

MULTI-FOLD GABOR FILTER CONVOLUTION DESCRIPTOR FOR FACE RECOGNITION

Cheng-Yaw Low, Andrew Beng-Jin Teoh, Cong-Jie Ng

School of Electrical and Electronic Engineering, College of Engineering
Yonsei University, Seoul, SOUTH KOREA
{chengyawlow, bjteoh, congjie} @yonsei.ac.kr

ABSTRACT

The standard multi-scale, multi-orientation Gabor filter ensemble (SGFE) in the face recognition task reposit 40 filters localized in 8 orientations and 5 scales, with a real and an imaginary constituent. This paper devises a simple means of filter diversification, dubbed as multi-fold Gabor filter convolution (\mathcal{M} -FGFC), where a set of pre-selected filters, e.g., single-scale Gabor filters across varying orientations, are self-cross convolved by \mathcal{M} folds to instantiate the offspring filters. To facilitate filter selection for \mathcal{M} -FGFC, this paper summarizes SGFE into the condensed Gabor filter ensemble (CGFE) of only 8 filters. In addition, an average histogram pooling operator is proposed to downsample and regulate the demodulated Gabor phase features prior to the final compression stage. The performance of a specific \mathcal{M} -FGFC instance, i.e., the 2-FGFC descriptor, is investigated on FERET I (frontal), FERET II (non-frontal) and AR datasets. The experimental results on FERET I substantiates that the 2-FGFC descriptor outperforms the leading state of the art face descriptors.

Index Terms— Gabor filters, self-cross filter convolution, face recognition, biometrics

1. INTRODUCTION

Face recognition is a challenging task due to the intra-class variability arisen from misalignment, non-rigid deformations, illuminations, occlusions, etc. An ideal face descriptor should be invariant to these intra-class variations whilst magnifying the inter-class margin. This paper focuses on the filter bank-based approaches that share the common 3-stage pipeline: (1) convolution of input images and filters, where the filters are either defined by mathematical functions [1-4], or pre-learned from the training specimens [5-7]; (2) a non-linear operation and encoding; (3) local histogramming; in which coined *spectral histogram technique* collectively [8-9].

The standard Gabor filter ensemble (SGFE) adopted in the face recognition literature encompasses 40 filters localized in 8 orientations and 5 scales [1], [10-15]. One of the earliest known works involving Gabor face representation is introduced by Liu and Wechsler [1]. Instead of histogramming the extracted features, Liu and Wechsler apply the enhanced Fisher linear discriminant (FLD) to the concatenation of Gabor responses. In contrast, Zhang et al. [10] and Zhang et al. [11] make use of the spectral histogram technique, where the former utilizes the local binary pattern (LBP) operator to encode Gabor magnitude while the latter exploits the local XOR pattern (LXP) operator on Gabor phase. Other similar

methods are: Xie et al. [12] fuses the local patterns derived from Gabor magnitude and phase, followed by block-based FLD; Lei et al. [13] conglomerates the local variations in the spatial, scale and orientation domains. This paradigm (applying LBP or LXP in the Gabor domain), on average, have shown significant improvement over their individual representation. Hussain [14], on the contrary, histograms the binary and ternary vectors extracted from the local neighborhoods of the Gabor responses with respect to a codebook learned using the K-mean algorithm.

The latest spectral histogram technique - PCANet [5] employs the cascaded principal component analysis (PCA) to pre-learn the multistage filter banks. The empirical performance discloses that a simplistic 2-stage net topology performs unexpectedly well on the generic image recognition, including face. Different from PCANet, the binarized statistical image features method (BSIF) [6] exercises the independent component analysis (ICA) to pre-learn a filter set from 13 natural images. BSIF, however, demands an illumination normalization step proposed by Tan and Triggs [15] to normalize the texture and face images. Other well-known filters in texture classification include Leung Malik [2], Schmid [3], MR8 [4], etc.

