Gabor and Maximum Response Filters with Random Forest Classifier for Face Recognition in the Wild

Yuen-Chark See[1], Eugene Liew[1], and Norliza Mohd Noor[2]
[1]Department of Electrical and Electronic Engineering, University Tunku Abdul Rahman, Malaysia
[2]Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Malaysia

Abstract: *Research on face recognition has been evolving for decades. There are numerous approaches developed with highly desirable outcomes in constrained environments. In contrast, approaches to face recognition in an unconstrained environment where varied facial posing, occlusion, aging, and image quality still pose vast challenges. Thus, face recognition in the unconstrained environment still an unresolved problem. Many current techniques are not performed well when experimented in unconstrained databases. Additionally, most of the real-world application needs a good face recognition performance in the unconstrained environment. This paper presents a comprehensive process aimed to enhance the performance of face recognition in an unconstrained environment. This paper presents a face recognition system in an unconstrained environment. The fusion between Gabor filters and Maximum Response (MR) filters with Random Forest classifier is implemented in the proposed system. Gabor filters are a hybrid of Gabor magnitude filters and Oriented Gabor Phase Congruency (OGPC) filters. Gabor magnitude filters produce the magnitude response while the OGPC filters produce the phase response of Gabor filters. The MR filters contain the edge- and bar-anisotropic filter responses and isotropic filter responses. In the face features selection process, Monte Carlo Uninformative Variable Elimination Partial Least Squares Regression (MC-UVE-PLSR) is used to select the optimal face features in order to minimize the computational costs without compromising the accuracy of face recognition. Random Forests is used in the classification of the generated feature vectors. The algorithm performance is evaluated using two unconstrained facial image databases: Labelled Faces in the Wild (LFW) and Unconstrained Facial Images (UFI). The proposed technique used produces encouraging results in these evaluated databases in which it recorded face recognition rates that are comparable with other state-of-the-art algorithms.*

Keywords: *Face recognition, labelled faces in the wild, unconstrained facial images.*

1. Introduction

Face recognition has been actively studied and researched for over half a century due to its benefits in different applications, predominantly in controlled environments. Recent studies have shifted to faces in the wild (uncontrolled environment or unconstrained database), such as live face images captured from the internet or surveillance cameras. These images comprised of different resolution, occlusion, illumination, and facial expressions poses challenges in face recognition [27].

The human face is subjected to many factors that cause variation in its appearance. This could be classified into two categories: intrinsic factors and extrinsic factors. Intrinsic factors include the varying facial appearance of the same individual due to facial expression. The extrinsic factors include the interaction of light with the face, such as brightness level, head scale and orientation, resolution and noise. These factors affect face detection and recognition rate. Feature extraction plays an important role in solving these challenges. It is usually divided into two mainstreams, namely holistic approach and local feature approach. The holistic approach uses statistical methods to derive an image (high-dimensional) into a feature space (low-dimensional). Examples of this approach are Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). PCA is one of the earlier approaches in face recognition using eigenfaces which achieve good verification accuracy for well-alighted frontal faces [44]. Turhan and Bilge [43] improvised the traditional PCA with a class-wise 2-D PCA-based technique that produced more than one subspace. Li *et al.* [22] combined the PCA technique with their early proposed Probabilistic Elastic Part (PEP) model to form a better technique called Eigen PEP for large database video face recognition. Li and Hua [21] further enhanced the PEP model and proposed the Hierarchical-PEP model for unconstrained face recognition. In this technique, the face is broken down into different sections with different details that were later used for recognition. The feature approach extracts local face features such as the eyes, nose, and mouth of an individual. This

approach allows flexible deformation at the feature points, which is good for face images with pose variation. Local Binary Pattern (LBP) is one of the methods to discriminate the textures and edges within an image [5, 28]. The LBP kernel operates on the change in intensity in the neighbourhood of a pixel. Ahonen *et al.* [1] had used the histogram of LBP values as the facial feature representation. However, the LBP-based feature extraction carries out histogram computation on a uniform and predetermined grid in the facial image and does not consider the properties of an image. Therefore, Lenc and Král [18] proposed an automatic face recognition system based on the LBP with Gabor wavelets and k-means clustering approach to solve the problem of feature position in the conventional LBP approach. The same authors [20] further enhanced the system to recognise-fixed coordinates and facial fiducial points even though there are large differences in the positions of facial features between images with large pose variations.

There have been many improvements made to the LBP. Vu *et al.* [46] proposed the Patterns of Oriented Edge Magnitudes (POEM) approach where the LBP based structure is applied to the oriented magnitudes. Vu [45] later further improved the POEM approach and proposed a novel feature set called Patterns of Orientation Difference (POD), which can acquire information on the self-similarity of an image. Lin and Chiu [23] used LBP edge-mapped descriptor that utilize the maxima of gradient magnitude points [12] for face recognition. This approach can exhibit facial contours with a low computational cost. In order to overcome the problem of noise and illumination changes in the face image, Kral and Vrba [17] enhanced the LBP by measuring the features from point-sets instead of the isolated points. Juefei-Xu *et al.* [15] proposed a technique capable of generating a highly discriminative matching score without precise face image alignment for unconstrained face images. The authors used only one image per subject for training, and with a wide range of Three Dimensional (3-D) rotations, a set of new face images are generated. The periocular regions of these images are segmented out. Walsh Local Binary Patterns (WLBP) is used to extract the periocular features from these images. They called the proposed technique as Spartans. Zhang *et al.* [51] found that by encoding the (n-1)[th]-order local derivative direction variations, more data are collected than the first-order local pattern as in LBP. The authors name this approach as Local Derivative Pattern (LDP). Based on the concept of LBP, Ylioinas *et al.* [50] developed a high dimensional features representation for face recognition by getting histograms of Binarized Statistical Image Features (BSIF) codes. This approach was claimed to provide an optimal discriminative vector representation for the face. Barkan *et al.* [4] modified the original LBP and introduced the Diffusion Maps technique for dimensionality

reduction. By fusing this technique with PCA, the authors had shown further improvement in the classification accuracy.

