Engineering Applications of Artificial Intelligence 64 (2017) 112–127

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

# Using a Probabilistic Neural Network for lip-based biometric verification

Krzysztof Wrobel [a,*], Rafal Doroz [a], Piotr Porwik [a], Jacek Naruniec [b], Marek Kowalski [b]

[a] University of Silesia, Institute of Computer Science, ul. Bedzinska 39, 41-200 Sosnowiec, Poland
[b] Warsaw University of Technology, Faculty of Electronics and Information Technology, ul. Nowowiejska 15, 00-665 Warsaw, Poland

ARTICLE INFO

*Keywords:*
Biometrics
Lip
Image processing
Probabilistic Neural Network
Particle swarm optimization

ABSTRACT

In classical recognition techniques only raw features of objects are employed. Our approach allows use the composed features — so called *Sim* coefficients and landmarks which determine the area where biometric features should be searched. Biometric composed features are associated with appropriate similarity coefficients. Such approach brings significant advantages — recognition level of objects is higher compared to method based on the raw data. In this paper, a novel and effective lip-based biometric recognition approach with the Probabilistic Neural Network (PNN) is proposed. Lip based recognition has been less developed than the recognition of other human physical attributes such as the fingerprint, voice patterns, blood vessel patterns, or the face. For this reason, achieved results on this field are still improved and new recognition techniques are searched. Results achieved by PNN were improved by the Particle Swarm Optimization (PSO) technique.

In the first step, lip area is restricted to a Region Of Interest (ROI) and in the second step, features extracted from ROI are specifically modeled by dedicated image processing algorithms. Extracted lip features are then an input data of neural network. All experiments were confirmed in the ten-fold cross validation fashion on three diverse datasets, Multi-PIE Face Dataset, PUT database and our own faces dataset. Obtained in researches result show that proposed approach achieves an average classification accuracy of 86.95%, 87.14%, and 87.26%, on these three datasets, respectively. Announced results were verified in the comparative studies and confirm the efficacy of the proposed lip based biometrics learned by PSO technique.



## 1. Introduction

Technologies for cyber–physical security have achieved rapid growth over the past two decades. There is mounting evidence that they are important in biometric-based access control. Today, fingerprints, footprints, pistol bullets with their ballistic traits, but also hair, voice, blood, semen, DNA and the fibers from clothes are widely employed in crime detection and some of them can be used in biometrics. The biometric fields are increasing each year and this tendency is well visible due to need for security at borders, buildings, airports, corridors, banking transaction, etc., as a method of screening test and help for security service (Doroz et al., 2016; Porwik et al., 2014; Doroz et al., 2014; Dollár et al., 2010; Bedagkar-Gala and Shah, 2014). There is still much room for improvement with respect to recognition the new biometric features. One of these domains is human recognition through distinctive facial features. Recognition algorithms can be divided into two main approaches: photometric or geometric. Photometric algorithms try to overlay the pictures for a match, whereas geometric algorithms extract landmarks or other features from an image. Our strategy concerns the second type of such algorithms. We will focus on lip-based biometric, where digital photos including images of lip are analyzed (Aarabi and Mungamuru, 2004; Cao et al., 2014; Raheja et al., 2010; Howell et al., 2016). It has a great practical meaning because biometric modalities can be deployed in many areas including remote access control, secure special areas in airports, banks, to identify the cell phone users or people search. For this reason, analysis of various regions of images including human faces is still an important task and can improve identification effectiveness level. We noticed that features selection procedure can be linked with various similarity measures and classifiers (Doroz et al., 2016, 2014; Czabanski et al., 2012; Krawczyk and Wozniak, 2016). It has improved a machine-learning method of data classification, what is presented in this paper. Recognition of lip images can be conducted on the basis of groups of features — global and local lip features which are located inside of the contour of the lips.

Global features treat patterns as a whole, while local features are extracted from a limited area of pattern. A global feature is a contour of the lip image (latin: *rubor labiorum*). This contour comprises upper and

* Corresponding author.
*E-mail address:* krzysztof.wrobel@us.edu.pl (K. Wrobel).

http://dx.doi.org/10.1016/j.engappai.2017.06.003
Received 3 January 2017; Received in revised form 29 April 2017; Accepted 3 June 2017
0952-1976/

lower lip. Local features are determined on the basis of measurements of angles and distances between characteristic points of the lip. The lip contour can be considered as a unique signature of face, and we assumed that some characteristic points of the lip contour can be exploited as biometric features. In this paper all measurements on the shape of lip image contour and inside of contour area have been considered in the Cartesian coordinate plane.

## 2. Related works

Lip visual features are generally grouped into three categories: (a) appearance-based features; (b) shape-based features and (c) a combination of both appearance and shape features (Howell et al., 2016; Ibrahim and Mulvaney, 2015). Depending on the method of lip features extraction, researches are focused on the various domains. Part of the works concentrates around automatic speech recognition in human computer interfaces (Aarabi and Mungamuru, 2004; Ibrahim and Mulvaney, 2015), where image containing the lips region is processed and appropriate ROI is established. It was achieved by the Viola–Jones object detection framework. Later, the geometrical information obtained from the lips contour was used for the lips reading process. In the work (Newman and Cox, 2009) authors try to determine the language in a given person speaks. For this task the lip movements is registered during utterance of text. The authors use Active Appearance Models to locate the face and mouth, and build a vector that represents the lip shape for video frame.

