Pattern Recognition Letters 000 (2015) 1–11

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Fusion of operators for heterogeneous periocular recognition at varying ranges☆

Zhicheng Cao*, Natalia A. Schmid

Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26505, USA

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Heterogeneous recognition
Cross-spectral
Periocular
Varying ranges
Short and long standoff distances
Local operators

ABSTRACT

Cross-spectral matching of active and passive infrared (IR) periocular images to a visible light periocular image gallery is a challenging research problem. This scenario is motivated by a number of surveillance applications such as recognition of subjects at night or in harsh environmental conditions. This problem becomes even more challenging with a varying standoff distance. To address this problem a new compound operator named GWLH that fuses three local descriptors – Histogram of Gradients (HOG), Local Binary Patterns (LBP) and Weber Local Descriptors (WLD) – applied to the outputs of Gabor filters is proposed. The local operators encode both magnitude and phase information. When applied to periocular regions, GWLH outperforms other compound operators that recently appeared in the literature. During performance evaluation LBP, Gabor filters, HOG, and a fusion of HOG and LBP establish a baseline for the performance comparison, while other compound operators such as Gabor followed by HOG and LBP as well as Gabor followed by WLD, LBP and GLBP present the state-of-the-art. The active IR band is presented by short-wave infrared (SWIR) and near-infrared (NIR) and passive IR is presented by mid-wave infrared (MWIR) and long-wave infrared (LWIR). In addition to varying spectrum, we also vary the standoff distance of SWIR and NIR probes. In all but one case of the combination of spectrum and range, GWLH outperforms all the other operators. A sharpness metric is introduced to measure the quality of heterogeneous periocular images and to emphasize the need in development of image enhancement approaches for heterogeneous periocular biometrics. Based on the statistics of the sharpness metric, the performance difference between compound and single operators is increasing proportionally with increasing sharpness metric values.

1. Introduction

Periocular recognition as a research topic has been known for several years [3,35,46]. It can be categorized as a part of face, perhaps most visible in the presence of facial occlusion. However, compared to other facial parts such as the forehead, nose and nose bridge, chin and cheeks, the periocular region contains a lot of fine textural and geometric information. Since the periocular region surrounds the area of the eye (but excludes the eye), when fused with the iris it can result in boosted matching performance (in visible and NIR spectra), especially when the quality of iris images is low [2,37]. Although many research challenges such as unconstrained subject's presentation, uneven illumination, and partial occlusions have been previously addressed in the literature [17,25,34,35,43], periocular biometrics faces many new challenges. As new practical applications evolve, new challenges offered by the applications arise and hence a need for development of new algorithms to mitigate them.

Surveillance at night or in harsh environments is one of the most recent applications. Latest advancements in manufacturing of small and cheap imaging devices sensitive in active infrared range (NIR and SWIR) [14,16] and the ability of these cameras to see through fog, rain, at night and operate at long ranges provided researchers with new type of imagery and posed new research problems [4,20,31,32]. As observed, active IR energy is less affected by scattering and absorption by smoke or dust than visible light. Also, unlike visible spectrum imaging, active IR imaging can be used to extract not only exterior but also useful subcutaneous anatomical information. This results in a very different appearance of images in active IR range compared to images in visible spectrum. Acknowledging these differences, many related questions can be posed. What type of information should be extracted from active IR images to successfully solve the problem of periocular recognition? How to match a periocular image in active IR or MWIR spectral band to a periocular image in visible spectral band? The latter falls in the scope of heterogeneous periocular

☆ This paper has been recommended for acceptance by Maria De Marsico Maria Frucci.
* Corresponding author. Tel.: +1 304 293 3017.
E-mail address: zcao1@mix.wvu.edu, zhichengtcao@gmail.com (Z. Cao).

http://dx.doi.org/10.1016/j.patrec.2015.10.018
0167-8655/© 2015 Elsevier B.V. All rights reserved.

ARTICLE IN PRESS
JID: PATREC [m5G;December 4, 2015;14:56]

Fig. 1. A block diagram of a typical periocular recognition system.

recognition. So far very few studies have been conducted on this topic [6,7,40]. Most relevant research work in the literature either studied periocular recognition within the same spectrum or fused information from different spectra. This paper is dedicated to this new topic and focuses on developing and combining local operators for heterogeneous periocular recognition.

A few publications on cross-spectral face recognition have appeared in the literature. Most of them were focused on algorithms for and analysis of matching NIR, SWIR, MWIR and LWIR face images to a gallery of visible face images [8,18,21,22,26,32,49]. Some publications assumed short standoff distances, whereas others explored the case of varying standoff distances [4,8,18,32]. Popular algorithms such as Local Binary Pattern (LBP), Scale-Invariant Feature Transform (SIFT), Histogram of Gradients (HOG), Gabor ordinal measures [9] and their variants have been used for feature extraction and matching in the past [47].