Another local feature descriptor, a histogram of gradients descriptors such as Scale Invariant Feature Transform (SIFT) [25] and Histogram of Oriented Gradients (HOG) [10], have shown encouraging results in face verification. Simonyan *et al.* [36] proposed to use Fisher vectors on densely sampled SIFT features that achieved a good face verification performance on large scale databases like Labeled Faces in the Wild. Seo and Milanfar [33] invented a new descriptor called Locally Adaptive Regression Kernel (LARK) that can determine the self-similarity between the centre pixel and surrounding pixels.

Štruc and Pavešić [37] integrated Gabor magnitude and phase feature information with LDA to devise a new method known as complete Gabor-Fisher classifier. Both authors claimed that the method outperformed Principal Component Analysis. Yi *et al.* [49] utilized the Gabor filter for feature extraction and proposed Pose Adaptive Filter (PAF), which converted the Gabor filter based on the pose variation of the face images. This was achieved by using the 3-D deformable model created according to the face image. 3-D deformable model or manual annotation on the respective face image was needed for facial landmark localization. Sagonas *et al.* [32] utilized a statistical model based on hundreds of frontal images to perform landmark localization. With this technique, the frontal view of a face image in unconstrained situations can be reconstructed. They named the technique Robust Statistical Frontalization (RSF). Based on the concept of the Gabor filter, Pinto *et al.* [29] introduced modern multiple kernel learning (MKL) techniques using V1-like features. This technique is unsusceptible to lighting and image variations. Arashloo and Kittler [2, 3] proposed a nonlinear binary class-specific kernel discriminant analysis classifier fused with the Markov Random Fields (MRF) approach. In this approach, the used image is inserted into a discriminative subspace.

In recent years researchers explored the neural network approach in facial recognition. Numerous mathematical models were introduced under this approach. The deep convolutional neural network approach has produced encouraging results in face recognition in the unconstrained environment [26, 40]. Sun *et al.* [39] proposed a hybrid model of the neural networks approach. They created a hybrid Convolution Neural Network (CNN) as feature extractors and Restricted Boltzmann Machine (RBM) as classifiers. The proposed model had achieved a good result under the LFW database [14] in unrestricted with label-free outside data and labelled outside data. Xi *et al.* [48] introduced an approach called Local Binary Pattern Network (LBPNet) based on unsupervised deep learning, which adapts the concept of CNN. This model has two sections, the deep network, which

utilized the LBP and PCA for feature extraction and the regular network for classification. Devi and Hemachandran [11] proposed a system that utilized PCA, Wavelet Transformation and Gabor wavelet for feature extraction. The modular neural network is used for image retrieval, and the Support Vector Machine (SVM) is used as a classifier. With this combination, the training time for the large database is reduced.

Based on the literature as discussed early, face recognition in the wild continues to be a challenging task due to pose, illumination variations, occlusion, etc. Researchers are still finding the balance between computational complexity and recognition accuracy [2, 3, 4, 15, 23, 40, 45, 46, 50]. In the recent literature, the deep learning technique [11, 39, 48] gives promising results in face recognition. But this technique needs millions of parameters that lead to high requirements of processing power, memory size etc.

In this paper, the following contributions are made,

- Two filters are used in the features extractions process. Gabor filter is to extract the features in spatial and frequency domains. A Maximum Response Filter (MR) is used to respond to oriented image patches and anisotropic textures. MR filter uses both Gaussian and Laplacian of Gaussian (LoG) filters. Compared with the Gabor filter, the MR filter has an additional LOG element that detects edges by looking for zero crossings in the image. The combination of Laplacian and Gaussian functions assists in smoothening the image and edge detection.
- The feature selection using Monte Carlo Uninformative Variable Elimination Partial Least Squares Regression (MC-UVE-PLSR) is included to trim the large dimensionality of feature sets to save computation costs while maintaining the accuracy rate. The implemented method is applied to two unconstrained databases, LFW [14] and UFI [19] to test the reliability of the face recognition system.
- The complete system (training and testing process) is discussed and evaluated.

The rest of the paper is organized as follows. Section 2 presents a detailed description of the proposed system. In section 3, the experimental settings, the experimental datasets and simulation results are discussed. Finally, section 4 concludes this paper.