Inspired by the PCANet performance and the Gabor filters that model the receptive field profiles of mammalian cortical simple cells [16-17], this paper outlines a new means of simple filter diversification by self-cross convolving the Gabor filters via multiple folds. The major contribution is threefold: (1) \mathcal{M} -fold Gabor filter convolution (\mathcal{M} -FGFC) is devised to yield offspring filters of numerous traits from a set of pre-selected Gabor filters. (2) this paper represents the 40 SGFE filters as a condensed Gabor filter ensemble (CGFE) of only 8 elementary filters; (3) the Gabor phase feature is leveraged by an average pooling unit to alleviate the dimensionality problem. The generic \mathcal{M} -FGFC pipeline is portrayed in **Fig. 1**.

2. GABOR WAVELET

Gabor wavelets, acclaimed to be optimally localized in the spatial and frequency domains, permit the local structures of input images to be extracted with respect to the tuned spatial frequencies (scale) and orientations [1]. The 2D Gabor wavelets of u orientations and v scales, in practice of face recognition, are defined as follows:

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{(-\|k_{u,v}\|^2 \|z\|^2 / 2\sigma^2)} [e^{ik_{u,v}z} - e^{-\sigma^2/2}] \quad (1)$$

where $u \in \{0, \dots, 7\}$, $v \in \{0, \dots, 4\}$, $z = (x, y)$, σ refers to the Gabor envelop width set to 2π in the experiments. The wavelet vector, on the other hand, is denoted as $k_{u,v} = k_v e^{i\theta_u}$, where $k_v = k_{max}/f^v$, $k_{max} = \pi/2$, $f = \sqrt{2}$, and $\theta_u = u\pi/8$.

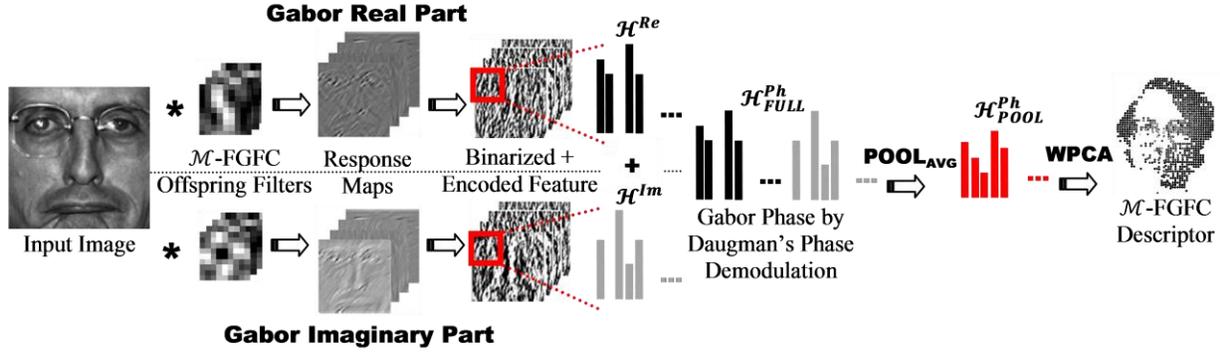


Fig. 1. The generic \mathcal{M} -FGFC pipeline: convolution, non-linear operation and encoding, local histogramming. The demodulated Gabor phase representation is averagely pooled via POOL_{AVG} and PCA whitened to construct the compact \mathcal{M} -FGFC descriptor.

The Gabor response of an arbitrary image $I(\cdot)$ is described as the convolution of $I(\cdot)$ with each Gabor filter in the spatial domain as follows:

$$\mathcal{G}_{u,v}(z) = I(z) * \psi_{u,v}(z) \quad (2)$$

where $*$ denotes the convolution operator. For each Gabor filter, a real and an imaginary map are responded: $\mathcal{G}_{u,v}^{\text{Re}}(z)$ and $\mathcal{G}_{u,v}^{\text{Im}}(z)$.

