Facial Skin Texture as a Source of Biometric Information

Garsah Farhan Al-Qarni and Farzin Deravi

School of Engineering and Digital Arts, University of Kent
Canterbury, CT2 7NT, UK
{gfa4,f.Deravi}@kent.ac.uk

Abstract

This paper investigates the possibility of exploiting facial skin texture as a source of biometric information to facilitate automatic recognition of individuals. Such ability may be particularly important in circumstances when a full view of the face may not be available. The proposed algorithm automatically segments the forehead region and divides it into non-overlapping patches. Two state-of-the-art families of texture feature extraction approaches, namely Gabor wavelet filter and Local Binary Pattern operator, are compared for extracting features from these patches which are classified using a k-NN classifier. The identification and verification performance is evaluated for different patch sizes using the XM2VTS database. For the verification experiments an EER of 0.065 using Gabor features and 0.083 using LBP features is obtained for forehead regions with pure skin. Additionally a novel classifier is presented for automatically detecting pure skin patches in the forehead region.

Keywords: *Skin texture, Gabor filter, Local Binary Pattern, k-Nearest Neighbor*

1. Introduction

There have been a number of research studies utilizing human skin texture for facial synthesis, facial feature tracking and automatic dermatological diagnosis. However, there has been relatively little work reported in the literature on investing skin texture for person recognition.

Lin, et al., [1] presented a multilayer framework for face recognition with high-resolution images. Each face image was factorized into four layers: global appearance, facial organs, skin regions, and irregular details. In the third layer, a new skin texture representation was used based on texton distributions. The procedure comprised three stages: Filtering, Dictionary Building, and Discriminant Learning. Their technique was tested on the XM2VTS [2] database and showed improvements in matching accuracy when incorporating skin texture.

Some researchers have focused their work on studying only skin marks. Pierrard, et al., [3] presented a method to extract irregularities in facial skin, in particular nevi (moles, birthmarks), using normalized cross correlation matching and a morphable model. Lee, et al., [4] proposed a content-based image retrieval system based on SIFT features, for matching and retrieval of tattoo images which are considered as soft biometrics [5]. In Jain, et al., [6], an Active Appearance Model was used to locate and segment facial organs. Then, Laplacian-of-Gaussian and morphological operators were used to detect facial marks.

This paper focuses on the extraction of useable forehead regions from facial images and their initial pre-processing to provide useful identity information. The paper is organized as follows: The proposed algorithm is described in Section 2. The skin/hair classifier that will be able to automatically detect skin areas is presented in Section 3. The experimental results based on the XM2VTS database are reported and discussed in Section 4, and Section 5 provides conclusions and suggestions for future work.

2. Recognition Based on Local Skin Patches

The overall block-diagram of the proposed system is shown in Figure 1. This section will give step-by-step illustrations of the proposed algorithm.