2. The Proposed System

The proposed system is divided into two main phases, training and testing phases. The flow diagram of these two phases is shown in Figures 1-a) and 1-b), respectively. The training stage trains the program to implement the classification process of the sampled face image based on the labelled data. The testing stage involves the application of the trained module to carry out the classification of the test image. In the training process, the feature vectors generated are fed into the classifiers using Random Forest algorithms to produce a trained and learned model. The feature vectors are utilized to grow a Random Forest. In the testing process, the prediction is based on the obtained feature vectors of the testing images using the trained decision trees. The prediction function generates a matrix of matching scores that state the probability of that testing image belonging to a specific class. The class candidate with the highest matching scores suggests that the testing image most probably belongs to that class. The performance of random forest constructed is evaluated by analysing the Receiver Operating Characteristic (ROC) curve. ROC curve is the evaluation graph that investigates the interaction between true-positive-rate and false-positive-rate. The recognition rate will also be carried out to measure the classification accuracy rate.
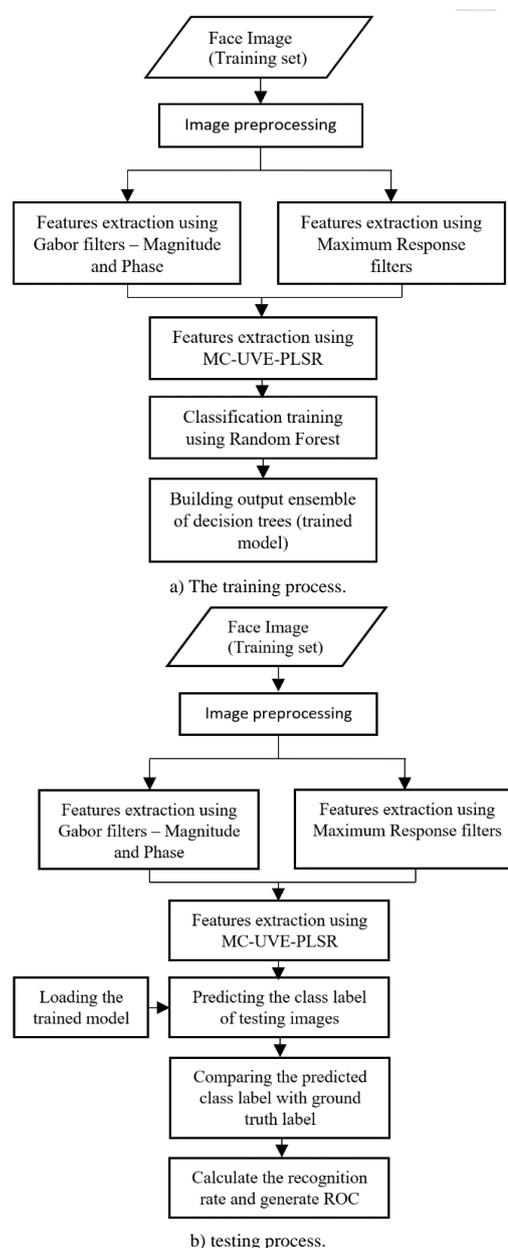

a) The training process.

b) testing process.

Figure 1. Flow diagram.

2.1. Image Pre-Processing

The input face image undergoes pre-processing in the first place by converting the Red, Green, and Blue (RGB) scale into grey scale. The image pre-processing is to refine the quality of an image. Filtering is applied to images to reduce image details to speed up the computational process. The Tan and Triggs [41] illumination normalization is then applied to the grey-scaled image. This process minimizes the adverse effects of illumination variations, shading or flashing of the images while still retaining the important features in the recognition process.

2.2. Feature Extraction

The image undergoes a feature extraction process using Gabor filters and Maximum Response filters. There are two components of Gabor filters which are Gabor magnitude components and Gabor phase components. The feature vectors generated by the filters are large in dimensionality and consume high computation costs and time. Thus, feature selection is being carried to use top importance ranking features to reduce the number of features for the next stage. Facial features extracted using the Gabor filter have robustness against illumination and minor facial expression variation. The Gabor wavelet, $\emptyset_{m,n}(x,y)$ [35, 38]

$$\emptyset_{m,n}(x,y) = \frac{f_m^2}{\pi As} e^{-\left(\frac{f_m^2}{A^2}\right)x'^2 + \left(\frac{f_m^2}{n^2}\right)y'^2} e^{j2\pi f_m x'} \tag{1}$$

Where,
$x' = x\cos\theta_n + y\sin\theta_n$
$y' = -x\sin\theta_n + y\cos\theta_n$
f_m = Gaussian center frequency = $f_{max}/2^{m/2}$
m = Number of scales
θ_n = Gaussian orientation = $n\pi/8$
A = Gamma, ratio between the center frequency
s = size of Gaussian envelope
f_{max} = Maximum frequency of the filter

Štruc and Pavešić [37] recommends that $A=s=\sqrt{2}$ and $f_{max} = 0.2$. Filter bank of five scales and eight orientations is constructed with $m \in \{0,1,2,3,4\}$ and $n \in \{0,1,2,3,4,5,6,7\}$. The filter bank has the real and imaginary terms of the Gabor wavelet. The real term is used in the facial feature extraction process.

The input image is a greyscale image of a face having the size of $p \times q$ pixels. Gabor filter is denoted as $\emptyset_{m,n}(x,y)$ with centre frequency f_m and orientation θ_n. The transfer function of the filter is the convolution between the greyscale image, $B(x,y)$ and Gabor filter, $\emptyset_{m,n}(x,y)$ [38, 52].

$$H_{m,n}(x,y) = B(x,y) * \emptyset_{m,n}(x,y) \tag{2}$$

$H_{m,n}(x,y)$ refers to the complex value output of the filter function. $H_{m,n}(x,y)$ is broken down into real term $R_{m,n}(x,y)$ and imaginary term, $I_{m,n}(x,y)$.

$$R_{m,n}(x,y) = Re[H_{m,n}(x,y)] \tag{3}$$

$$I_{m,n}(x,y) = Im[H_{m,n}(x,y)] \tag{4}$$

The magnitude of the output filter, $J_{m,n}(x,y)$ is obtained using Pythagorean Theorem.

$$J_{m,n}(x,y) = \sqrt{R_{m,n}^2(x,y) + I_{m,n}^2(x,y)} \tag{5}$$

Gabor magnitude changes slowly across the spatial location, but Gabor phase information can have very different values even if measured a few pixels apart. Zhang *et al.* [52] demonstrated that the face representation based on the phase section of the Gabor feature contributes to face recognition accuracy. While Kovesi [16] employed the phase congruency model to look for points within an image where the log-Gabor filter outputs are maximally in phase with each other. The phase congruency is computed for every filter orientation and is combined to obtain a final phase congruency output image. His work claimed that phase congruency output robustly detects the edges and corners. However, doubts are posed by Liu and Wang [24] that Kovesi's [16] method might not be capable of extracting features of multi-orientations effectively.