3. MULTI-FOLD GABOR FILTER CONVOLUTION

3.1. Condensed Gabor Filter Ensemble Formulation

Gabor filter-bank methods, e.g., [1], [10-15], suffer from the severe dimensionality issue due to the concatenation of filter responses. Hence, this paper condenses the 40 multi-scale, multi-orientation Gabor filters in SGFE (refer to Ψ in (1)) by averaging the filters of certain orientation across varying spatial frequencies (scales) to define CFGE, Φ , of only u_{max} filters as follows:

$$\Phi = \left\{ \varphi_{\mathcal{F}} : \frac{1}{v_{\text{max}}} \sum_{v=0}^{v_{\text{max}}-1} \psi_{u,v}, u = \mathcal{F} - 1 \right\}_{\mathcal{F}=1}^{u_{\text{max}}} \quad (3)$$

where $u_{\text{max}} = 8$, $v_{\text{max}} = 5$ and $\psi_{\mathcal{F}} \in R^{k \times k}$ in this paper. **Fig. 2(a)** depicts that the 8 elementary CGFE filters are in the initial 8 orientations.

3.2. \mathcal{M} -FGFC Filtering

For \mathcal{M} -FGFC, let $\left\{ \varphi_{\mathcal{F}_m}^{(m)} \in R^{k \times k} \right\}_{\mathcal{F}_m=1}^{\mathcal{F}_m}$ be \mathcal{M} sets of CGFE filters defined in (3), where $m = 1, \dots, \mathcal{M}$ and $\mathcal{F}_m \leq u_{\text{max}}$; these CGFE filters are *self-cross convolved* by m -fold to yield $\mathcal{L} = \prod_{m=1}^{\mathcal{M}} \mathcal{F}_m$ offspring filters \mathcal{O} as follows:

$$\mathcal{O} = \left\{ \sigma_{\ell} \in R^{\mathcal{K} \times \mathcal{K}} : \varphi_{\mathcal{F}_1}^{(1)} * \varphi_{\mathcal{F}_2}^{(2)} * \dots * \varphi_{\mathcal{F}_M}^{(M)} \right\}_{\ell=1}^{\mathcal{L}} \quad (4)$$

where $\mathcal{F}_1 = \{1, \dots, \mathcal{F}_1\}$, $\mathcal{F}_2 = \{1, \dots, \mathcal{F}_2\}$, ..., $\mathcal{F}_M = \{1, \dots, \mathcal{F}_M\}$ and $\mathcal{K} = \mathcal{M}(k-1) + 1$. For 2-FGFC with $\mathcal{M} = 2$ and $\mathcal{F}_1 = \mathcal{F}_2 = 8$, the input filters of size 3×3 (in each fold) return $\mathcal{L} = 64$ ($=8 \times 8$) offspring filters of size 5×5 ; on the other hand, for 3-FGFC with $\mathcal{M} = 3$ and $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{F}_3 = 8$, the input filters of the same size produces $\mathcal{L} = 512$ ($8 \times 8 \times 8$) offspring filters of size 7×7 . Let an arbitrary image $I \in R^{\ell \times w}$ be a set of zero-mean local patches $\{ \mathcal{P}_n \in R^{\mathcal{K} \times \mathcal{K}} \}_{n=1}^{\mathcal{N}}$ and $\mathcal{N} = \ell \times w$; I is convolved with each offspring σ_{ℓ} to yield \mathcal{L} responses as follows:

$$\mathcal{G} = \{ \mathcal{G}_{\ell} : I * \sigma_{\ell} \}_{\ell=1}^{\mathcal{L}} \quad (5)$$

Fig. 2(b) depicts the Gabor offspring set of 64 ($=8 \times 8$) filters resulted from 2-FGFC with respect to the 8 CGFE filters (in each

fold) provided in **Fig. 2(a)**. It is observed that these offspring filters are implanted with distinctive characteristics in the spatial domain. The filters of checkerboard-like structure, for example, are formed as a consequence of the filter convolution in two directions. Hence, aside from being sensitive to the local edges (just as 1-fold Gabor filters), the offspring filters are also capable of featuring the latent textures consistent with the corresponding spatial frequencies. In other words, the offspring filters serve dual role: edge and texture detectors. This remark is to be validated through extensive studies and experiments. Since the offspring filters are redundant (the self-cross filter convolution operation is commutative), the offspring set can be pruned to 36 ($=(8+1)8/2$) unique filters.

