DEEP FEATURES CREATION FOR SMILE CLASSIFICATION IN BIOMETRIC SYSTEMS

Boris Assanovich, Yury Veretilo
YK State University of Grodno
National Anti-Doping Laboratory
Grodno, Minsk, Republic of Belarus
E-mail: bas@grsu.by, y.veretilo@antidoping.by

The new face deep features obtained from video with use of autoencoder have been proposed and a smile imprint for the biometric cryptosystem was constructed. The performance of the proposed system has been verified with the use of posed smiles dataset from UvA-NEMO database.

Introduction

Recently, the analysis of human emotions has attracted more and more researchers. Often a person's smile is a key factor not only in determining the person's psychological state, but also has become one of the most preferred biometric methods both in video surveillance and in digital banking [1].

To ensure high accuracy of recognition and resistance to various attacks (cropping, spoofing, etc.), a number of authors switched to the analysis of the dynamic characteristics of a person's face or to soft biometrics. Recently Murat Taskiran et all [2] have applied dynamic face features extraction from videos and used them for face recognition. Authors performed the segmentation of smile phases from video and extraction of the main face landmarks allowed to recognize face with the equal error rate $EER = 30\%$ (threshold 2.8) for the spontaneous smile and with $EER = 21.25\%$ (threshold 2.39) for the posed one with the use of 648 feature vector. Different techniques use facial dynamics to identify a person with the use of various spatio-temporal parameters of face. Classical linear methods of image processing and feature extraction based on principal components analysis (PCA) are replaced by non-linear transformations. Compared with PCA, the use of an autoencoder significantly increases the classification accuracy.

In this paper, we will consider the use of auto-encoders to extract soft biometric data from a person's smile and use them for authentication and identification to provide access to digital services.

The article provides the analysis of classification efficiency of face deep features and demonstrates the creation of a digital smiling face imprint to provide a secure biometric interface.

I. AUTOENCODERS

Actually an autoencoder is the network that is forced to learn to select the most important features from data [3]. To exclude the process of overfitting in ML sparse autoencoders are used.

In our system we applied so-called stacked autoencoder (SAE) that is a neural network including several layers of sparse autoencoders where output of each hidden layer is connected to the input of the successive hidden layer. In this case the hidden layers are trained in an unsupervised way and then fine-tuned by a supervised method.

The use of SAE will allow the use of the trained data of the output layer of the neural network as features for biometric separation of users into genuine and imposters (persons who pretends to be somebody else).

The more the two distributions are separated and the smaller the standard deviation for each distribution, the better the separation of the classified classes. This property of distributions is estimated by such a parameter as decidability index

$$DI = \frac{|\mu_g - \mu_i|}{\sqrt{(\sigma_g^2 + \sigma_i^2)/2}}, \quad (1)$$

where μ_g, μ_i and σ_g, σ_i are the means and standard deviations of genuine and imposter distributions.

In addition to the decidability index, an equal error rate (ERR), which is the error rate at which a false accept rate (FAR) is equal to a false rejection rate (FRR), is normally used as a measure of biometric system verification accuracy. In biometrics FAR is the rate at which an imposter print is incorrectly accepted as genuine and FRR is the rate at which a genuine print is incorrectly rejected as imposter.

Then, a comparison of users by their feature representation can be performed by the evaluation of a receiver operator characteristic (ROC) curve that demonstrates how a genuine acceptance rate (GAR) determining a correctly received genuine fingerprint depends on a change in FAR.

Biometric authentication is the process of establishing user identity by measurements of his biological characteristics. To handle the variability inherent in biometric authentication, it is necessary to create and store a template for each user. To create it, one can use the fuzzy commitment scheme with application of error correcting codes (ECC) [4].

II. PROPOSED SYSTEM

As soon as autoencoders are widely used in many applications many researches have tested their efficiency. The experiments of face recognition based on SVM and softmax classifiers have shown that in recognition of 51 classes autoencoder achieved the error rate 9.53% while PCA 13.5% [5].

Inspired by a possible improvement in the classification characteristics, we used an auto-encoder to obtain biometric data on a person's smile and bind them to a secure user key. According to our knowledge, the use of auto-encoders to create a biometric imprint from face dynamics like a human smile has not been considered before.

Thus, the user smile imprint captured from video can serve as the biometric key to organize the access to different external digital services.

