Signal Processing 93 (2013) 1547–1556

Contents lists available at SciVerse ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Intelligent pixels of interest selection with application to facial expression recognition using multilayer perceptron

Taner Danisman*[,1], Ioan Marius Bilasco, Jean Martinet, Chabane Djeraba

Lille University of Science and Technology, CNRS UMR 8022, Parc Scientifique de la Haute Borne, 50 Avenue Halley, 59655 Villeneuve D'Ascq, France

ARTICLE INFO

Article history:
Received 25 January 2012
Received in revised form
3 August 2012
Accepted 12 August 2012
Available online 21 August 2012

Keywords:
Facial expression recognition
Multi layer perceptron
Feature selection

ABSTRACT

This paper presents an automatic way to discover pixels in a face image that improves the facial expression recognition results. Main contribution of our study is to provide a practical method to improve classification performance of classifiers by selecting best pixels of interest. Our method exhaustively searches for the best and worst feature window position from a set of face images among all possible combinations using MLP. Then, it creates a non-rectangular emotion mask for feature selection in supervised facial expression recognition problem. It eliminates irrelevant data and improves the classification performance using backward feature elimination. Experimental studies on GENKI, JAFFE and FERET databases showed that the proposed system improves the classification results by selecting the best pixels of interest.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Facial expression recognition (FER) is a hot research topic and a challenging problem on different domains including face recognition, human computer interaction, facial animation as well as social interaction. In the last decade, researchers from various disciplines focused on efficient, accurate and fast recognition of facial expressions. Emotions can be detected from physical sensors, image and video. Each sensor type has its own challenges such as noisy signals, high dimensionality and quality of selected features. There are many automatic FER studies achieving high accuracy on well-defined datasets. However, these studies still perform poor results under real world situations. Therefore there is still a considerable accuracy gap for realistic classification scenarios. One solution for this problem is to improve the classification results in terms of objective measures. Among others, feature selection is an important step towards better classifiers. Feature selection and reduction strategies are used to select relevant features to create robust models. In this scope, majority of the previous studies on FER considered the face and facial features as a combination of coarse rectangular units [1,2]. These units are used to locate or extract valuable facial feature information. Although its implementation simplicity, it includes useless and noisy data for the machine learning step. Therefore there is a need to find local pixel of interests (POI) to be used in FER. Group of POI provide non-rectangular masks that can be used to improve the classification performance.

Selection of the best variable and feature become the focus in classification research where there are thousands of different possibilities. Feature selection is the technique for selecting a subset of relevant features from original data to reduce feature size while maximizing the classifier output. Wrapper and filter based feature selection are the most common two approaches in the field. Wrappers evaluates the importance of specific features considering a particular learning algorithm [3] whereas filter based methods reduce the features space using a specific filter.

* Corresponding author. Tel.: +33 676303332; fax: +33 328778537.
E-mail addresses: danisman.taner@gmail.com, taner.danisman@lifl.fr (T. Danisman), marius.bilasco@lifl.fr (I.M. Bilasco), jean.martinet@lifl.fr (J. Martinet), chabane.djeraba@lifl.fr (C. Djeraba).
[1] Present address: University of Angers, LISA Laboratory, 62 Avenue Notre Dame du Lac, 49000 Angers, France.

0165-1684/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2012.08.007

Regardless of the fact that computational complexity of wrapper based methods put aside, the facial area used in FER is a small region that can be represented by regions as small as 20×20 to 50×50 pixels for vision based algorithms.

In this study, we used an analytic approach that performs wrapper based feature selection by exhaustive searching of all possible set of feature windows to find informative pixels to improve the results of FER. For a given emotion class, we created corresponding emotion mask to improve the Multilayer Perceptron (MLP) model's performance. Our experiments on different datasets showed that proposed method gives better results than full frame and the best traditional feature window based classification.

The rest of this paper is organized as follows. In Section 2, we briefly overview related works. Database material, mask generation and its application to FER is described in Section 3. Experimental results and discussion are presented in Section 4, followed by the conclusion.