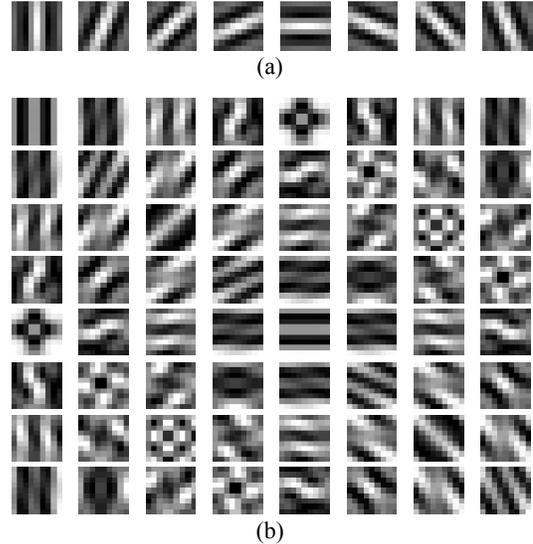


Fig. 2. The 8 elementary CGFE filters (a) are diversified via 2-FGFC to generate 64 offspring (b), with only 36 unique filters.

In summary, for \mathcal{M} -FGFC, \mathcal{L} responses are obtained for each real and imaginary constituent. In practice, \mathcal{M} is to be determined empirically for certain degree of performance gain resulting from filter diversification. The major downside of \mathcal{M} -FGFC is that \mathcal{L} increases exponentially with respect to \mathcal{M} . This paper, however, only restricts the discussion to 1-FGFC and 2-FGFC.

3.3. Binarization and Histogramming

Subsequent to (5), the resultant real and imaginary filter responses are binarized via a thresholding phase with respect to 0, i.e., a non-

linear operation, followed by feature encoding to bin the extracted local structures into histograms. This is primarily due to the reason that histogram representation offers translation invariance, to some extent. The full properties of the spectral histogram technique are detailed in [8].

\mathcal{M} -FGFC yields \mathcal{L} responses to be binned into $\mathcal{J} = \prod_{m=1}^{\mathcal{M}-1} \mathcal{F}_m$ histograms. To be specific, \mathcal{G} in (5) is first quantized into \mathcal{F}_m -bit integers, ranging from $[0, 2^{\mathcal{F}_m} - 1]$ to define $\mathcal{D}^{(i)}$ as follows:

$$\mathcal{D}^{(i)} = \sum_{\beta_{\mathcal{M}}=1}^{\mathcal{F}_m} \mathcal{S}(\mathcal{G}_{(i \times \mathcal{F}_m) - \mathcal{F}_m + \beta_{\mathcal{M}}}) \cdot 2^{\beta_{\mathcal{M}} - 1} \quad (6)$$

where $i = 1, \dots, \mathcal{J}$ and $\mathcal{S}(\cdot)$ refers to the Heaviside step function. Each $\mathcal{D}^{(i)}$ is subsequently separated into \mathcal{B} non-overlapping blocks of $x \times y$, where $b = 1, \dots, \mathcal{B}$. For each b block, the occurrences for the $2^{\mathcal{F}_m}$ bins are aggregated as follows:

$$\mathcal{H}_b^{(i)} = \sum_{x,y} \delta(\alpha, \mathcal{D}_b^{(i)}(x, y)), \alpha = 0, \dots, 2^{\mathcal{F}_m - 1} \quad (7)$$

where $\delta(\cdot)$ is the Kronecker delta function. The \mathcal{B} local histograms for $\mathcal{D}^{(i)}$ are concatenated to form $\mathcal{H}^{(i)} \in R^{(2^{\mathcal{F}_m}) \times \mathcal{B}}$ and the global histogram defining the \mathcal{M} -FGFC descriptor is derived as follows:

$$\mathcal{H} = [\mathcal{H}^{(i=1)}, \dots, \mathcal{H}^{(i=\mathcal{J})}] \in R^{(2^{\mathcal{F}_m}) \times \mathcal{B} \times \mathcal{J}} \quad (8)$$

In our complex Gabor case, the global histogram features for the real and imaginary parts, denoted by \mathcal{H}^{Re} and \mathcal{H}^{Im} in **Fig. 1**, are cascaded and vectorized to represent the corresponding Gabor phase feature $\mathcal{H}_{FULL}^{Ph} \in R^{(2^{\mathcal{F}_m}) \times \mathcal{B} \times \mathcal{J} \times 2}$ based on the Daugman's phase-quadrant demodulation code described in [18]. As the \mathcal{H}_{FULL}^{Ph} formulation doubles the feature dimensionality, an average pooling operator is devised to alleviate this issue.

