Pattern Recognition Letters 000 (2015) 1–14

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

A survey on periocular biometrics research☆

Fernando Alonso-Fernandez*, Josef Bigun

School of Information Science, Computer and Electrical Engineering, Halmstad University, Box 823, Halmstad SE 301-18, Sweden

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Periocular
Biometrics
Iris
Eye
Face

ABSTRACT

Periocular refers to the facial region in the vicinity of the eye, including eyelids, lashes and eyebrows. While face and irises have been extensively studied, the periocular region has emerged as a promising trait for unconstrained biometrics, following demands for increased robustness of face or iris systems. With a surprisingly high discrimination ability, this region can be easily obtained with existing setups for face and iris, and the requirement of user cooperation can be relaxed, thus facilitating the interaction with biometric systems. It is also available over a wide range of distances even when the iris texture cannot be reliably obtained (low resolution) or under partial face occlusion (close distances). Here, we review the state of the art in periocular biometrics research. A number of aspects are described, including: (*i*) existing databases, (*ii*) algorithms for periocular detection and/or segmentation, (*iii*) features employed for recognition, (*iv*) identification of the most discriminative regions of the periocular area, (*v*) comparison with iris and face modalities, (*vi*) soft-biometrics (gender/ethnicity classification), and (*vii*) impact of gender transformation and plastic surgery on the recognition accuracy. This work is expected to provide an insight of the most relevant issues in periocular biometrics, giving a comprehensive coverage of the existing literature and current state of the art.

1. Introduction

Periocular biometrics has been shown as one of the most discriminative regions of the face, gaining attention as an independent method for recognition or a complement to face and iris modalities under non-ideal conditions [52,74]. The typical elements of the periocular region are labeled in Fig. 1, left. This region can be acquired largely relaxing the acquisition conditions, in contraposition to the more carefully controlled conditions usually needed in face or iris modalities, making it suitable for unconstrained and uncooperative scenarios. Another advantage is that the problem of iris segmentation is automatically avoided, which can be an issue in difficult images [28].

This paper presents a survey of periocular research works found in the literature. We provide a comprehensive framework covering different aspects, from existing databases (Section 2), to algorithms for detection of the periocular region (Section 3), and features for recognition (Section 4). Databases utilized include face and iris databases (since the periocular area appears in such data), as well as newer databases capturing specifically the periocular area. Although initial studies have made use of annotated data, detection and segmentation of the periocular region has become a research target in itself. We also provide a taxonomy of the features employed for periocular recognition, which can be divided between those performing a *global* analysis of the image (extracting properties describing an entire ROI) and those performing *local* analysis (extracting properties of the neighborhood of a set of sparse selected key points).

Most recognition algorithms work by applying feature extraction and/or key points detection to a predefined ROI around the eye (Fig. 1, right). This holistic approach implies that some components not relevant for identity recognition, such as hair or glasses, might be erroneously taken into account [66]. It can also be the case that a certain feature is not equally discriminative in all parts of the periocular region. Some works have addressed these problems, as presented in Section 5. Since the periocular area appears in face and iris images, comparison and fusion with these modalities has been also proposed, which is the focus of Section 6. Besides personal recognition, a number of other tasks have been also proposed using features extracted from the periocular region. In this direction, Section 7 deals with issues like soft-biometrics (gender/ethnicity classification), and impact of gender transformation and plastic surgery on the recognition accuracy. We finally conclude the paper by highlighting current trends and future directions in periocular biometrics.

2. Databases

Table 1 summarizes the databases used in periocular research. Some sample images are shown in Fig. 2. Very few databases have been designed specifically for periocular research, with face and iris

☆ This paper has been recommended for acceptance by Maria De Marsico.

* Corresponding author. Tel.: +46 35 167304; fax: +46 35 120348.
E-mail address: feralo@hh.se (F. Alonso-Fernandez).

http://dx.doi.org/10.1016/j.patrec.2015.08.026

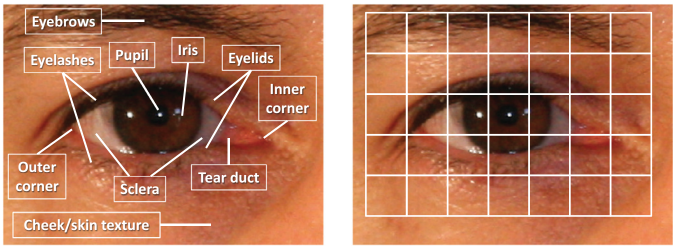

Fig. 1. Left: elements of the periocular region. Right: region of interest around the eye for feature extraction. Image from UBIRIS v2 database.

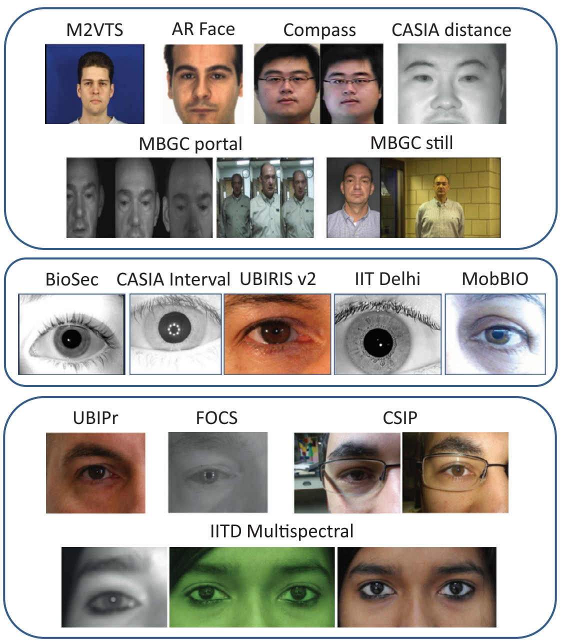

Fig. 2. Samples of databases used in periocular research. Top row: facial databases. Middle: iris databases. Bottom: periocular databases.