For the implementation of feature extraction in this project, Gabor phase congruency information for multi-orientations is computed, and then augmentation of the phase congruency feature vectors is carried out. Instead of the log-Gabor function used by Kovesi, the conventional Gabor function is used for filters with five scales and eight orientations. The definition of Oriented Gabor Phase Congruency (OGPC) is defined as follows [38]

$$OGPC = \frac{\sum_{m=0}^{p-1} A_{m,n}(x,y)\Delta\phi_{m,n}(x,y)}{\sum_{m=0}^{p-1}(A_{m,n}(x,y)+\epsilon)} \tag{6}$$

Where ϵ is a constant to prevent equation divided by zero; $\phi_{m,n}(x,y)$ is the phase angle of Gabor function; m is the scales and n is the orientations; p is the total number of scales;

$\Delta\phi_{m,n}(x,y)$ is the phase deviation given by

$$\Delta\phi_{u,v}(x,y) = \cos\left(\phi_{u,v}(x,y) - \bar{\phi}_{u,v}(x,y)\right) - \left|\sin\left(\phi_{u,v}(x,y) - \bar{\phi}_{u,v}(x,y)\right)\right| \tag{7}$$

Where $\bar{\phi}_{u,v}(x,y)$ is the mean phase angle at u-the orientation and $\phi_{u,v}(x,y)$ is the phase angle of the Gabor filter given as

$$\phi_{m,n}(x,y) = \tan^{-1}\left(\frac{I_{m,n}(x,y)}{R_{m,n}(x,y)}\right) \tag{8}$$

The output of OGPC is the illumination and contrast independent facial feature representations. The OGPCs are then downsized by the down-sampling factor. Z-score normalization is applied to the downsized

OGPCs, and the OGPC feature column vectors are concatenated together becoming augmented OGPC feature vectors.

The MR filter has a total of 38 filters. The filters include one Gaussian filter and one LoG isotropic filter at scale $\sigma = 10$. Gaussian filter helps to smooth the image to reduce noise before Laplacian filter is used for edge detection [6]. There is the edge (first derivative) anisotropic filters with six orientations and three scales. Similarly, there is a bar (second derivative) anisotropic filters with six orientations and three scales. Figure 2 shows the filter response at different orientations and scales of the anisotropic and isotropic filters. The isotropic filter responses are used without any further processing, but for the anisotropic filter, only the maximum filter responses at every scale across entire orientations are chosen. This generates eight filter responses, and they are rotationally invariant. The respective equations are shown as follows [13, 42]:

$$Edge\ filter = -\exp\left[-\frac{input^2}{2\sigma^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\right]\left(\frac{input}{\sigma^2}\right) \quad (9)$$

$$Bar\ filter = \exp\left[-\frac{input^2}{2\sigma^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\right] \cdot \left(\frac{input^2 - \sigma^2}{\sigma^4}\right) \quad (10)$$

$$input = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} * (pixel\ value\ of\ image) \quad (11)$$

where $\theta = \frac{n\pi}{6}$, n = 0,1,2, …, 5 and σ is the standard deviation along x-axis and y-axis.
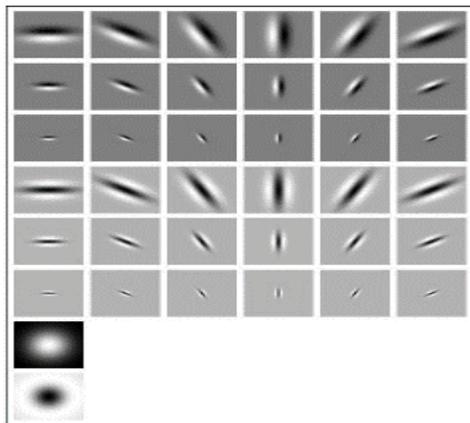

Figure 2. The filter output of the MR filters.

*The first three rows are edge filters. The fourth until sixth rows are bar filters. The last two rows are isotropic Gaussian and Laplacian of Gaussian filters.

The output response is the convolution between the facial images and the MR filter banks. The reason for using MR filter is because it has a bar filter, isotropic and Laplacian filters that can extract features that could not be executed by the Gabor filter.

2.3. Feature Selection

MC-UVE-PLSR is used in the feature selection process. The combination of the

Monte Carlo (MC) method and Uninformative Variable Elimination (UVE) method is used to select features generated by Gabor filters. Usually, the UVE employs the leave-one-out procedure. However, the Monte Carlo method is used instead of the leave-one-out procedure [7, 30]. The samples are divided randomly into a training set, evaluation set and prediction set. The Monte Carlo randomly chooses a few subsamples from the training set (at 75%) to build the Partial Least Square (PLS) model, and this process repeats 1000 times. The PLS regression coefficients and the stability of each feature set are determined as:

$$\gamma = \alpha X + \kappa \quad (12)$$

Where γ is the prediction; X is the information of the feature sets; α is the regression coefficients; κ is the offset.

The reliability of a feature is determined by its stability level, given as in Equation (13). The $\mathscr{s}_i$ is the regression coefficients that contribute the selected features to the prediction model.

$$\varsigma_i = \frac{mean\ of\ \alpha_i}{standard\ deviation\ of\ \alpha_i} \quad (13)$$

Where $i = 1,2, …,$ number of features.

The level of stability determines the importance of the features. The features are rank based on the stability level from highest to lowest. The total features after going through the down sampling process are 4000 features (full features). From the features generated by Gabor and Maximum response filters, the top-ranked 2000 features (best features) are selected to be fed into the Random Forest classifier function. The reason for selecting 2000 features is that it is the balance point between maintaining optimal accuracy level and the computation costs in implementing the training process.

2.4. Gabor and Maximum Response Filters

According to Štruc and Pavešić [38], the fusion between magnitude and phase information of the Gabor filter is build a robust facial feature extraction. The effectiveness of the feature extraction method is measured by the face recognition rate.

