Biomedical Signal Processing and Control 11 (2014) 53–62

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Eye-state analysis using an interdependence and adaptive scale mean shift (IASMS) algorithm

Masrullizam Mat Ibrahim [a], John J. Soraghan [b], Lykourgos Petropoulakis [b,*]

[a] Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
[b] Centre for Excellence in Signal and Image Processing (CeSIP), Electronic and Electrical Engineering Department, University of Strathclyde, 204 George St., Glasgow G1 1XW, United Kingdom

ARTICLE INFO

Article history:
Received 17 July 2013
Received in revised form 20 January 2014
Accepted 17 February 2014
Available online 19 March 2014

Keywords:
Mean shift algorithm
Eye tracking
Eye state analysis
Adaptive scale
Fatigue assessment

ABSTRACT

Eye state analysis in real-time is a main input source for Fatigue Detection Systems and Human Computer Interaction applications. This paper presents a novel eye state analysis design aimed for human fatigue evaluation systems. The design is based on an interdependence and adaptive scale mean shift (IASMS) algorithm. IASMS uses moment features to track and estimate the iris area in order to quantify the state of the eye. The proposed system is shown to substantially improve non-rigid eye tracking performance, robustness and reliability. For evaluating the design performance an established eye blink database for blink frequency analysis was used. The design performance was further assessed using the newly formed Strathclyde Facial Fatigue (SFF) video footage database[1] of controlled sleep-deprived volunteers.

1. Introduction

This paper introduces a new approach which is capable of tracking and quantifying eye state in real-time. Through the innovative use of an adaptive scale mean shift tracking algorithm, the resulting system is also shown to be scale invariant. In general, eyes' state or blink detection methods can be classified into three categories: (i) knowledge-based methods [1] which encode human knowledge regarding the eye state; (ii) feature-based methods [2] which employ features from the eye such as colours, shapes, and edges; and (iii) appearance-based methods [1] which use the competency of machine learning algorithms to evaluate the eye state from specifically extracted features. There are several types of classifiers commonly deployed in these applications such as boosted classifiers, Support Vector Machines (SVM) and Neural Networks.

Knowledge-based methods typically use template matching algorithms using pre-prepared eye model templates. Chau et al. [3] and Wu et al. [4] created the open eye template, where the eye is already located, and they used it to track the eyes and the correlation of the template regions in between two frames. The state of the eye is determined by the computed correlation with a high correlation score signifying an opened eye and a low score for a closed eye. However, this approach has difficulties when it deals with multi-size of eyes and multi-pose of faces. In papers [5,6] statistical active appearance shape models are utilised in order to remedy the face multi-pose problem. This technique requires a set of the annotated eye open and closed images, which is applied in a statistical model training algorithm. The experimental results demonstrate that this technique is robust to the face multi-pose issues. Nevertheless, this technique encounters problems when it deals with face orientation changes such as face look up with partial head rotation.

The eyes have various features such as colour, texture and edges that can be utilised to evaluate the eye state. For instance, Liying et al. [7] use of skin colour to distinguish between eye open and close states [8,9], exploits the changes of intensity within the eye region as a method to detect the blink. Moris et al. [10] compute the variance map of the eye in order to obtain the distinctive features of the eye state. However, these approaches are particularly sensitive to illumination variations. Fast eye state detection algorithms reported in [11–13] analyse the movement of the eyelid using an optical flow algorithm. These techniques are effectively detecting the movement of the eyelid, but they are sensitive to face movements. Techniques that are based on appearance and use machine intelligence require a large number of images of eye states to train the classifiers. For example, Wang et al. [14] extract the eye features using the Gabor wavelet technique and classify them through

[1] The Strathclyde Facial Fatigue (SFF) video footage database was developed in collaboration with the Psychology Department, University of Strathclyde and the Sleep Centre, University of Glasgow, and it was approved by the Ethics Committee of the University of Strathclyde.

* Corresponding author. Tel.: +44 141 548 2935.
E-mail address: l.petropoulakis@strath.ac.uk (L. Petropoulakis).

http://dx.doi.org/10.1016/j.bspc.2014.02.007
1746-8094/

Neural Networks, whilst in [15,16] a closed eye state is determined using the Adaboost classifier [43]. Research in [17] determines the eye state based on rectangular eye position pixels; this position is trained and classified using SVM. The performance of this type of method relies on the features and the number of images available for training.

The mean shift tracking algorithm is a popular method of object tracking that is based on the probability of colour histogram [18,19]. This algorithm has been successfully applied to track distinctive colour objects in a variety of background scenes [20–22]. In the eye tracking technique, there are several examples which have applied mean shift to track the eye [23–25]. However, these eye tracking techniques are specifically based on the coloured region of the eye and lack any mechanisms to measure the eye state.

In this paper, a novel technique is introduced that integrates a conventional mean shift tracking algorithm with an adaptive scale scheme. This results in a system which improves on previous approaches as it is also able to measure the state of the eye during tracking in a scale invariant way. The resulting interdependence and adaptive scale mean shift (IASMS) algorithm tracks non-rigid eye movement by evaluating the eye position on the face and the distance between the eyes from successive frame sequences. The remainder of this paper is organised as follows: Section 2 provides an overview of the complete IASMS based eye state analysis system; Section 3 explains the IASMS algorithm in detail; Section 4 discusses the obtained experimental results; conclusions to the paper are provided in Section 5.