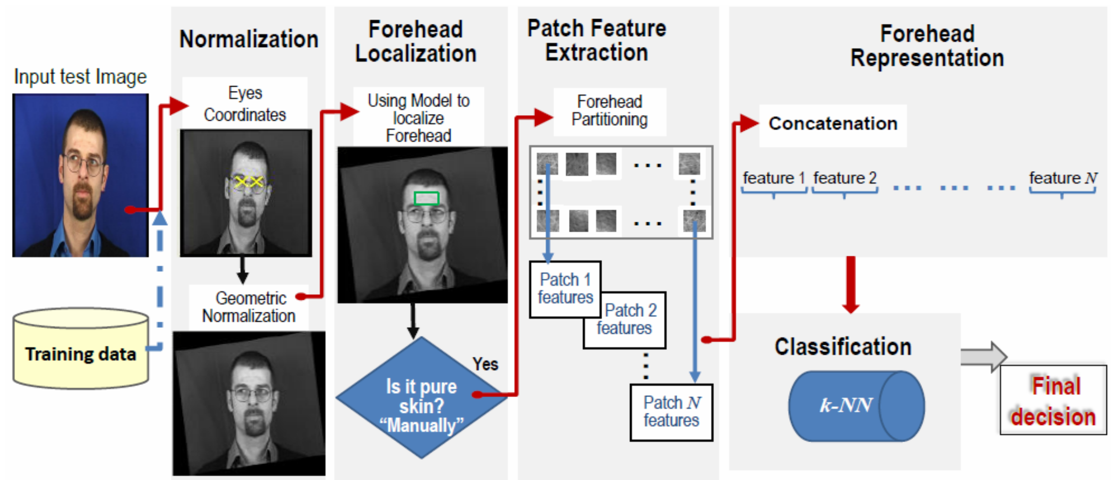

Figure 1. The System Block Diagram Extraction

2.1 Face Normalization

This study focuses on the forehead region of each face as it is likely to be less affected by facial expressions. Therefore the first step in the processing chain is to extract a normalised forehead region from face images. Geometric normalization is one of the key steps in most face recognition systems. It is used in order to rotate, crop and scale a face so that image comparison and feature extraction may be conducted on corresponding areas of facial images. In this paper only rotation of the face and then cropping of the forehead region are performed. It is assumed that all the subjects are roughly placed at the same distance from the camera and therefore no scale adjustments are necessary. Formally, let (x_r, y_r) and (x_l, y_l) be the locations of the centers of right eye and left eye respectively, then the face is rotated by angle θ that is computed as:

$$\theta = \tan^{-1}\left(\frac{y_r - y_l}{x_r - x_l}\right) \qquad (1)$$

In order to align the eyes, the new coordinates of every point in the face and therefore, as shown in Figure 2, the new locations of eye centers are given by:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad (2)$$

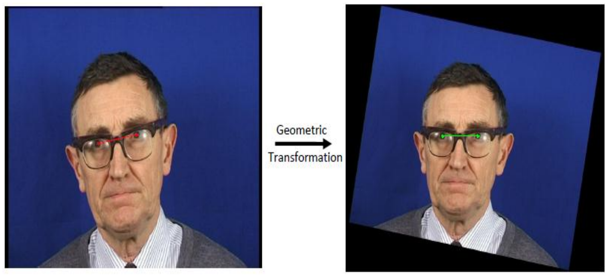

Figure 2. Forehead Region Extraction

2.2. Forehead Localization

Once the positions (x_r, y_r) and (x_l, y_l) of the right and left eye respectively have been determined, we simply calculate the distance between them as follows:

$$d = \sqrt{(x_r - x_l)^2 + (y_r - y_l)^2} \tag{3}$$

Based on the geometric model of the face in [7], the vertical distance between eyes and eyebrows may be set to $0.4d$. Therefore, the proposed forehead windows are rectangles of size $d \times 0.5d$ above the eyebrows region as shown in Figure 3.

Figure 3. Forehead Region Localization

2.3. Forehead Partitioning and Feature Extraction

In the proposed algorithm, the variability of forehead images is modelled by treating each forehead image as a collection of smaller sub-images, referred to here as patches. For the purpose of feature extraction it is necessary to establish the optimum patch size for texture analysis. This could be the whole of the extracted forehead. Alternatively a number of smaller patches can be chosen inside the forehead region for extracting features which can then be concatenated to form longer feature vectors.

In this work we explore subdividing the pure skin forehead region into non-overlapping patches of equal sizes with varying dimensions: 1×2, 2×2, 2×4, 2×8, 4×4 and 4×8 as shown in Figure 4. The 1×1 subdivision indicates the whole forehead being used for analysis without any partitioning. Each patch is then analyzed to extract separately the Gabor and LBP-based features that will be used for classification.