Fusing the Gabor filters (Gabor Magnitude filters + OGPC filters) and MR filters raises the recognition rate of the proposed face recognition system compared with the Gabor-only feature extraction. The fusion matching scores are computed as:

*Gabor-only matching scores= (1-q)*Gabor magnitude scores + q * OGPC scores*
(14)

Fusion matching scores=(1-r) Gabor-only matching scores + r * MR matching scores*
(15)

Where q sets the ratio between Gabor magnitude and phase components and r sets the ratio between Gabor filter and Maximum Response components, $q \in [0, 1]$, $r \in [0, 1]$.

Figure 3 shows the fusion process of the testing phase that utilizes the Gabor magnitude, OGPC and MR filters as features extractors.
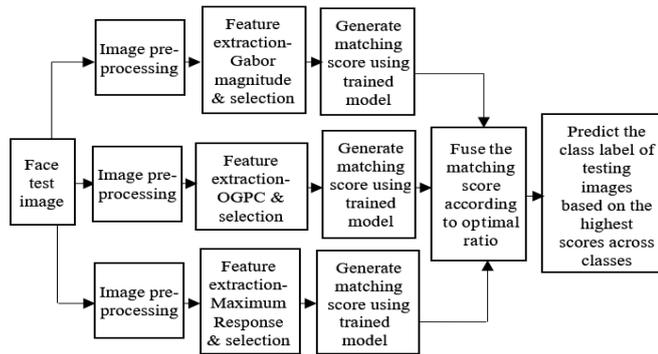

Figure 3. Block diagram of fusing between Gabor magnitude, OGPC and MR filters in the testing phase.

3. Experimental Results and Discussion

3.1. Datasets

The proposed hybrid technique is evaluated by piloting face recognition experiments on two popular face databases namely, LFW [14], and UFI [19].

The LFW dataset consists of 13,233 face images collected in the wild on 5,749 people. Sample images of the LFW database are shown in Figure 4-a). The words "in the Wild" means the face images are obtained without going through any parameter adjustments and taken in "natural" conditions as in daily life routine images. The UFI dataset is a real-world dataset containing cropped images at the size of 128 x 128 pixels from 605 people. Reporters of the czech news agency collected these images. Sample images of the UFI database are shown in Figure 4-b).

a) Labeled Faces in the Wild (LFW).

b) Unconstrained Facial Image (UFI).

Figure 4. Faces in the wild database.

3.2. The Effect of Gabor Magnitude Filter Parameters on Recognition Rate

Referring to Equation (1), Gamma(*A*) determines the ellipticity of the Gaussian function and the width of the Gaussian window. *s* specifies the linear size of the visual receptive field simulated by the Gabor filter. It is found that by setting to lower Gamma values and *s*, the finer discrimination of the texture of the facial region is obtained. It is also found that the higher value of the maximum central frequency improves the recognition rate. The lower Gamma value gives smaller Gaussian bandwidth and a sharper filter so that the tails of the two Gaussians do not overlap much at the origin, which produces only a few non-zero DC components. The lower value of Gamma will also help to obtain maximal spatial localization of frequency information of the facial image.

The higher value of maximum central frequency, *fmax* moves the two Gaussian functions further apart so that the overlap does not happen excessively. This is to restrain the frequency value within the scope of Nyquist frequency. The excessive overlapping and higher frequency bandwidth (due to higher Gamma) will cause smaller coverage of the spectrum in the spatial domain. Since 40 Gabor filters are used, the excessive overlapping will cause a narrower spectrum of detected and extracted features, thus lowering the recognition rate. In our design, the maximum frequency of the filter is set to $\frac{\sqrt{2}}{4}$ Hz.

The number of features generated by Gabor filters and MR filters is still large, and it causes high computation costs during the classification process. Thus, 2000 most useful features (best features) are selected through the Monte Carlo Uninformative Variable Elimination PLS Regression method (MC-UVE-PLSR). The selection of the 2000 highest importance features is decided to strike an optimal balance between maintaining optimal classification accuracy level and controlling the computation costs incurred during the classification process. The MC-UVE-PLSR process ranks the features according to their information importance. A subset of features with the highest information importance is selected. The reason for using feature selection is explained as follows. If the number of features fed into the classifier is too many, the classification process consumes disproportionately longer time, and it is unfeasible in terms of limited computation resources. The performance of the proposed algorithm is compared with the full feature selection, where all the features, 4000 features are selected.

3.3. Face Verification Results

The following experiments are conducted on Intel I5-4430S 2.7 GHz processor with an 8 GB memory PC under the Windows 10 operating system. The

algorithms are implemented using MATLAB, R2016. The default MATLAB accuracy is used to report the results. Three decimal points are considered in reporting the verification performance. Four algorithms are proposed and tested on LFW and UFI databases. These are Gabor-Random Forest (full features selection-4000 features), Gabor-Random Forest (best features selection-2000 features), Gabor-MR-Fusion-Random Forest (full features selection-4000 features) and Gabor-MR-Fusion-Random Forest (best features selection-2000 features.

In the first experiment, the image-restricted, no outside training data protocol results are used to compare with our proposed algorithms. Based on Table 1 (in bold text), it is found that the full features selection performs second-best among the compared algorithms in terms of verification accuracy, standing at 95.872% accuracy for full features selection. If the best features are selected, the accuracy is reduced to 92.285%. In the second experiment, the MR filter is omitted, and only Gabor Magnitude Response and OGPC are used in the feature extraction process. The reason for this is to determine the impact of the MR filter against the verification rate. Based on the tabulated results in Table 1, it is found that the proposed algorithm (Gabor-Random Forest) performs worse than the proposed algorithm (Gabor-MR-Random forest) in terms of verification accuracy, standing at 90.746% accuracy, whereas for Gabor-MR-Random forest, standing at 95.872%. If the best features are selected (Gabor-Random forest), the accuracy is reduced to 88.893% and for Gabor-MR-Random forest at 92.285%. These results support our hypothesis that the MR filter contributes in terms of the verification rate. Although full-features selection gives the highest accuracy in both experiments, the computation cost is high. For the Gabor-MR-Random Forest, the best features consume 14471.71 seconds for the feature extraction process, feature selection process and feature classification process. On the other hand, the Gabor-MR-Random Forest (full features selection) consumes 37496.71 seconds for the feature extraction process and feature classification process. It was an increment of 159.1% for full features selection.