Another field of investigations is recognition of facial expression (Raheja et al., 2010). In this method the lip contour is extracted from the object. In the next stage histogram of the lip contour is built and classified. It allows to achieve a lip gesture recognition rate of up to 95%.

Nowadays, global and local lip features are also used as genetic or morphological marker of ethnic differentiation in people populations in many countries (Domiaty et al., 2010; Kapoor et al., 2015; Rastogi and Parida, 2012). Separate researches are focused on medical treatments, where dental and surgical as well as post mortem analysis are preformed (Sharma et al., 2009; Utsuno et al., 2005). Some researches describe also biometric-based identification methods centered on static mouth or face images. It brings a new impulse to improve previously announced results.

There have been various investigations into recognizing a person directly from the lips shape and contour, where lip ROI is determined on the basis of colors distribution around the lip area (Choraś, 2010; Jin et al., 2007). It follows from the fact, that lip color usually has a stronger red component than other parts of the face. These principles were applied in Choraś (2010).

In the biometric system in the first step appropriate set of discriminative biometric features have to be determined and then these features can be classified (Bolle, 2004). The mouth area position on the face image can be sought by means of various methods (Ibrahim and Mulvaney, 2015): (a) matching technique; (b) feature invariant; (c) machine learning and (d) knowledge based.

Matching technique is known method used in image processing where correlation between a template image and an unknown image have to be determined and then classified. Such a technique has been used to detect the face (Ibrahim and Mulvaney, 2015; Jin et al., 2007; Marzec et al., 2015), as well as the eyes and mouth (Bartlett et al., 2006; Naruniec, 2014).

Invariant-based methods employed features which are stable, although some appreciable changes can be observed in the whole image such as its brightness or the pose of the subject, for example. For face, invariant features can be found by color analysis in the mouth ROI (Ibrahim and Mulvaney, 2015; Choraś, 2010; Jin et al., 2007). Machine learning methods are applied to recognizing the different face poses and head rotation registered in image video sequences. The most popular machine learning approaches are Artificial or Convolutional Neural Networks (Tadeusiewicz, 2015; Specht, 1990; Li et al., 2015), decision trees (Markus et al., 2013), the Viola–Jones object recognizer (Viola and Jones, 2004; Murphy et al., 2016) and HyperFace methods (Ranjan et al., 2016) for face detection and fiducial point extraction. The HyperFace detector automatically extracts all the faces from a given image.

Knowledge-based technique employs knowledge about human face and mouth location — the mouth is positioned in the lower part of the head below the nose and on a line of vertical symmetry and skin color (Ibrahim and Mulvaney, 2015). It is sufficient to determine the mouth ROI. The set of biometric features of a given object is mostly formed as a multidimensional normalized vector. It can be reduced (Porwik et al., 2014; Wang and Liew, 2012), and classified by means of the various methods such as dynamic time warping (DTW) (Ibrahim and Mulvaney, 2015), as well as combination of classifiers (Doroz et al., 2014; Kurzynski et al., 2016; Tadeusiewicz and Horzyk, 2014; Ooi et al., 2016). To the best of our knowledge lips biometrics for person verification is still developed domain. In this paper we show that combining face together with lip landmarks descriptors improving announced in the literature achievements and we demonstrate the validity of this arguments in performed experiments.

## 3. Proposed innovation

The surface of lips is covered by contours and numerous reticular depressions (grooves) that form system of lines. The lines of a lip pattern can be illustrated as shapes and can have various representations. As was previously stated, to date most of the existing work concentrates on lip movement of speaker or language recognition, where combination of lip movement and static lip features are analyzed. In such systems lip dynamics is sufficient to limited number of spoken words recognition (Aarabi and Mungamuru, 2004; Ibrahim and Mulvaney, 2015; Newman and Cox, 2009). This follows from the fact that this domain is still relatively new, being explored and developed only over the recent years. Popularity of lip print sensors is not yet high and investigations in this area are still being developed. It is the second reason why biometric systems based on lip analysis are not widespread.

It should be noted that also multimodal biometric systems are applied because high performance of the system can be easier to get, when various modalities will be employed. Designing of the biometric system with high accuracy is more difficult when only single modality is available. It is usually a more challenging task. In this work, we focus on the single modality approach, based on static shape information. These data can be collected either from photographs or video cameras as single video frames.

We present a novel biometric system based only on lip contours and new lip geometrical measurements. It means that in contrast to other methods texture of the lip surface is ignored. It simplifies computations without deterioration of the recognition quality. Lip extraction is possible through the two new independently working classifiers which operate on sets of the different features extracted from the image. Proposed verification model is trained by Probabilistic Neural Network (PNN) and optimized by Particle Swarm Optimizer (PSO) (Hashemi and Meybodi, 2011; Garg, 2016; Ali et al., 2015). It should be emphasize that the proposed technique achieves a classification accuracy of up to 87.26%. To the best of our knowledge, a study in lip-based biometrics using PNN + PSO methods has not been investigated before.

It is well known, that there are dozens of biometric modalities because any measurable biological or behavioral characteristic can be a biometric modality. In our case instead of face-based modality we propose lip-based modality. It causes that some problems which occur in the biometric systems based on face analysis can be eliminated. For example, if a given person wears glasses or beard, face recognition process can be unsettled in cases when other face image of the same person will be taken without glasses or facial hair. Our approach allows also correctly to designate all landmarks for ethnic faces because, in