Figure 4. Partitioning Pure Skin Forehead Region

2.4 Gabor Filter for Extracting Skin Texture Features

Texture features based on Gabor functions belong to frequency-based approaches. They are widely used for texture segmentation [8] and fingerprint recognition [9] because Gabor functions have the attractive mathematical property of minimizing the joint uncertainty in frequency and space [10]. A two-dimensional Gabor function ψ which is a Gaussian kernel function modulated by a sinusoidal plane wave, can be defined as [11]:

$$\psi(x, y; f, \theta) = \frac{f^2}{\pi \gamma \eta} e^{-(\alpha^2 x' + \beta^2 y')} e^{j 2\pi f x'}, \quad (4)$$
$$x' = x\cos\theta + y\sin\theta, \quad y' = -x\sin\theta + y\cos\theta$$

f (cycles/pixel) is the central frequency of the sinusoidal plane wave, θ is the anticlockwise rotation of the Gaussian and the plane wave, α is the sharpness of the Gaussian along the major axis parallel to the wave, and β is the sharpness of the Gaussian minor axis perpendicular to the wave, $\gamma = f/\alpha$ and $\eta = f/\beta$ are defined such that the ratio between frequency and sharpness is constant. To obtain Gabor features of an image ζ, $\zeta(x, y)$ is convolved with a bank of scale and orientation selective Gabor filters:

$$r_{uv}(x, y; f_u, \theta_v) = \psi(x, y; f_u, \theta_v) * \zeta(x, y) \quad (5)$$

u, v indicate the filter scale and orientation respectively.

$$f_u = f_{max}\left(\sqrt{2}\right)^{-u}, \text{ for } u = 0, \cdots, m-1,$$
$$\theta_v = \frac{k\pi}{n}, \quad k = \{0, \cdots, n-1\} \quad (6)$$

where m is the total number of scales and n is the total number of orientations in the filter bank. Then all means, μ_{uv}, and standard deviations, σ_{uv}, of the magnitude $|r_{uv}|$ are computed. Finally, a feature vector $\overline{\mathbf{F}}$ is constructed using the concatenation of all the means and standard deviations. This is denoted as:

$$\overline{\mathbf{F}} = [\mu_{00}\, \sigma_{00}\, \mu_{01} \cdots \mu_{m-1n-1}\, \sigma_{m-1n-1}] \quad (7)$$

In Section 4.1 we report on an investigation to establish the optimal parameters used for the class of images that are to be classified with these filters in our work.

2.5 Local Binary Pattern (LBP) Operator for Extracting Skin Texture Features

The LBP operator is an efficient gray-scale invariant texture operator. It is robust to monotonic gray-scale changes caused, for example, by lighting variations and it has been shown to have high discriminative power for texture classification. The basic LBP operator labels the pixels of an image by thresholding the eight-neighbourhood of each pixel with the value of the centre and considers the result as a binary code [12]. Formally, it takes the form:

$$LBP(x_c, y_c) = \sum_{p=0}^{7} s(g_p - g_c)\, 2^p \tag{8}$$

where g_c , g_p correspond to the grey value of the central pixel $c\,(x_c, y_c)$ and the eight neighbouring pixels respectively. The function $s(x)$ is defined as follows:

$$s(x) = 1 \;\; if \;\; x \geq 0, \;\; s(x) = 0 \;\; if \;\; x < 0 \tag{9}$$

The LBP operator was extended to allow any radius and any number of pixels in the neighbourhood by using a circular neighbourhood and interpolating values at non-integer pixel coordinates [13]. Another extension to the original operator is the definition of the so called Rotation-Invariant LBP and Uniform LBP [13]. With uniform LBP ($LBP^{u2}_{P,R}$) it is possible to detect local primitives in the image texture such as edges, spots and corners. When an image is rotated in the plane, the neighbourhoods, g_p around the centre will be rotated in the same direction. The rotation invariant LBP operator ($LBP^{r2}_{P,R}$) can be implemented to remove this rotation effect. Ahonen, et al., [l4] proposed the basic methodology for LBP-based face description. It divides the facial image into local regions and LBP texture descriptors are extracted from each region independently. The descriptors are then concatenated to form a global description of the face.