In the third experiment, the proposed algorithm is compared with methods under the unsupervised protocol. The comparison results are shown in Table 2. These results are obtained from the respective cited papers. It is found that Gabor-MR-fusion with Random Forest (full features selection) performs second-best among the compared algorithms with 0.9887 Area Under the ROC Curve (AUC). If only half features are selected, the proposed algorithm performs third-best among the existing algorithms in terms of AUC at 0.9865. Although full-features selection gives the highest accuracy in both experiments, the computation cost is high.

Based on Table 3, it is shown that Gabor-MR-Random Forest (full features selection) obtains the highest recognition accuracy at 74.961% compared to other reported existing algorithms for the UFI database. In Table 4, the implemented Gabor-MR-Random Forest algorithm gets 96.01% Area Under ROC (AUC) which is higher than the Hybrid of CBIR.

Table 1. Mean Verification Accuracy on the LFW database (No outside training data used).

Algorithm	Mean Verification Accuracy(%) $\pm S_E$	Computation Time (second)
MRF-Fusion-CSKDA [3]	95.891 ± 0.0194	N/A
ConvNet-RBM [39]	93.831 ± 0.0052	N/A
CVPR13' high-dim LBP + JB [9]	93.182 ± 0.0107	N/A
DM+PCA fusion [4]	92.051 ± 0.0045	N/A
Hierarchical-PEP (layers fusion) [21]	91.106 ± 0.0147	N/A
Joint Bayesian [8]	90.908 ± 0.0148	N/A
Eigen-PEP [22]	88.972 ± 0.0132	N/A
RSF [32]	88.812 ± 0.0078	N/A
Spartans [15]	87.553 ± 0.0021	N/A
BMVC13' Fisher vector faces [36]	87.471± 0.0149	N/A
V1-like/MKL, funneled [29]	79.351 ± 0.0055	N/A
MRF-MLBP [2]	79.082 ± 0.0014	N/A
Hybrid descriptor-based, funnelled [47]	**78.47 ± 0.0051**	**N/A**
Proposed Method-Gabor-Random Forest (full features selection) [1]	**90.746 ± 0.0602**	**33087.575**[3]
Proposed Method-Gabor-Random Forest (best features selection) [2]	**88.983 ± 0.0243**	**12933.979**[4]
Proposed Method-Gabor-MR-Fusion-Random Forest (full features selection) [1]	**95.872 ± 0.0197**	**37496.72**[3]
Proposed Method-Gabor-MR-Fusion-Random Forest(best features selection) [2]	**92.285 ± 0.0198**	**14471.72**[4]

[1]4000 features are selected, [2]2000 features are selected N/A=not available.
[3]Computation time for Full features=Feature extraction time+Classification time.
[4]Computation time for Best features=Extraction time+Selection time+Classification time).

Table 2. Comparison of AUC among algorithms on the LFW database (unsupervised setting).

Algorithm	AUC
MRF-Fusion-CSKDA [3]	0.9894
Spartans [15]	0.9428
Pose Adaptive Filter (PAF) [49]	0.9405
LBPNet [48]	0.9404
SA-BSIF, WPCA, aligned [50]	0.9318
MRF-MLBP [2]	0.8994
LHS, aligned [34]	0.8107
LARK unsupervised, aligned [33]	0.7830
H-XS-40, 81x150 [31]	0.7547
GJD-BC-100, 122x225 [31]	0.7392
Proposed Method-Gabor-MR-Fusion-Random Forest (full features selection) [1]	**0.9887**
Proposed Method-Gabor-MR-Fusion-Random Forest (best features selection) [2]	**0.9865**

[1]4000 features are selected, [2]2000 features are selected.

Table 3. Mean recognition accuracy on the UFI database.

Algorithm	Mean Recognition Accuracy (%)
POEMHS [46]	67.11
Enhanced Local Binary Patterns [17]	65.28
FS-LBP [18]	63.31
SIFT [25]	61.32
LBPHS [1]	55.04
MGM+LBP Edge-mapped [23]	51.07
LDPHS [51]	50.25
Proposed Method-Gabor-Random Forest (full features selection) [*1]	**69.348**
Proposed Method-Gabor-Random Forest (best features selection) [*2]	**66.242**
Proposed Method-Gabor-MR-Fusion-Random Forest (full features selection) [*1]	**74.961**
Proposed Method-Gabor-MR-Fusion-Random Forest (best features selection) [*2]	**71.891**

[*1]4000 features are selected, [*2]2000 features are selected.

Table 4. Comparison of AUC among algorithms on UFI database.

Algorithm	AUC
Hybrid of CBIR and SVM [11]	0.9542
Proposed Method-Gabor-MR-Fusion-Random Forest (full features selection) [*1]	**0.9601**
Proposed Method-Gabor-MR-Fusion-Random Forest (best features selection) [*2]	**0.9325**

[*1]4000 features are selected, [*2]2000 features are selected.

4. Conclusions

In conclusion, the proposed method is Gabor-Maximum Response Filters with Random Forests Classifiers. The Gabor filters utilize the features generated from magnitude and phase responses. Maximum Response filters exploit features generated by edge, bar and isotropic filters. Random Forests is used as a classifier. The biometric performance is evaluated on two unconstrained facial image databases which are LFW and UFI. The proposed method showed better performance than several existing algorithms such as Enhanced Local Binary Patterns and SIFT. It is also found that the trade-off between selecting the full features and best features in terms of accuracy and computational cost. In future works, face recognition using deep learning can be further studied as deep learning has advantages in processing unstructured data such as visual images. The method of fusing between Gabor feature extraction and deep learning is suggested because this method trains the system to learn a small set of filters to produce an enhanced deep learning model. The fusion has the advantages of high accuracy of deep learning and the capability of Gabor filters in extracting the key features in a faster way. This could save computation costs and time even further.

