Authentication Based on Face Recognition under Uncontrolled Conditions

Niloofar Tvakolian
School of Electrical and
Computer Engineering
Shiraz University
Iran
Email: niloofar-ta@shirazu.ac.ir

Azadeh Nazemi
School of Electrical and
Computer Engineering
Shiraz University
Iran
Email: a.nazemi@cse.shirazu.ac.ir

Zohreh Azimifar
School of Electrical and
Computer Engineering
Shiraz University
Iran
Email: z.azimifar@cse.shirazu.ac.ir

***Abstract*—This paper describes a method to address issues regarding uncontrolled conditions in face recognition. This method using mask projection, extracts affecting factor from the test sample and adds it to all normal training samples then compares test sample with all synthetic affected training samples. The method has been applied for multi-factor authentication/verification based on face biometric. Obtained results indicate high accuracy in the lake of sufficient training samples for each class(single sample classes).**

Keywords*—*occlusion, face authentication, face identification, face verification, mask projection

I. Introduction

Generally, the Multi-Factor Authentication (MFA) system works based on the candidate identification through logging in with registered username and password (something the user knows) and recognizing him/her utilizing a biometric factor such as the face (something the user is). In the multi-factor authentication system, the user must present a proof of presence to avoid cheating and fraud. This system increases the layer of security to electronic authentication[**1**].
The main concern for identification based on face biometric is that face has been strictly affected by some uncontrolled conditions and is not reliable. This disadvantage leads to false penetration to the system in two ways:
False Match (FM) that allows a wrong person to access to system.
False Non-Match (FNM) that does not confirm a valid logged on person is same as claimed person.
MFA systems consist of four units: Registration, Confirmation ,Identification and Verification.

- Registration: in biometric systems each user has to enrol by creating an account to provide some personal information and a standard facial photo. This photo makes a template of face biometric and must be current, valid, and authentic.
- Confirmation means acceptance of the logged on information comparing to registered information.
- Identification means the image of the present user who is already logged in is matched with the image of the claimed user.
- Verification: Users can access to a remote online secure system as soon as be identified through their user name and password (E-authentication). The access to the system will be continued as long as successful verification results obtained by face recognition indicates that the logged in user is current, enrolled and constant during the access.

In the most cases making obligation for user to submit more than one normal frontal facial photo during registration/enrolment is not practical or ethics and causes to reduce system functionality. In the results training data set in MFA is single sample for each class and designing a face recognition system with a small size data set is required.

The rest of this paper was organized as follows:
Section II: Related Work
Section III: Methodology
Section IV: Experimental Results and Evaluation
Section V: Conclusion and Further Developments

II. Related Work

The most problematic challenges in face recognition is occlusion. Sparse Representation based Classification (SRC) methods claim for high accurate and robust face recognition under occlusion[**2**]. Gabor feature based sparse representation for face recognition (GSRC)[**3**] Extended SRC(ESRC)[**4**] Structured SRC (SSRC)[**5**] Structured Occlusion Coding (SOC)[**6**].

The following formulae are the basis of all stated sparse methods:
Assuming $d_{(i,j)}$ is a sample in training data set where
i: the lable of class in training data set
j: the lable of sample in class

$$D_T = [d_{(1,1)}, d_{(1,2)}, ..., d_{(l,k)}] \tag{1}$$

where
$k = number\ of\ samples\ in\ each\ class$

l =*the number of classes*
D_T represent training sample dictionary in SRC family that is generated by reshaping and concatenating all samples.
O denotes occlusion dictionary in SRC family methods which is created in offline mode with different approaches depends on method.
For non occluded testing sample :

$$y_{input} = D_T \alpha + \varepsilon \tag{2}$$

For occluded testing sample :

$$y_{input} = D\alpha + O\beta + \varepsilon \tag{3}$$

$error = \varepsilon = 0$

According to published results by these researches SOC is the most accurate one as it shown in fig1.[6]

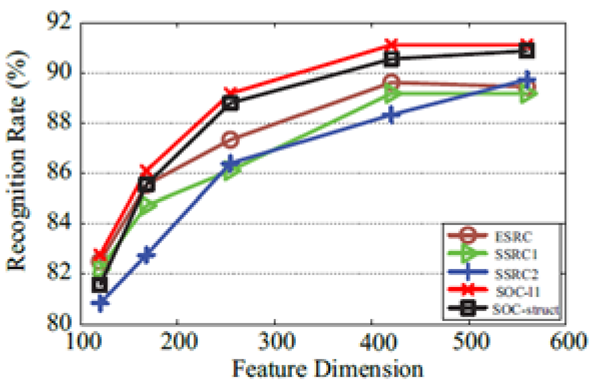

Fig. 1. Accuracy comparison in SRC family methods

III. METHODOLOGY

Generally, any face recognition system regardless being offline or online contains the following sub modules: Face detection, prepossessing, feature extraction/making descriptor and finally classification. Since this research makes effort to address occlusion issue then it supposes all training and testing samples are properly face detected, land marked and aligned.