Zou, et al., [15] compared PCA, Gabor wavelets and LBP for FERET and AR databases. The comparison of the three local feature representations is conducted on four 37×37 windows centred at four facial landmarks (two eyes, nose, and mouth). They concluded that LBP is a good local feature, but it is inadequate for non-monotonic illumination changes, which often appear in facial regions such as the nose. However, the Gabor features were found to be more robust to illumination variations.

In Section 4.1 we, nevertheless, investigate the use of LBP features for extracting identity information from pure-skin forehead images and report on experiments to establish the optimal parameters for their use.

2.6 Forehead Representation and Classification

Finally, the feature vectors extracted from every patch are concatenated to form a *forehead feature vector*. As a result, every forehead image can be represented either by Gabor-based feature vectors that have $2Lmn$ elements – where m is the total number of scales and n is the total number of orientations in the filter bank and L is the total number of patches - or by LBP-based feature vectors with 256 L elements.

A *k-Nearest-Neighbor* (*k*-NN) classifier with *k*=1 is used as a common platform to compare the different feature extraction methods.

3. Developing Hair/Skin Patch Classifier

The forehead region is sometimes partially or completely covered with hair, head covering, hats, scarves, .., etc. Therefore, it is essential to build a good classifier that will be able to automatically detect the useable skin regions. We propose to use a classifier such as that shown in Figure 5 for automatic classification of pure-skin patches in the forehead region. The cropped forehead region is divided into a number of non-overlapping patches. The results of Experiment 2 in Section 4.2 suggest a suitable partitioning of this region to be into 2×4 patches. In the training phase, the classifier is trained to recognize the categories of "Hair"/"Skin" patches. In the classification phase, unknown patches which were not part of the training dataset are assigned the class label of the nearest category according to the data available from the training patches. This procedure uses the Gabor features presented in Section 2.4 to learn the texture of Hair and Skin patterns.

Once trained, the classifier can be used to establish those forehead images that have all pure-skin patches. Only those images are then used for further processing using the approach developed in this paper. The experiments reported in this paper use a manual pure-skin forehead region detection approach to ensure there are only such patches used for calculating the error rates. However, the performance of this automated Hair/Skin classifier is also reported in Section 4.3 and it is seen to be quite effective.