The Matlab code of the proposed method can be obtained at the following link

https://drive.google.com/drive/folders/1GVvDJNah Ag1JJDwuGvp19HYcjYqzurFE.

References

[1] Ahonen T., Hadid A., and Pietikainen M., "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, 2006.

[2] Arashloo S. and Kittler J., "Efficient Processing of Mrfs for Unconstrained-Pose Face Recognition," *in Proceedings of 6th International Conference on Biometrics: Theory, Applications and Systems*, Arlington, pp. 1-8, 2013.

[3] Arashloo S. and Kittler J., "Class-Specific Kernel Fusion of Multiple Descriptors for Face Verification Using Multiscale Binarised Statistical Image Features," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2100-2109, 2014.

[4] Barkan O., Weill J., Wolf L., and Aronowitz H., "Fast High Dimensional Vector Multiplication Face Recognition," *in Proceedings of the IEEE International Conference on Computer Vision*, Sydney, pp. 1960-1967, 2013.

[5] Benitez-Garcia G., Nakano-Miyatake M., Olivares-Mercado J., Perez-Meana H., Sanchez-Perez G., and Toscano-Medina K., "A Low Complexity Face Recognition Scheme Based on Down Sampled Local Binary Patterns," *The International Arab Journal of Information Technology*, vol. 16, no. 3, pp. 338-347, 2019.

[6] Caenen G. and Gool L., "Maximum Response Filters for Texture Analysis," *in Proceedings of Conference on Computer Vision and Pattern Recognition Workshop*, Washington, pp. 58-58, 2004.

[7] Cai W., Li Y., and Shao X., "A Variable Selection Method Based on Uninformative Variable Elimination for Multivariate Calibration of Near-Infrared Spectra," *Chemometrics and Intelligent Laboratory Systems*, vol. 90, no. 2, pp. 188-194, 2008.

[8] Chen D., Cao X., Wang L., Wen F., and Sun J., "Bayesian Face Revisited: A Joint Formulation," *in Proceedings of European Conference on Computer Vision*, Florence, pp. 566-579, 2012.

[9] Chen D., Cao X., Wen F., and Sun J., "Blessing of Dimensionality: High-Dimensional Feature and its Efficient Compression for Face Verification," *in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Portland, pp. 3025-3032, 2013.

[10] Déniz O., Bueno G., Salido J., and Torre F., "Face Recognition using Histograms of Oriented Gradients," *Pattern Recognition Letters*, vol. 32, no. 12, pp. 1598-1603, 2011.

[11] Devi N. and Hemachandran K., "Content based Feature Combination Method for Face Image Retrieval using Neural Network and SVM Classifier for Face Recognition," *Indian Journal of Science and Technology*, vol. 10, no. 24, pp. 1-11, 2017.

[12] Faraji M., Shanbehzadeh J., Nasrollahi K., and Moeslund T., "Extremal Regions Detection Guided by Maxima of Gradient Magnitude," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5401-5415, 2015.

[13] Geusebroek J., Smeulders A., and Van-De-Weijer J., "Fast Anisotropic Gauss Filtering," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 938-943, 2003.

[14] Huang G., Ramesh M., Berg T., and Learned-Miller E., "Labeled Faces in The Wild: A Database for Studying Face Recognition in Unconstrained Environments," Technical Report, University of Massachusetts, 2007.

[15] Juefei-Xu F., Luu K., and Savvides M., "Spartans: Single-Sample Periocular-Based Alignment-Robust Recognition Technique Applied to Non-Frontal Scenarios," *IEEE Transactions on Image Processing,* vol. 24, no. 12, pp. 4780-4795, 2015.

[16] Kovesi P., "Phase Congruency: A Low-Level Image Invariant," *Psychological Research*, vol. 64, no. 2, pp. 136-148, 2000.

[17] Král P. and Vrba A., "Enhanced Local Binary Patterns for Automatic Face Recognition," *arXiv preprint arXiv:1702.03349*, 2017.

[18] Lenc L. and Král P., "Automatically Detected Feature Positions for LBP Based Face Recognition," *in Proceedings of IFIP International Conference on Artificial Intelligence Applications and Innovations*, Rhodos, pp. 246-255, 2014.

[19] Lenc L. and Král P., "Unconstrained Facial Images: Database for Face Recognition Under Real-World Conditions," *in Proceedings of Mexican International Conference on Artificial Intelligence*, Cuernavaca, pp. 349-361, 2015.

[20] Lenc L. and Král P., "Local Binary Pattern Based Face Recognition with Automatically Detected Fiducial Points," *Integrated Computer-Aided Engineering*, vol. 23, no. 2, pp. 129-139, 2016.

[21] Li H. and Hua G., "Hierarchical-PEP Model for Real-World Face Recognition," *in Proceedings of the Conference on Computer Vision and Pattern Recognition*, Boston, pp. 4055-4064, 2015.

[22] Li H., Hua G., Shen X., Lin Z., and Brandt J., "Eigen-Pep for Video Face Recognition," *in Proceedings of Asian Conference on Computer Vision*, Singapore, pp. 17-33, 2014.

[23] Lin J. and Chiu C., "LBP Edge-Mapped Descriptor Using MGM Interest Points for Face Recognition," *in Proceedings of International Conference on Acoustics, Speech and Signal Processing*, New Orleans, pp. 1183-1187, 2017.

[24] Liu W. and Wang Z., "Facial Expression Recognition Based on Fusion of Multiple Gabor Features," *in Proceedings of 18th International Conference on Pattern Recognition*, Hong Kong, pp. 536-539, 2006.

