Applied Soft Computing 34 (2015) 72–93

Contents lists available at ScienceDirect

Applied Soft Computing

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

Intelligent facial emotion recognition using a layered encoding cascade optimization model

Siew Chin Neoh[a], Li Zhang[a,*], Kamlesh Mistry[a], Mohammed Alamgir Hossain[b], Chee Peng Lim[c], Nauman Aslam[a], Philip Kinghorn[a]

[a] Computational Intelligence Research Group, Department of Computing Science and Digital Technologies, Faculty of Engineering and Environment, University of Northumbria, Newcastle, NE1 8ST, UK
[b] Anglia Ruskin IT Research Institute, Faculty of Science and Technology, Anglia Ruskin University, Cambridge CB1 1PF, UK
[c] Centre for Intelligent Systems Research, Deakin University, Waurn Ponds, VIC 3216, Australia

ARTICLE INFO

Article history:
Received 6 September 2014
Received in revised form 15 March 2015
Accepted 5 May 2015
Available online 14 May 2015

Keywords:
Facial expression recognition
Feature selection
Cascade optimization
Layered representation structure
Evolutionary algorithm
Ensemble classifier

ABSTRACT

In this research, we propose a facial expression recognition system with a layered encoding cascade optimization model. Since generating an effective facial representation is a vital step to the success of facial emotion recognition, a modified Local Gabor Binary Pattern operator is first employed to derive a refined initial face representation and we then propose two evolutionary algorithms for feature optimization including (i) direct similarity and (ii) Pareto-based feature selection, under the layered cascade model. The direct similarity feature selection considers characteristics within the same emotion category that give the minimum within-class variation while the Pareto-based feature optimization focuses on features that best represent each expression category and at the same time provide the most distinctions to other expressions. Both a neural network and an ensemble classifier with weighted majority vote are implemented for the recognition of seven expressions based on the selected optimized features. The ensemble model also automatically updates itself with the most recent concepts in the data. Evaluated with the Cohn–Kanade database, our system achieves the best accuracies when the ensemble classifier is applied, and outperforms other research reported in the literature with 96.8% for direct similarity based optimization and 97.4% for the Pareto-based feature selection. Cross-database evaluation with frontal images from the MMI database has also been conducted to further prove system efficiency where it achieves 97.5% for Pareto-based approach and 90.7% for direct similarity-based feature selection and outperforms related research for MMI. When evaluated with 90° side-view images extracted from the videos of the MMI database, the system achieves superior performances with >80% accuracies for both optimization algorithms. Experiments with other weighting and meta-learning combination methods for the construction of ensembles are also explored with our proposed ensemble showing great adpativity to new test data stream for cross-database evaluation. In future work, we aim to incorporate other filtering techniques and evolutionary algorithms into the optimization models to further enhance the recognition performance.

1. Introduction

Facial expression is one of the effective channels to convey emotions and feelings. Research shows that facial expression contributed to about 55% effect of overall emotion expression during social interactions [1]. Many geometric-based, appearance-based or hybrid methods for automatic facial expression recognition have been proposed [2–4]. However, high feature dimensionality is a challenging issue for texture-based automatic facial expression perception. Many optimization algorithms have been proposed to reduce feature dimensionality [5–7]. However, it is difficult to select significant discriminating facial features that could represent the characteristic of each expression because of the subtlety and variability of facial expressions.

In this research, we propose a facial expression recognition system with a layered encoding cascade optimization model.

* Corresponding author. Tel.: +44 0191 243 7089.
E-mail addresses: siew.neoh@northumbria.ac.uk (S.C. Neoh), li.zhang@northumbria.ac.uk (L. Zhang), kamlesh.mistry@northumbria.ac.uk (K. Mistry), alamgir.hossain@anglia.ac.uk (M.A. Hossain), chee.lim@deakin.edu.au (C.P. Lim), nauman.aslam@northumbria.ac.uk (N. Aslam), philip.kinghorn@northumbria.ac.uk (P. Kinghorn).