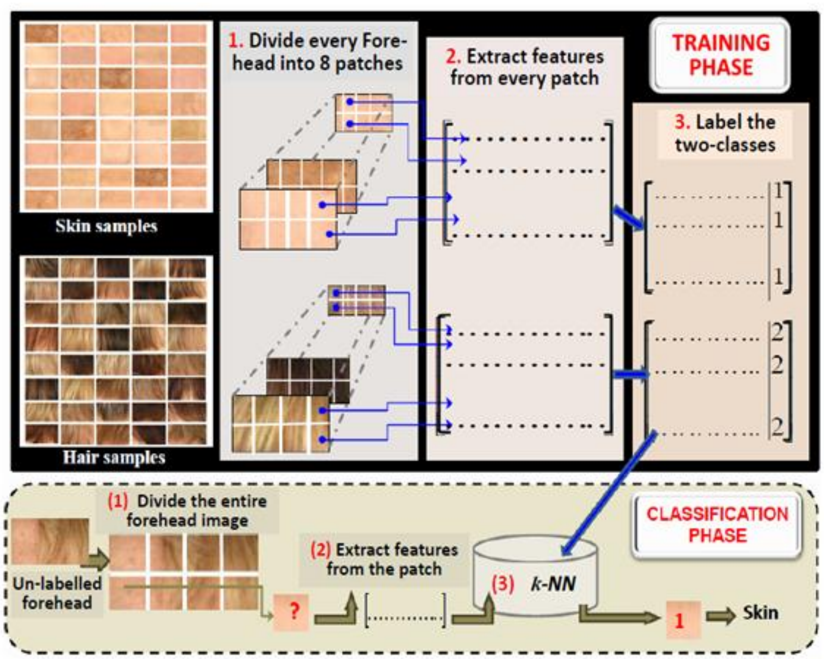

Figure 5. Scheme of the Two-class Skin/Hair Classifier

4. Experimental Results

The XM2VTS database [2] is a multi-modal database consisting of face images, video sequences and speech recordings taken of 295 people of both sexes and different ethnic origins with 4 face images for each person. Over a period of five months, recordings were acquired during four sessions under controlled conditions (uniform illumination, blue background) for the standard set. Two head rotation and "speaking" shots were taken during each session. Since the data acquisition was distributed over a long period of time, significant variability of appearance of people e.g. changes of hair style, facial hair, shape and presence or absence of glasses, is presented in the recordings. The size of images in this database is 720×576 pixels and inter-eye distance ranges from 86 pixels to 120 pixels.

4.1 Experiment 1: Identification Scenario

This experiment has been conducted on a subset of the XM2VTS database to evaluate the performance of the Gabor-based features for skin texture classification as presented in Section 2.4. The data used consists of pure skin patches as determined manually. The number of images in this group is 336 pure skin patches which came from 84 different persons. Only three partitionings for forehead region were evaluated. These were: 1×2, 2×4 and 4×8 with patch size of $0.5d\times0.5d$, $0.25d\times0.25d$ and of $0.125d\times0.125d$ respectively. Table 1 lists the recognition accuracy achieved for different parameter settings.

The results indicate that the partitionings with larger analysis windows significantly improve performance by increasing the number of orientations/scales. For example, when the number of scales/orientations has changed from 4/4 to 12/12, the classification accuracy improves only from 60.12% to 64.29% for the 4×8 partitioning but from 58.33% to 67.86% for the 2×4 partitioning and from 35.12% to 56.55% for the 1×2 partitioning. Additionally, it is noted that increasing the number of scales and orientations of the Gabor filter bank will add to the complexity of the system and its computational time and will not always result in an increase in the classification accuracy.

Table 1. Classification Accuracy using Gabor Features

Orientation	Patch	Scales				
		4	6	8	12	16
4	1×2	35.12%	37.20%	38.99%	44.05%	45.54%
	2×4	58.33%	62.80%	65.18%	65.77%	66.96%
	4×8	60.12%	62.20%	61.91%	62.50%	62.20%
6	1×2	36.61%	42.26%	43.75%	49.11%	50.59%
	2×4	61.01%	63.09%	65.48%	66.07%	67.56%
	4×8	60.12%	62.80%	63.39%	63.39%	63.39%
8	1×2	42.26%	45.83%	48.81%	53.57%	54.76%
	2×4	61.61%	65.48%	66.96%	67.86%	68.16%
	4×8	60.42%	63.09%	63.99%	63.99%	63.69%
10	1×2	41.67%	48.51%	51.49%	55.06%	56.84%
	2×4	62.50%	65.48%	66.67%	67.86%	67.86%
	4×8	61.31%	62.20%	63.10%	64.29%	63.99%
12	1×2	44.34%	49.40%	53.57%	56.55%	57.14%
	2×4	62.50%	65.18%	67.56%	67.86%	66.96%
	4×8	61.91%	63.09%	63.69%	64.29%	63.99%
16	1×2	46.73%	49.40%	55.06%	57.44%	58.04%
	2×4	62.50%	65.18%	67.56%	67.26%	67.86%
	4×8	61.91%	63.39%	63.99%	64.29%	63.99%

4.2. Experiment 2: Verification Scenario

The aim for conducting this experiment is threefold: (*i*) to determine the appropriate number of scales and orientations that can be used for Gabor features and (*ii*) to investigate the different LBP operators (standard, Uniform '*u2*', Rotation invariant '*ri*') using different radius and sample points (*iii*) to decide the patch size for achieving the best performance. This experiment has been conducted on a part of the XM2VTS database: 460 images that came from 115 different persons have been manually determined as faces with pure skin forehead regions. Half of the entire dataset is used as the training set and the remaining represents the test set. Table 2 lists the Equal Error Rates (EER) obtained for both Gabor and LBP approaches for varied patches size.