During verification the system does not have any knowledge of type of occlusion thus making occlusion dictionary is not feasible in this research. Therefore, it assumes any input test sample is occluded and tries to extract occlusion mask from it.
Fig2 illustrates the overview of implementation this method.
For verification purpose, first RGB aligned detected face image must be converted to HSV. In the HSV representation of color, Hue determines the color, Saturation determines color intensity and Value determines the image lightness. Then to isolate the colors multiple masks have been applied to HSV image. The following snippet describes how to occlusion mask is extracted from input image[**4**][**8**].

T_i = trained front facial normal image. $i = 1, ..., l$ where l is number of classes in training Data set. R is RGB

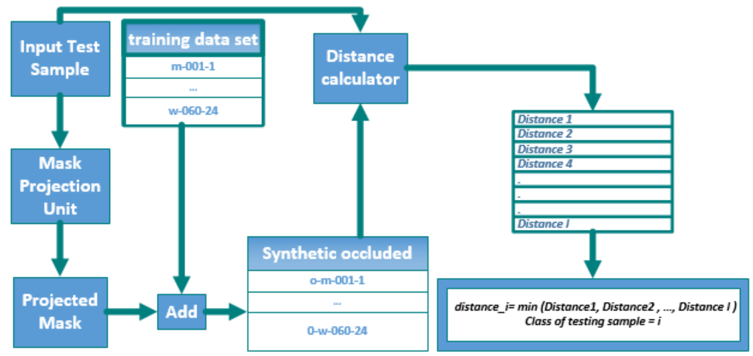

Fig. 2. proposed method overview

received rample image for verification captured automatically by assesement service request.

$Function\ Mask\ ExtractorR$

$R \rightarrow Mask(HSV)$

$Mask(HSV) \rightarrow Mask(RGB)$

$Mask(RGB) \rightarrow Mask(YCRCB)$

$Mask(R) = \frac{Mask(HSV)+Mask(RGB)+Mask(YCRCB)}{3}$

$Return\ Mask(R)$

In the next step occlusion mask will be applied to all members of all classes in training data set to generate synthetic occluded face. Finally, the distance between histogram of oriented gradients (HOG) features of input test sample and each synthetic occluded training data is calculated, the shortest distance represents the class of input sample. The following sppinet describes how this distance is measured.

$$Distance_{R1} = |MaskExtractor(R) - R|$$

$$Distance_{R2} = |MaskExtractor(T_i) - R|$$

$$Affected_{Received} = BitwiseOr(Distance_{R1}, Distance_{R2})$$

$$Distance_{T1} = |MaskExtractor(R) - T_{i)}|$$

$$Distance_{T2} = |MaskExtractor(T_i)) - T_i|$$

$$Affected_{(Normal_i)} = BitwiseOr(Distance_{T1}, Distance_{(T_2)})$$

$$Hog_{T_i} = HOG(HistogramEqualization(Affected_{(Normal_i)}))$$

$$Hog_R = HOG(HistogramEqualization(Affected_{(Recievesd)}))$$

$$EuclidianDistance_i = EuclidianDistance_{(Hog_{T_i}, Hog_R)}$$

$1 <= i <= l$ represents the class of testing image if $EuclidianDistance_i$ is minimum.
In the following formulae $\forall i = 1, ..., l; T_1, T_2, ..., T_l$ denote all training samples $m_1, m_2, , m_l$ represent mask of train samples. m_R is mask of received image (mask of testing samples).