In this work, we describe a new composite operator that involves a bank of Gabor filters followed by the application of three local operators. They are WLD, LBP and HOG applied to magnitude and/or phase of the output of Gabor filters. The composite operator is used to encode heterogeneous periocular regions. We compare performance of the composite operator with the performance of (1) single local operators such as LBP, HOG, and Gabor filters, (2) a combination of LBP and HOG [45], and (3) two state-of-the-art composite operators such as Gabor filters followed by WLD, LBP and GLBP [32] and Gabor filters followed by LBP and HOG [41]. We further demonstrate that the new composite operator outperforms the other operators when applied to match NIR, SWIR, MWIR or LWIR probe images to a gallery of visible light periocular images. In addition to varying the spectral band of the probes, we also consider a short (1.5 m) and long (50 and 105 m) standoff distances in the case of NIR and SWIR. The poorer performance of the operators at the long standoff distances is linked to the lower quality of IR periocular images at that distances – a larger quality disparity between the IR and the visible images. We therefore suggest the need of enhancement on the IR images to reduce the disparity. We also show that the performance gap in matching heterogeneous periocular images increases as the quality of active IR periocular images increases. The quality of the images is measured in terms of the sharpness metric introduced by [48] applied to heterogeneous images.

The remainder of the paper is organized as follows. Functional blocks of the recognition system used in this work are described in Section 2. It also provides a brief summary of all single local operators involved in performance analysis as well as introduces the new compound operator GWLH. Section 3 characterizes datasets used for performance evaluation. Section 4 describes experiments and presents the results of our numerical analysis. A brief summary of the work is provided in Section 5.

2. Heterogeneous periocular recognition

A typical system for heterogeneous periocular recognition can be described by three consecutive blocks: a preprocessing block, a feature extraction block, and a matching block (See Fig. 1). In this work, the preprocessing block implements an alignment, cropping and contrast enhancement of heterogeneous periocular images. The feature extraction block performs filtering, encoding and converts encoded data into a histogram representation. The matching block applies I-divergence metric to the histogram representations of heterogeneous face images to generate a matching score [30,32].

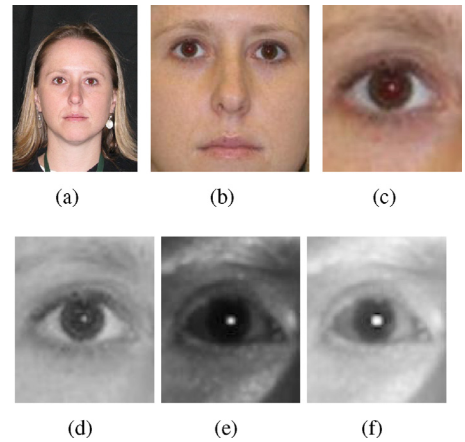

Fig. 2. Illustration of preprocessing: (a) original image, (b) aligned and cropped face, (c) aligned and cropped periocular region in visible light, (d) grayscale conversion of (c), (e) periocular region in SWIR, (f) logarithm-transformation of (f).

For alignment, all original images are mapped into a canonical representation resulting in images of size 120 × 112. The images are aligned using the position of the eyes – the pupils to be exact, which are manually selected. Geometric transformations such as rotation, scaling and translation are applied to each image such that the eyes are in fixed positions (See Fig. 2(a) and (b)). The aligned face images are then cropped to generate images of periocular regions for the usage in the following experiments. In this work, we crop the right eye from the face image with the pupil as its center (See Fig. 2(c)). For contrast enhancement, color images are first converted to grayscale images using a simple linear combination of the three channels. The transformed IR periocular images is further normalized between [0, 255].

The remainder of this section introduces our newly assembled compound operator composed of a bank of Gabor filters followed by the set of WLD, LBP and HOG applied to magnitude and (or) phase of Gabor filter outputs. We name the operator GWLH. Since the building of the new compound operator involves single operators, namely Gabor filter, WLD, LBP and HOG, we first provide a brief description of these single operators and then present a block-diagram explaining how these operators are fused to form GWLH.

2.1. Gabor filter, Weber local descriptor and local binary patterns

We begin with Gabor filters. Each normalized image (any spectrum) is passed through a bank of Gabor filters at 2 different scales and in 8 orientations, resulting in a total of 16 filter responses. A 2D Gabor filter is defined as

$$G(z,\theta,s) = \frac{\|\mathcal{K}_s e^{i\phi_\theta}\|}{\sigma^2}\exp\left[\frac{\|\mathcal{K}_s e^{i\phi_\theta}\|^2\|z\|^2}{2\sigma^2}\right]\left[e^{iz\mathcal{K}_s e^{i\phi_\theta}} - e^{-\frac{\sigma^2}{2}}\right], \quad (1)$$

where $z=(x,y)$, $\mathcal{K}_s e^{i\phi_\theta}$ is a wave vector and σ^2 is the variance of the Gaussian kernel. $\mathcal{K}_s$ is known as the scale parameter and ϕ_θ is the orientation parameter. The adopted parameters for the complex vector in the experiments of this paper are set to $\mathcal{K}_s=(\pi/2)^{s/2}$ with $s\in\mathbb{N}$ and $\phi_\theta=\theta\pi/8$ with $\theta=1,2,\ldots,8$. The Gaussian kernel has the standard deviation $\sigma=\pi$.

For WLD [11], the differences between the neighboring pixels of a central pixel are calculated and normalized by the pixel value itself. The summation of these normalized differences is further normalized by a monotonic function such as a tangent function. Finally, quantization is performed to output the WLD value. The mathematical