This table suggests that the achievable performance with Gabor features is better than with LBP features, showing that the best EER of 0.065 for Gabor features has been obtained using 6 scales, 6 orientations with a 4×4 partitioning. If no patch partitioning is used, the resulting EER is significantly higher at 0.213.

It also shows that the performance of standard LBP approach using 8 sample points with radius equal to 2 outperforms other LBP operators for this dataset. If no patch partitioning is used an EER of 0.165 is achieved with this method. It is clear that the performance of LBP approach does not significantly improve using rotation invariant features because all images were normalized in the pre-processing stage.

Table 2. EER using Gabor Features with Different Scales and Orientations and Different LBP Operators

Approach	S	O	1×1	1×2	2×2	2×4	2×8	4×4	4×8
GABOR	4	6	2.13e-01	1.83e-01	1.48e-01	1.30e-01	1.13e-01	1.00e-01	1.01e-01
		8	2.14e-01	1.70e-01	1.38e-01	1.13e-01	1.09e-01	9.13e-02	9.90e-02
		12	2.01e-01	1.52e-01	1.13e-01	1.01e-01	1.08e-01	7.87e-02	9.97e-02
	6	6	**2.13e-01**	1.75e-01	1.27e-01	1.09e-01	8.37e-02	**6.52e-02**	6.52e-02
		8	2.04e-01	1.57e-01	1.14e-01	9.57e-02	8.26e-02	6.52e-02	6.96e-02
		12	1.83e-01	1.26e-01	1.00e-01	8.27e-02	8.26e-02	6.52e-02	7.02e-02
	8	6	1.96e-01	1.53e-01	1.01e-01	8.26e-02	7.39e-02	7.44e-02	8.79e-02
		8	1.79e-01	1.35e-01	9.57e-02	7.49e-02	6.86e-02	7.01e-02	9.57e-02
		12	1.73e-01	1.26e-01	9.06e-02	6.59e-02	7.36e-02	6.62e-02	9.57e-02
LBP	1,8		1.61e-01	1.44e-01	1.09e-01	9.57e-02	1.00e-01	9.57e-02	1.23e-01
	2,8		**1.65e-01**	1.38e-01	1.13e-01	**8.26e-02**	1.10e-01	1.17e-01	1.87e-01
	$1,8^{u2}$		1.69e-01	1.55e-01	1.35e-01	1.05e-01	1.13e-01	1.09e-01	1.17e-01
	$1,8^{ri}$		2.09e-01	2.13e-01	2.00e-01	1.91e-01	2.31e-01	2.13e-01	2.40e-01

Features using patch partitionings of 2×4 and 4×4, achieved the best results for LBP and Gabor approaches respectively. Their DET curves are presented in Figure 6. Since the error that is produced from Gabor and LBP-based features at 2×4 is less than the error that is produced from them at 4×4, this experiment suggest that the 2×4 is an appropriate choice for the partitioning scheme for this dataset.