$$d_{(1T_i)} = T_i - m_i \tag{4}$$

Fig. 3. T_i, m_i and $d_{(1T_i)}$ from left to right. related to equation(5)

$$d_{(2T_i)} = T_i - m_R \tag{5}$$

Fig. 4. $T_i, m_R and d_{(2T_i)}$ from left to right. related to equation(6)

$$d_{1R} = R - m_R \tag{6}$$

Fig. 5. $R, m_R and d_{(1R)}$ from left to right. related to equation(7)

$$d_{(2R)} = R - m_i \tag{7}$$

Fig. 6. $R, m_i and d_{(2R)}$ from left to right. related to equation(8)

$$\begin{aligned} & d_{(1T_i)} OR d_{(2T_i)} \\ & = \sum_{n=0}^{[log_2(x)]} [2^n [[(\left\lfloor \frac{d_{1T_i}}{2^n} \right\rfloor mod2) + \left\lfloor \frac{d_{2T_i}}{2^n} \right\rfloor mod2) \\ & \quad + (\left\lfloor \frac{d_{1T_i}}{2^n} \right\rfloor mod2) + (\left\lfloor \frac{d_{2T_i}}{2^n} \right\rfloor mod2)] mod2]] \end{aligned} \tag{8}$$

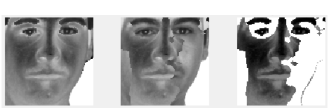

Fig. 7. cd˙(1T˙i), d˙(2T˙i) and d˙(1T˙i)ORd˙(2T˙i) from left to right. related to equation(9)

$$\begin{aligned} & d_{(1R)} OR d_{(2R)} \\ & = \sum_{n=0}^{[log_2(x)]} [2^n [[(\left\lfloor \frac{d_{1R}}{2^n} \right\rfloor mod2) + (\left\lfloor \frac{d_{2R}}{2^n} \right\rfloor mod2) \\ & \quad + (\left\lfloor \frac{d_{d_1}}{2^n} \right\rfloor mod2) + (\left\lfloor \frac{d_{2R}}{2^n} \right\rfloor mod2)] mod2]] \end{aligned} \tag{9}$$

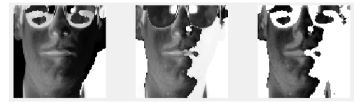

Fig. 8. d˙(1T˙i), d˙(2T˙i) and d˙(1R) OR d˙(2R) from left to right. related to equation(10)

$$\begin{aligned} & D_i = EuclidianDistance \\ & (HistogramEqualization(9), HistogramEqualization(10)) \end{aligned} \tag{10}$$

visually illustrates formula (11). Fig8 the demonstrates D_i value in different condition supposing D_i is distance of HOG feature vectors (1x64) rather than intensity values.

Another approach to measure distance is uesing Gabor fil-

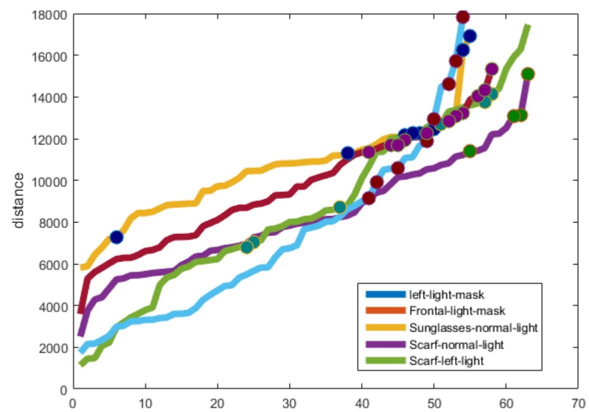

Fig. 9. comparison of distance between testing sample and synthetic occluded training sample in various condition using HOG feature vector

tering as it shown in fig9. As it observed from fig2 and fig3

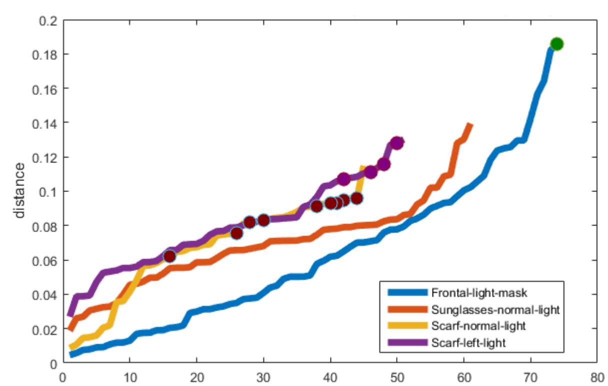

Fig. 10. comparison of distance between testing sample and synthetic occluded training sample in various condition using Gabor feature vector

Gabor has better result than HOG. The dots on graph represent non correct recognition.