3.4. Average Spatial Histogram Pooling

POOL_{AVG} is an average pooling operator proposed to downsample and regulate the underlying probability density function defined by the global histogram feature \mathcal{H} . This signifies that, in addition to dimension reduction, the POOL_{AVG} operator also uniformizes the histogram distribution such that the burstiness and the zero-valued bins, to some extent, are rectified.

Let \mathcal{H} be a D -dimensional histogram feature to be pruned onto \mathcal{H}' of D' dimensions; the POOL_{AVG} operator manipulates every single element of \mathcal{H} as follows:

$$\mathcal{H}'_d = \frac{1}{P} \sum_{n=1}^P \mathcal{H}_{(d-1) \times S + n}, d = 1, \dots, D' \quad (9)$$

where $D = (2^{\mathcal{F}_m}) \times \mathcal{B} \times \mathcal{J}$, $D' = \frac{D-P}{S} + 1$, and P and S are pooling window size and the stride step, respectively. In the experiments, the $(2^{\mathcal{F}_m}) \times \mathcal{B} \times \mathcal{J} \times 2$ -dimensional \mathcal{H}_{FULL}^{Ph} is in turn square-rooted and averagely pooled by POOL_{AVG} to elicit \mathcal{H}_{POOL}^{Ph} of $(2^{\mathcal{F}_m}) \times \mathcal{B} \times \mathcal{J}$ dimensions. After that, \mathcal{H}_{POOL}^{Ph} is L2-normalized and whitening PCA (WPCA) is followed to yield a globally compact \mathcal{M} -FGFC descriptor.

4. EXPERIMENTS

4.1. Benchmarking Datasets: FERET I, FERET II, AR

The performance of the \mathcal{M} -FGFC descriptor is benchmarked based on the FERET evaluation protocol on both FERET I (frontal) and FERET II (non-frontal) [19], and our own protocol on AR [20].

- FERET I furnishes a reference gallery set: FA, with only a single frontal image per subject (1196 images in total); and

4 probe sets: FB, FC, DUP I and DUP II, with expression, illumination and time span variations (with 1195, 194, 722 and 234 images, respectively). These images are re-aligned and cropped into size 128×128 based on the annotated eye coordinates.

- FERET II consists of non-frontal probe faces captured from viewpoint angles of $\pm 40^\circ$, $\pm 25^\circ$, $\pm 15^\circ$, which are labeled as BC, BD, BE, BF, BG and BH. In addition, it also includes frontal images in the BA, BJ and BK repositories employed as references. Note that, each repository accommodates 200 images provided by the same 200 subjects. All images are geometrically pre-processed into 128×128 according to the eyes and mouth coordinates.
- AR, on the other hand, contains frontal faces acquired from 100 subjects in 2 occasions. Each subject supplies 26 faces of various facial expressions: neutral, smile, anger, scream; illumination conditions: left or/and right lighting on; and occlusions: either sun-glasses, or scarf. In our experiments, each image is cropped into 165×120 . For each subject, only 2 images of neutral expression are on the gallery list; and the remaining 24 images act as probes.

4.2. Implementation Summary

In a nutshell, for 2-FGFC that accommodates 8 filters in each fold, a probe is convolved with each of the offspring filters to produce 36 distinctive responses to be in turn zero-thresholded and encoded into the 8-bit integer outputs. Subsequent to that, these outputs, for all FERET I, FERET II and AR images, are regionalized into 8×8 non-overlapping blocks and the discriminative features are locally quantized into 8 histograms. These histograms are concatenated to form a global feature of 131, 072 dimensions (for each real and imaginary part): $64 \text{ blocks} \times 2^8 \text{ histogram bins} \times 8 \text{ histograms}$. The demodulated Gabor phase feature, thus, encapsulates 262, 144 dimensions, to be averagely halved via POOL_{AVG} with $P = S = 2$. The 2-FGFC performance is evaluated on the WPCA compressed feature based on the Cosine similarity scores.