http://dx.doi.org/10.1016/j.asoc.2015.05.006
1568-4946/

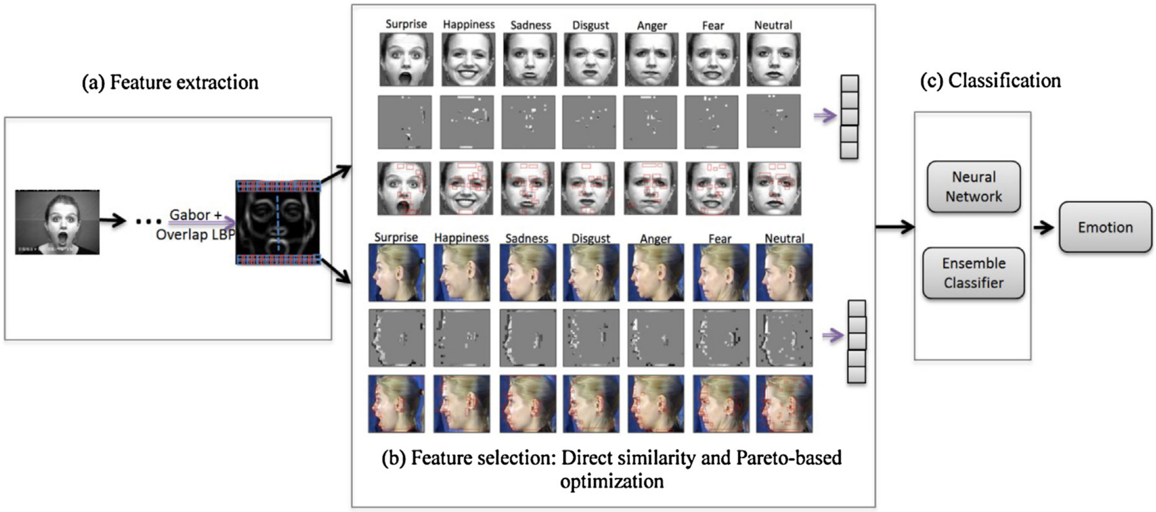

Fig. 1. The system architecture.

Especially, we propose two evolutionary optimization algorithms, i.e. direct similarity and Pareto-based feature selection, under the layered cascade model to extract the most significant facial features for each expression to benefit facial emotion recognition. The proposed direct similarity feature selection algorithm focuses on searching for common features purely from images within the same emotion category whereas the Pareto-based feature selection takes both within-class and between-class variations into account, i.e. the algorithm considers the capabilities of finding both similarity within one expression and the maximum differences to other emotion categories as two simultaneous optimization targets. Fig. 1 shows the system architecture.

Our system consists of three steps: (a) feature extraction, (b) feature dimensionality reduction and (c) expression recognition as shown in Fig. 1. First of all, (a) we implement a modified Local Gabor Binary Patterns (LGBP) with overlapping characteristics to extract the initial texture features in order to provide a more detailed facial representation. Then (b) the two proposed evolutionary algorithms are respectively used to select the most significant and discriminating facial features and to conduct feature dimension reduction for facial emotion recognition. Subsequently, (c) a neural network and an adaptive ensemble classifier are used to recognize seven expressions respectively including ‘happiness’, ‘fear’, ‘disgust’, ‘surprise’, ‘sadness’, ‘anger’ and ‘neutral’. To evaluate the efficiency of the proposed system, both evolutionary algorithms have been applied to frontal as well as 90° side-view images. Specifically, the system has been trained and tested with frontal images extracted respectively from the Cohn–Kanade extended (CK+) [8] and MMI facial expression databases [9] and 90° side-view images extracted from the videos of the MMI database [9]. In comparison with other typical approaches and the state-of-the-art research in the field, our system with both direct similarity and Pareto-based feature selection outperforms other research reported in the literature.

The research contributions of this research are summarized in the following.

1. A modified LGBP model that overlaps the corners of each sub-region in Local Binary Pattern (LBP) is proposed to retrieve more refined textural information for facial representation. The empirical findings indicate that it contains more discriminating power than the original LGBP for facial expression recognition.
2. Two evolutionary algorithms are proposed for feature optimization under a layered cascade evolutionary framework. The framework accords the feature selection process with great flexibility not only to separate facial features into specific areas for in-depth local search, but also to combine facial features for overall global search.
3. The direct similarity feature selection algorithm integrates the concept of micro Genetic Algorithm (microGA) [10] on top of the layered cascade framework to promote small population evolution and preserve search diversity with a non-replaceable memory. The algorithm focuses on finding common features within each emotion category to enhance the recognition capability of a specific facial expression.
4. Considering that misclassification often arises when the selected features do not possess sufficient distinctive characteristics among different emotion categories, the Pareto-based optimization algorithm is proposed to select features with maximal between-class variation as well as minimal within-class variation. This algorithm possesses more discriminative power in which more significant facial features for different emotion categories are revealed.
5. An ensemble classifier is implemented with the aim to create an improved composite model to improve classification accuracy. It employs the weighted majority vote for expression classification. This ensemble model consists of a series of classifiers complementing each other and shows great adaptive power to the diversity of the test data stream. Other weighting and meta-learning combination methods of ensembles are also compared and explored in order to further improve classification accuracy.
6. When tested with frontal images extracted from the CK+ and MMI databases, our system with the two proposed optimization algorithms outperforms other research reported in the literature. Since 90° side-view images may pose challenges to automatic facial expression perception because of significant information loss caused by head rotations, such images are extracted from videos of the MMI database to further test the efficiency of the proposed system. The results obtained from experimental analysis indicate that our system has a superior performance and achieves the best accuracy rates in comparison with other common approaches.

The paper is organized in the following way. Section 2 discusses related work including diverse feature selection techniques and their applications to face and facial expression recognition. Section 3 provides the system methodology for each key system stage. It includes the discussion of facial feature extraction using the extended modified LGBP in Section 3.1 and the theoretical