2. Related works

Considering either analytic or holistic classification problem, there is a need to eliminate redundant and noisy information. Analytic approaches are widely use in face recognition domain and they are based on the detection of specific facial features such as eyes, eyebrows, nose, mouth and the locations of facial fiducial points such as corner positions of eye, mouth and their geometric relationships. Here a system is solved by considering its subparts and how they work together to produce particular results. For this reason, researchers use high level features of the face such as position of eyes, eyebrow, noise and mouth corners [4]. Among others, Nonnegative Matrix Factorization (NMF) is a popular dimension–reduction method that approximates a nonnegative matrix by the product of two other low-rank matrices. NMF distinguishes from others by its use of the non-negativity property. In contrast to holistic methods, NMF is able to learn parts of image. Local Binary Patterns (LBP) [5] is another analytic method used in texture classification and face recognition [6] which summarizes local structure of an image. LBP is a powerful gray-scale invariant texture feature providing more discriminative and invariant features for the recognition phase. Here the input face is divided into non-overlapping regions to compute an enhanced feature histogram.

Holistic approaches consider the encoding of the entire facial image into a point on a high dimensional space. In holistic methods, the problem cannot be solved by its subcomponents. Instead, the problem as a whole involves the solution. In the literature, Principal Component Analysis (PCA) [7], Linear Discriminant Analysis (LDA) [8] and Independent Component Analysis (ICA) [9] are common holistic methods extensively used for finding more informative features which reduce the dimensionality of original data by rejecting low variance features. For example PCA reduces the dimensionality by assuming that variance implies the importance and finds the reduced set of features that mimics the original data. However, ability to reduce original data is not appropriate for all problem scenarios. For the supervised FER problem we need local representative areas and pixels in a face image that can be used for the recognition.

Guan et al. [10] proposed an efficient online NMF variant called OR-NMF that can be applied to large-scale datasets. Slow convergence problem is addressed by use of NeNMF [11] and MD-NMF [12]. Choi et al. [13] analyzed the relationship between the input space and feature space using discriminant analysis and provided an input variable selection method. They selected variables that contain large amount of discriminative information and less discriminative information is discarded. Gong et al. [14] used LBP to obtain more complete description of the face using shifted and scaled sub-windows over the face images. Jeong et al. [15] propose a pattern recognition method and its application to face recognition problem. In their method, extracted features are analyzed in the original space using feature feedback and then they identify the important segments of the original data that affects the classification performance. Praseeda Lekshmi et al. [16] present a holistic method to analyze facial expressions by focusing on the regions such as eyes, mouth, etc. whose geometries are mostly affected by variation in facial expressions. They used PCA to recognize different parts of the face.

FER algorithms can be classified as feature-based and appearance-based methods. Feature based methods first detects facial feature points then classification is performed by considering geometric information. Appearance-based methods generally use the texture information for the classification. Hybrid methods like image ratio features [17] combine both the local texture information and the geometric information. Zhang et al. [18] studied geometry and gabor-wavelet based FER using MLP and observed that gabor coefficients are more powerful than geometric positions. Aroussi et al. [19] proposed an efficient local appearance feature extraction method based on Steerable Pyramid(S-P) wavelet transform for face recognition problem.

In a recent study, [20] showed that FER is an analytic rather than holistic process. They used an elliptical mask to remove useless background pixels from the face image as a preprocessing step before their classification method. This type of masks eliminates the appearance of the hair and neck from the face image as seen in Fig. 1. However, there exists many other non-informative skin pixels on this masked image which behaves like a noisy pixel in future processing steps. Therefore additional masks or methods are needed to eliminate these pixels during the preprocessing step. Fig. 2 shows different masks generated by using different feature selection algorithms for smiling and non-smiling classification.

Correlation-based Feature Subset Selection (CFS) shown in Fig. 2(A) uses a search algorithm along with a

Fig. 1. Elliptical mask for faces from JAFFE dataset [34].

Download English Version:

https://daneshyari.com/en/article/562617

Download Persian Version:

https://daneshyari.com/article/562617

Daneshyari.com