2. System design overview

The overall design, depicted in Fig. 1, presents four algorithm stages. The first stage is a front face detection process. The output of this stage contains distinct face features which are useful for estimating the position of the eyes, nose and mouth. The locations of the eyes are evaluated using a profile of intensity values projection technique as implemented in [26,27]. Following this, a classification stage is used to decide whether the eye is opened or closed. The third stage is an iris detection and localisation process. By implementing eye-open classification, the number of false-positive errors in the iris localisation algorithm is shown to be substantially reduced. The iris localisation process produces the location and size of the iris which is then used for tracking and quantifying the eye state in the IASMS algorithm. In the front face detection step of the design a hybrid method that comprises an association of the appearance based method with the features based method, is applied. This is a combination of a skin colour segmentation technique, connected components of binary images and a state of the art learning machine classifier [28,29] as indicated in Fig. 2. The skin colour segmentation method uses the combination threshold of multi-colour space value. In this approach, RGB, YCbCr and HSV colour spaces have been utilised. The output of a segmented colour skin image is in binary format (Fig. 2(a)), and consists of eight connected components for each of the segmented skin regions. From these components, a rectangular box is formed (Fig. 2(b)) in order to create a specific region of interest to be subsequently processed. Each of the rectangular boxes is then examined for shape and size to ensure that only boxes likely to contain face information remain for the next process. As outlined in [28] this is achieved using a Support Vector Machine (SVM) classifier and the horizontal projection features which, for a face, have a distinctive pattern.

The final step involves applying a Viola Jones [30] classifier on the remaining rectangular boxes to evaluate whether a box contains face information or not. This classifier is an integration technique between the AdaBoost learning algorithm [43] and a cascade classifier. This technique extracts the features vectors using a Haar-like features technique from integral images [30]. The hybrid method that is employed obviously can reduce the number of false-positive errors that normally occur in a complex background image.

When the front face is detected, the process to evaluate facial features is easier compared to the multi-poses face detection algorithm where not all facial features are visible. In such cases, the location of the eyes is identified from horizontal and vertical profile

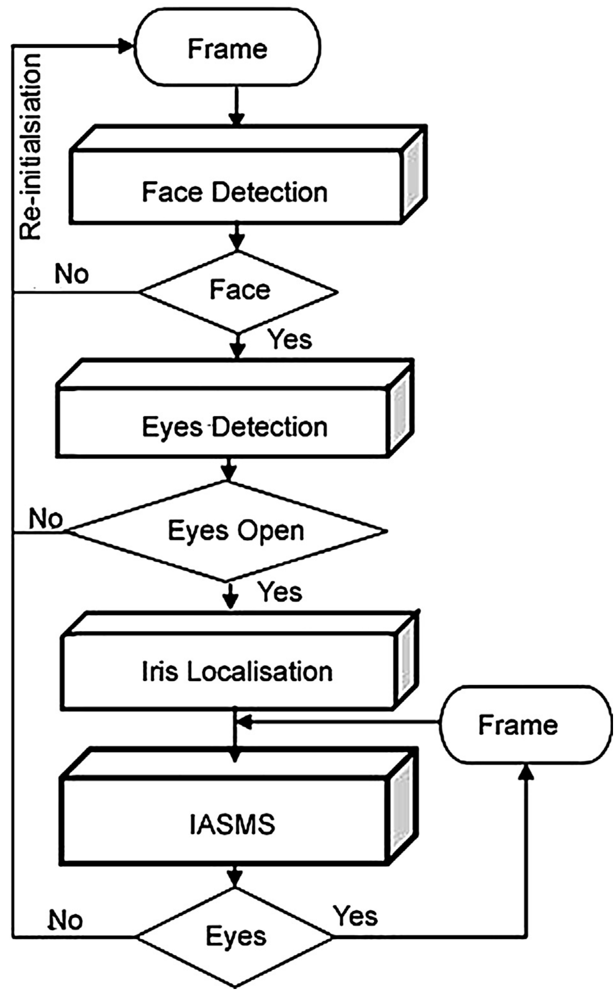

Fig. 1. Overall system design.

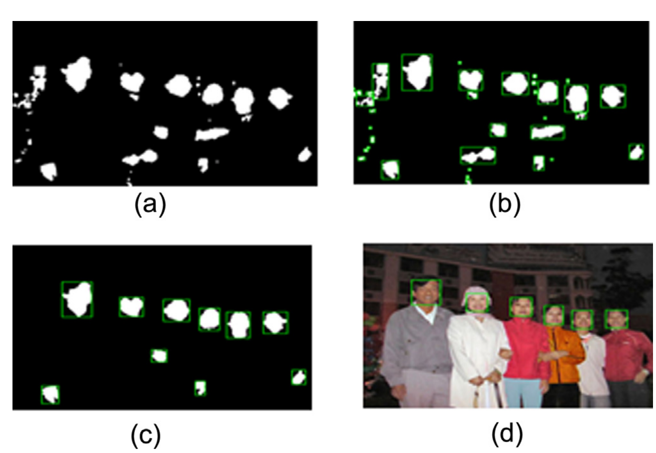

Fig. 2. Face detection algorithm: (a) skin colour segmentation; (b) rectangle bounding formation; (c) discarding boxes unlikely to contain face information; (d) SVM classification.