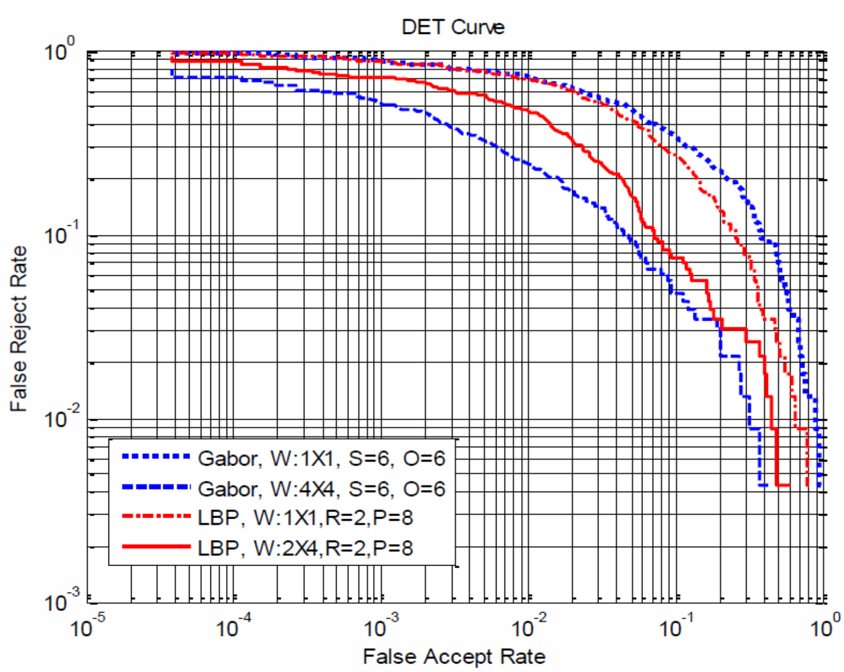

Figure 6. DET Curves for Gabor and LBP Approaches; *W* indicates window size, Gabor parameters are: *S*: Scale and *O*: Orientations, LBP parameters are: *R*: Radius and *P*: Sample Points

4.3. Test of the Proposed Skin/Hair Classifier

The proposed Skin/Hair classifier, explained in Section 3, was also tested using the XM2VTS data. The training data set is composed of the patches that are produced from partitioning the forehead window which is extracted from the XM2VTS facial images, and then these patches have been classified into two groups, (i) first group includes 800 patches that are partially or completely covered by hair and (ii) second group includes 800 pure skin patches. We have used 10-fold cross-validation analysis for testing. Table 3 shows that proposed classifier is able to distinguish between skin and hair texture. The best result of 96.97% correct classification rate was obtained using a Quadratic classifier. Samples from the output of this classifier are shown in Figure 7.

Table 3. Classification Accuracy for Different Classifiers

Classifier	Accuracy
Quadratic classifier	**96.97%**
k-nearest neighbour classifier	94.24%
Logistic linear classifier	96.41%
Parzen classifier	94.12%
Naive Bayes classifier	94.99%

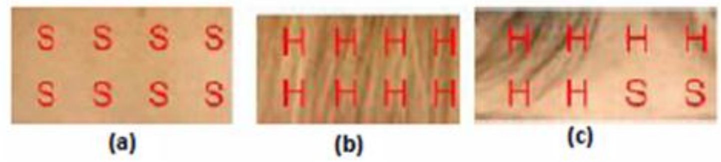

Figure 7. Outputs of the Hair/Skin Classifier

5. Conclusion and Future Work

We have presented a novel technique for extracting biometric information from the forehead region which may help to improve face recognition accuracy especially in applications where a full view of the face may not be available. Skin texture patches are detected using features based on Gabor filters. Gabor-based and LBP-based features are then used to classify forehead regions containing pure skin. The results indicate that the forehead region alone provides useful biometric information for person recognition. Future work will explore the development of adaptive algorithms to cope with forehead regions that may have partial hair coverage and will include the study of information available from skin texture in other facial regions.