databases mostly used for this purpose. The 'best accuracy' shown in Table 1 should be taken as an approximate indication only, since different works may employ different subsets of the database or a different protocol. A general tendency, however, is that facial databases exhibit a better accuracy. These are the most used databases, so each new work builds on top of previous research, resulting in additional improvements. The accuracy with newer periocular databases are only some steps behind, demonstrating the capabilities of the periocular modality even in difficult scenarios, where new research leaps are expected to bring accuracy to even better levels. The following is a short description of each database, highlighting the features not contained in Table 1.

2.1. Facial databases

M2VTS has video of people counting '0'–'9' in their native language and rotating the head left-right. **AR** has frontal view with different expressions, illumination, and occlusions (sun glasses, scarf). **GTDB: Georgia Tech** has frontal/titled faces with cluttered background, four expressions and lightning/scale change. **Caltech** has frontal pose under with different lighting/expressions/backgrounds. **FERET: Facial Recognition Technology** has variations of illumination, expression, pose (frontal, left/right), race, glasses, etc. **CMU-H: CMU Hyperspectral** has videos in the range 450nm-1100 nm, in steps of 10nm. Three halogen lamps surrounding the face was used individually one at a time, and all together (four lightning conditions). **FRGC: Face Recognition Grand Challenge** has controlled/uncontrolled and 3D images. Controlled images were taken in a studio setting, and uncontrolled images in hallways, atria, or outdoors, with varying lightning and distance. **MORPH aging** (Album1) has scanned mug-shots taken between 1962 and 1998, with age of the subjects ranging 15–68 years old. The gap between first and last images is from 46 days to 29 years. Images are near-frontal, with many types of illumination and eye occlusions. **PUT** has partially controlled illumination, uniform background and pose variation. Most images have neutral expression, although a small set has no constraints on pose or expressions. **MBGC v2: Multiple Biometric Grand Challenge** is organized into 3 challenges: (*i*) Portal, (*ii*) Still Face and (*iii*) Video. Only *i* and *ii* have been used in periocular research. Portal data has subjects walking naturally through a portal, acquired simultaneously with NIR and VW video cameras. Therefore, many image perturbations appear. In the NIR sequences, some frames are too dark or too bright since the NIR lights shine only for a short time. Still Face data has high resolution images with controlled/uncontrolled illumination and frontal/non-frontal collected both in a studio environment and in hallways/outdoors. **Plastic Surgery** has one pre- and one post-surgery image for each person, both frontal, with proper illumination and neutral expression. **ND-twins** has images of twins under varying lighting (indoor/outdoor), expression (neutral/smile), and pose (frontal/non-frontal). **Compass** has four manners (neutral, smiling, eyes closed, facial occlusion) at two distances (10 m and 20 m) acquired with a pan-tilt-zoom (PTZ) camera. **FG-NET Aging** has subjects from multiple race, large variation of lighting, expression, and pose. The age range is 0–69 years, with images taken years apart. **CASIA v4 Distance** has high-resolution frontal NIR images with neutral expression acquired at ∼3 m. **FaceExpressUBI** has seven expressions, with location/orientation of the camera and light sources changed between sessions.

2.2. Iris databases

BioSec, **CASIA Interval v3** and **IIT Delhi v1.0** have NIR images acquired with close-up iris cameras. **UBIRIS v2** has VW images acquired between 3–8 m with a digital camera. The 1st session has controlled conditions, and the 2nd session was captured in a real-world setup (natural light, reflections, contrast change, defocus, occlusions, blur and off-angle). **MobBIO** has VW images from a Tablet PC with two lighting conditions, variable eye orientations and occlusions. Distance to the camera was kept constant. Annotation of the iris databases described, or a subset of them, have been made available [5,24].

2.3. Periocular databases

UBIPr was acquired with a digital camera, with distance, illumination, pose and occlusion variability. The distance varies between 4–8 m in steps of 1 m, with resolution from 501 × 401 pixels (8 m) to 1001 × 801 (4 m). **FOCS: Face and Ocular Challenge Series** has images from NIR videos of subjects walking through a portal (as in MBGC). A large number of images are of very poor quality, with high variations in illumination, out-of-focus blur, sensor noise, specular reflections, partially occluded iris and off-angle. The iris is very small (∼50 pixels wide). **IMP: IITD Multispectral Periocular** has three spectrums: NIR, VW, and Night Vision. The NIR dataset is created with a close-up iris scanner, the VW dataset with a digital camera at 1.3 m, and the night dataset with a handycam in night mode. **CSIP: Cross-Sensor Iris and Periocular** has images with four different smarphones. Ten different setups are included by capturing with both frontal/rear cameras and with/without the flash embedded in the device. The resolution of each camera is different, ranging from 640 × 480 to 3264 × 2448. Participants were captured at different sites with artificial, natural and mixed illumination. Noise factors include multiple scales, chromatic distortions, rotation, poor lightning, off-angle, defocus, and iris obstructions (including reflections).

Please cite this article as: F. Alonso-Fernandez, J. Bigun, A survey on periocular biometrics research, Pattern Recognition Letters (2015), http://dx.doi.org/10.1016/j.patrec.2015.08.026