[25] Lowe D., "Distinctive Image Features From Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[26] Lu C. and Tang X., "Surpassing Human-Level Face Verification Performance on LFW with Gaussian Face," *in Proceedings of 29th AAAI Conference on Artificial Intelligence*, pp. 3811-3819, 2015.

[27] Muruganantham S. and Jebarajan T., "A Comprehensive Review of Significant Researches on Face Recognition Based on Various Conditions," *International Journal of Computer Theory and Engineering*, vol. 4, no. 1, pp. 7-15, 2012.

[28] Ojala T., Pietikainen M., and Maenpaa T., "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.

[29] Pinto N., DiCarlo J., and Cox D., "How Far Can You Get with A Modern Face Recognition Test Set using Only Simple Features?," *in Proceedings of Conference on Computer Vision and Pattern Recognition*, Miami, pp. 2591-2598, 2009.

[30] Quah K. and Quek C., "MCES: A Novel Monte Carlo Evaluative Selection Approach for Objective Feature Selections," *IEEE Transactions on Neural Networks*, vol. 18, no. 2, pp. 431-448, 2007.

[31] Ruiz-del-Solar J., Verschae R., and Correa M., "Recognition of Faces in Unconstrained Environments: A Comparative Study," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, pp.1-19, 2009.

[32] Sagonas C., Panagakis Y., Zafeiriou S., and Pantic M., "Robust Statistical Frontalization of Human and Animal Faces," *International journal of Computer Vision*, vol. 122, no. 2, pp. 270-291, 2017.

[33] Seo H. and Milanfar P., "Face Verification Using The Lark Representation," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 4, pp. 1275-1286, 2011.

[34] Sharma G., Hussain ul S., and Jurie F., "Local Higher-Order Statistics (LHS) for Texture Categorization and Facial Analysis," *in Proceedings of European Conference on Computer Vision*, Florence, pp. 1-12, 2012.

[35] Shen L. and Bai L., "A Review on Gabor Wavelets for Face Recognition," *Pattern Analysis and Applications*, vol. 9, no. 2, pp. 273-292, 2006.

[36] Simonyan K., Parkhi O., Vedaldi A., and Zisserman A., "Fisher Vector Faces in the Wild," *in Proceedings of British Machine Vision Conference*, pp. 1-13, 2013.

[37] Štruc V. and Pavešić N., "Gabor-Based Kernel Partial-Least-Squares Discrimination Features For Face Recognition," *Informatica*, vol. 20, no. 1, pp. 115-138, 2009.

[38] Štruc V. and Pavešić N., "The Complete Gabor-Fisher Classifier for Robust Face Recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 1, pp. 1-26, 2010.

[39] Sun Y., Wang X., and Tang X., "Hybrid Deep Learning for Face Verification," *in Proceedings of International Conference on Computer Vision*, Sydney, pp. 1489-1496, 2013.

[40] Taigman Y., Yang M., Ranzato M., and Wolf L., "Deepface: Closing The Gap to Human-Level Performance in Face Verification," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, pp. 1701-1708, 2014.

[41] Tan X. and Triggs B., "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635-1650, 2010.

[42] Tan Y., Qi J., and Ren F., "Real-Time Cloud Detection in High Resolution Images using Maximum Response Filter and Principle Component Analysis," *in Geoscience and Remote Sensing Symposium*, Beijing, pp. 6537-6540, 2016.

[43] Turhan C. and Bilge H., "Class-Wise Two-Dimensional PCA Method for Face Recognition," *IET Computer Vision*, vol. 11, no. 4, pp. 286-300, 2016.

[44] Turk M. and Pentland A., "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.

[45] Vu N., "Exploring Patterns of Gradient Orientations and Magnitudes for Face Recognition," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 2, pp. 295-304, 2013.

[46] Vu N., Dee H., and Caplier A., "Face Recognition using the POEM Descriptor," *Pattern Recognition*, vol. 45, no. 7, pp. 2478-2488, 2012.

[47] Wolf L., Hassner T., and Taigman Y., "Descriptor Based Methods in the Wild," *in Workshop on Faces In'real-Life'images: Detection, Alignment, and Recognition*, 2008.

[48] Xi M., Chen L., Polajnar D., and Tong W., "Local Binary Pattern Network: A Deep Learning Approach for Face Recognition," *in Proceedings of International Conference on Image Processing*, Phoenix, pp. 3224-3228, 2016.

[49] Yi D., Lei Z., and Li S., "Towards pose Robust Face Recognition," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Portland, pp. 3539-3545, 2013.

[50] Ylioinas J., Kannala J., Hadid A., and Pietikäinen M., "Face Recognition using Smoothed High-Dimensional Representation," *in Proceedings of Scandinavian Conference on Image Analysis*, Copenhagen, pp. 516-529, 2015.

[51] Zhang B., Gao Y., Zhao S., and Liu J., "Local Derivative Pattern Versus Local Binary Pattern: Face Recognition with High-Order Local Pattern Descriptor," *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 533-544, 2010.

[52] Zhang B., Shan S., Chen X., and Gao W., "Histogram of Gabor Phase Patterns (Hgpp): A Novel Object Representation Approach for Face Recognition," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 57-68, 2007.

Yuen-Chark See is an Assistant Prorfessor in Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Sungai Long Campus. He received his PhD from Universiti Teknologi Malaysia. His research interests include machine learning, embedded systems, and wireless sensor networks

Eugene Liew received B.Eng in Electrical and Electronic Engineering from Universiti Tunku Abdul Rahman.

Norliza Mohd Noor is a Professor in Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia (UTM), Kuala Lumpur Campus. She received her B.Sc. in Electrical Engineering from Texas Tech University in Lubbock, Texas, and Master (by research) and PhD both in Electrical Engineering from UTM. Her research is in machine learning and image analysis for medical and industry applications.