References

[1] D. Lin and X. Tang, "Recognize High Resolution Faces: From Macrocosm to Microcosm", Proc. IEEE Computer Vision and Pattern Recognition, **(2006)**, pp. 1355-1362.
[2] K. Messer, J. Matas, J. Kittler, J. Luettin and G. Maitre. "XM2VTSDB: The Extended M2VTS Database", In Second International Conference on Audio and Video-based Biometric Person Authentication, **(1999)**.
[3] J. -S. Pierrard and T. Vetter, "Skin detail analysis for face recognition", Proc. IEEE Computer Vision and Pattern Recognition **(2007)**, pp. 1–8.
[4] J. Lee, A. K. Jain and R. Jin. "Scars, marks and tattoos (SMT): Soft biometric for suspect and victim identification", Proc. Biometric Symposium, Biometric Consortium Conference, **(2008)**, pp. 1–8.
[5] A. K. Jain, S. C. Dass and K. Nandakumar, "Soft biometric traits for personal recognition systems", In Proc. ICBA, **(2004)**, pp. 731–738.
[6] A. K. Jain and U. Park. "Facial Marks: Soft Biometric for Face Recognition", In ICIP, **(2009)**.
[7] F. Shih and C. Chuang, "Automatic extraction of head and face boundaries and facial features", Information Sciences, vol. 158, **(2000)**, pp. 117–130.
[8] T. P. Weldon, W. E. Higgins and D. F. Dunn, "Efficient Gabor filter design for texture segmentation," Pattern Recognition, vol. 29, no. 12, **(1996)**, pp. 2005–2015.
[9] C. J. Lee and S. D. Wang, "Fingerprint feature extraction using Gabor filters", Electronics Letters, vol. 35, no. 4, **(1999)**, pp. 288–290.
[10] J. Daugman, "Complete discrete 2D Gabor transform by neural networks for image analysis and compression", IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 36, **(1988)**, pp. 1169-1179.
[11] V. Kyrki, J. -K. Kamarainen and H. Kälviäinen, "Simple Gabor feature space for invariant object recognition", Pattern Recognition letters**,** vol. 25, **(2003)**, pp. 311–318.
[12] T. Ojala, M. Pietikinen and D.Harwood, "A comparative study of texture measures with classification based on feature distribution", Pattern Recognition, vol. 19, **(1996)**, pp. 51-59.
[13] T. Ojala, M. Pietikäinen and T. Mäenpää, "Multiresolution grayscale and rotation invariant texture classification with local binary patterns", IEEE T-PAMI, vol. 24, no. 7, **(2002)**, pp. 971-987.
[14] T. Ahonen, A. Hadid and M. Pietikainen, "Face description with local binary patterns: Applications to face recognition", IEEE T-PAMI, vol. 28, **(2006)**, pp.2037-2041.
[15] J. Zou, Q. Ji and G.Nagy, "A comparative study of local matching approach for face recognition", IEEE Trans. Image Process, vol. 16, **(2007)**, pp. 2617-2628.

Authors

Garsah Farhan Al-Qarni received her B.Sc. (with highest honor) in Mathematics from Faculty of Science, Girls College in Dammam, Saudi Arabia, where she became a Graduate Assistance. In 1999 she was promoted to the rank of Lecturer at the same college after she obtained an M.Sc in Mathematics and Computer Science. She was awarded the University of Dammam scholarship to pursue her Ph.D degree in Computer Science. She is currently a PhD student at the School of Engineering and Digital Arts, University of Kent, Canterbury, UK. Her research interests include computer vision, biometrics, genetic algorithms and optimization.

Dr Farzin Deravi obtained his first degree in Engineering and Economics from the University of Oxford in 1981 and his M.Sc. in Electronic Engineering from Imperial College, University of London in 1982. From 1983 to 1987 he worked as a research assistant at the University of Wales, Swansea where he obtained my Ph.D. In 1987 he joined the academic staff at Swansea at the Department of Electrical and Electronic Engineering. In 1998 he joined the Department of Electronics – now School of Engineering and Digital Arts - at the University of Kent where he is Reader in Information Engineering. He is a fellow of the Institution of Engineering and Technology (IET), and Editor-in-Chief of the IET Image Processing journal. He is currently the Vice President of the BioSecure Association, Chair of the European Association for Biometrics' Training and Education Committee and also serves on BSI and ISO committees on biometric standardisation.