel using a difference matrix. This practice helps in local and global illumination compensation. A powerful tool PCA is used to reduce the number of dimensions without much loss of features obtained in previous stage. The trained classifier is employed to test the sample face expressions.

A comparative study of proposed technique MSBP with LBP [10], LGCHD [16], LGCHVD [16] and CLBP [16] is performed in this paper. Experiments for all the techniques are conducted on static facial expressions in wild [4] that is more challenging due to variation and complexity of different parameters that are present in real world images and it dataset is not used earlier with LBP, LGCHD, LGCHVD and CLBP techniques. The results exhibit that MSBP provides 96.5% accuracy for full face images and 67.2% accuracy for division based approach which are improved when compared with state of the art techniques using SFEW dataset. Key contributions of this research are mentioned below.