III. RESULTS AND CONCLUSIONS

In order to perform the features extraction and processing from captured images and video we developed the FaceAnalyzer Application (FAA) that has been applied to select and encode face HOG features, extract and normalize frames related to three main phases of video smile, calculate the statistical characteristic of biometric features.

The UvA-NEMO Smile Database has been used for experiments. It contained videos of spontaneous and deliberate smiles collected from 400 subjects.

Using FAA, video was analyzed with a person's smile and its three phases:

1. onset (neutral to expressive) phase;
2. apex;
3. offset phase (expressive to neutral), describing the state of a smiling person's face.

To compare the opportunities of binary and non-binary ECC by analogy with our earlier work [4], we extracted 4464-dimensional vectors from selected video smile frames, calculated their reliability and applied a 511-element mask for subsequent encoding with BCH codes (511,58,91) (511,28,111). The results showed a superposition of intraclass and intraclass distributions on each other and significant classification errors.

A series of experiments were performed with SAE to get good compact biometric features. To reduce time spent, in these experiments the subsets of 40 subjects randomly selected from the entire UvA-NEMO Database were used, reproducing a posed smile. Then normalized grayscale images from corresponding video of 112x112 pixels in size, scaled to 50%, creating a vector length of the input layer of 6272 elements have been used for unsupervised learning of SAE.

The training of the SAE layers consisted of 2 stages:

1. Unsupervised training on a 6x40 = 240 frames from training dataset, containing 3 phases of a smile and a neutral state;
2. Supervised training, when 4 additional video frames were added to SAE input from the control dataset for tuning testing.

Soft-Max classifier was used for test identification.

To assess the quality of training, on the basis of latent layer data y, such values as FRR, FAR, GAR, ERR and DI were calculated and the ROC-characteristic was monitored, as well as the values of MSE for controlling the intraclass and interclass distribution. The results of SAE training and tuning through the corresponding number of iterations are shown in Fig.1, and the histograms of the obtained distributions in Fig.2.

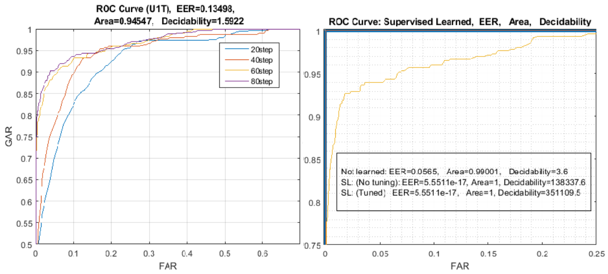

Fig. 1 – ROCs

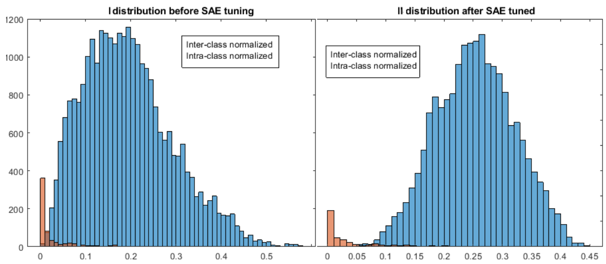

Fig. 2 – DIs

The experiments based on of UvaNemo and Caltech datasets performed have shown the improvement in the accuracy of classification genuine smile from posed one by 10%, as well as a reduction in the complexity of biometric templates design.

1. Cook, S. Selfie banking: is it a reality? Biometric Technology Today, 3(March). 2017 P. 9-–11.
2. Taskiran, M. et al. Face Recognition Using Dynamic Features Extracted from Smile Videos. In: IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), Sofia, 3-5 July 2019. P. 1-6.
3. Assanovich B. AutoEncoders for Denoising and Classication Applications. OSTIS 2020, M., 2020. – Вып. 4. – P. 309-312
4. Assanovich, B., Veretilo Yu. Biometric Database Based on HOG Structures and BCH Codes. In Proc. Information Techn. and Syst. BSUIR, 2017 P. 286-287.
5. Siwek, K., Osowski, S. Autoencoder versus PCA in face recognition. In: 18th International Conference on Computational Problems of Electrical Engineering (CPEE). Kutna Hora, 11-13 September 2017 . P. 1-–4.
6. Assanovich B., Bich N., Pronevich. A. Smile Biometric Imprint Creation with the Use of Autoencoder. ТСЗИ. – M., 2020 . – P. 8.