IV. Eperimental Result And Evaluation

To evaluate proposed method a subset of AR data set includes 74 classes was used.[**9**] Each class contains normal, expressed, illuminated, occluded by sunglasses and scarf images, but this research considers only normal images as training samples and uses others as testing samples

Fig10 shows the difference between color histogram of the testing sample in various conditions and the nearest synthetic occluded training sample. It respectively indicates expression, illumination, sunglasses and scarf occlusion from left to right. When the first row of figure is color histogram of original occluded testing sample and the second row is the nearest color histogram of synthetic occluded training sample.

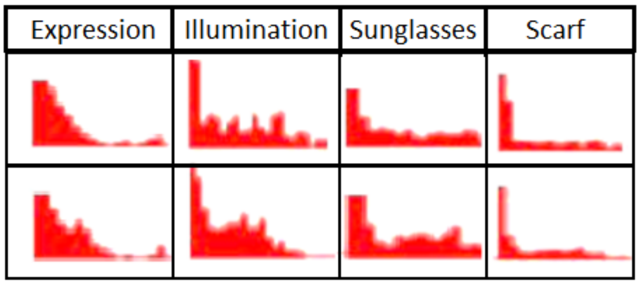

Fig. 11. the first row illustrates color histogram of original occluded testing sample and the second one illustrates the nearest color histogram of synthetic occluded training sample.

This section demonstrates some experimental results of implementation proposed method. As stated before in this study all uncontrolled conditions can be considered as occlusion. Therefore and fig8 is used to illustrate comparison the relation between recognition accuracy and the percentage occlusion in different conditions added some coordinate value to these figures. It can be concluded however it seems that illumination affects more than others but its accuracy is still high. By the result the worst case scenario is related to occluded face by scarf and affected by illumination at the same time.

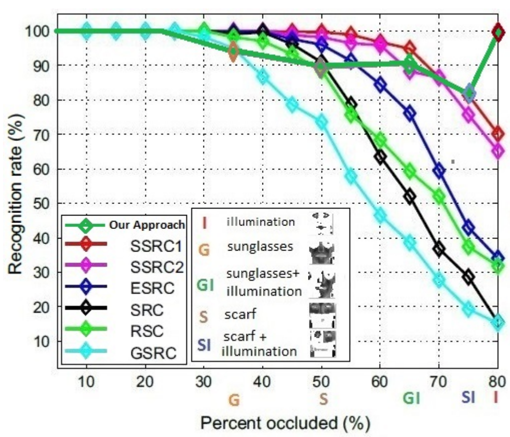

Fig. 12. recognition rate of proposed method on AR database

V. Conclusion

The research undertaken is part of ongoing study to address the issues regarding occlusion in real time face recognition in small sample size data set. The proposed method is based on finding occlusion mask, adding it to normal training face in data set and measuring the distance between synthetic and original occluded images. Further development will conduct to design a deep learning network for this purposes.

References

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.

References

[1] Weihong Deng, Jiani Hu, and Jun Guo. "Extended SRC: Undersampled face recognition via intraclass variant dictionary". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34.9 (2012), pp. 1864–1870.

[2] Tan Guo, Xiao Heng Tan, and Chen Xie Chao. "Robust Face Recognition via Occlusion Detection and Masking". In: *MATEC Web of Conferences*. Vol. 44. EDP Sciences. 2016.

[3] Anil K Jain, Arun Ross, and Salil Prabhakar. "An introduction to biometric recognition". In: *IEEE Transactions on circuits and systems for video technology* 14.1 (2004), pp. 4–20.

[4] Hongjun Jia and Aleix M Martinez. "Face recognition with occlusions in the training and testing sets". In: *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on*. IEEE. 2008, pp. 1–6.

[5] Weihua Ou et al. "Robust face recognition via occlusion dictionary learning". In: *Pattern Recognition* 47.4 (2014), pp. 1559–1572.

[6] Yandong Wen et al. "Structured occlusion coding for robust face recognition". In: *Neurocomputing* 178 (2016), pp. 11–24.

[7] John Wright et al. "Robust face recognition via sparse representation". In: *IEEE transactions on pattern analysis and machine intelligence* 31.2 (2009), pp. 210–227.

[8] Meng Yang and Lei Zhang. "Gabor feature based sparse representation for face recognition with gabor occlusion dictionary". In: *Computer Vision–ECCV 2010* (2010), pp. 448–461.