4.3. Performance Analysis on CGFE

FERET I is utilized as a testbed to compare the CGFE performance to SGFE, where 1-FGFC-8 and 1-FGFC-40 denote the descriptors resulted from the 8 CGFE and 40 SGFE filters, respectively. The 1-FGFC-40 descriptor, in general, resembles GGPP in [11]. What distinguish 1-FGFC-40 from GGPP is the former applies WPCA to compress the demodulated Gabor phase representation onto 1000 dimensions. The rank-1 recognition rates (%) summarized in **Table 1** disclose that, despite of only 8 filters, the performance of the 1-FGFC-8 descriptor is on par with 1-FGFC-40 for the filter sizes of 7×7 , 9×9 and 11×11 .

Table 1. Performance comparison for 1-FGFC-8 of 8 filters and 1-FGFC-40 of 40 filters, in terms of rank-1 recognition rates (%).

DESCR.	FB	FC	DUP I	DUP II	MEAN
1-FGFC-40 _{7x7}	99.16	98.97	91.69	88.03	94.46
1-FGFC-40 _{9x9}	99.00	98.97	92.52	89.32	94.95
1-FGFC-40 _{11x11}	98.91	99.48	91.14	86.75	94.07
1-FGFC-8 _{7x7}	99.16	99.48	91.83	86.75	94.31
1-FGFC-8 _{9x9}	99.25	98.97	91.41	87.61	94.31
1-FGFC-8 _{11x11}	99.16	99.48	90.58	86.75	94.00

4.4. Performance Evaluation on FERET I

Table 2 compares the performance of the 2-FGFC descriptor, in terms of rank-1 recognition rate (%), to the notable state of the arts on FERET I. Overall, the 2-FGFC descriptor of 1000 dimensions prevails over the traditional 1-fold Gabor descriptors, specifically, HGPP [11], LGBP+LGXP [12], E-GV-LBP [13] and G-LQP [14]. The merit of the 2-FGFC descriptor is that it pursues a simple 3-stage architecture with only 36 distinctive filters for each real and imaginary constituent. G-LQP, on the other hand, necessitates an external code book to encode the thresholded filter responses.

PCANet_{MP} [5], PCANet_{FERET} [5] and BSIF [6] are re-implemented using the MultiPIE-learned filters (shared by PCANet’s authors), the FERET-learned filters, and the ICA filters pre-learned from 13 natural images (shared by BSIF’s authors), respectively. Due to the reason that these filters are cultivated via an explicit learning phase in accordance with the training images available, the learning-free 2-FGFC descriptor, therefore, appears to be parsimonious. Another discovery is the BSIF performance drops drastically, from 92.02% to 69.62%, if the Tan & Trigg’s illumination normalization [15] is withdrawn from its pipeline. Moreover, it is also demonstrated that the 2-FGFC descriptor outperforms the recently proposed learning-based LBP descriptors: DFD [21] and Cbfd [22], especially in the most challenging DUP I and DUP II subsets.

Table 2. Performance summary for 2-FGFC and state of the arts, in terms of rank-1 recognition rate (%), on **FERET I**, where * denotes the results reported in the original papers.

DESCR.	FB	FC	DUP I	DUP II	MEAN
* HGPP [11] (TIP, 2007)	97.50	99.50	79.50	77.80	88.58
* LGBP+LGXP [12] (TIP, 2010)	99.00	99.00	94.00	93.00	96.25
* E-GV-LBP [13] (TIP, 2011)	98.41	98.97	81.99	81.62	90.25
* G-LQP [14] (BCMV, 2012)	99.99	100	93.20	91.00	96.05
BSIF [6] + Tan & Trigg [15] (ICPR 2012)	95.98	99.48	86.29	86.32	92.02
BSIF [6] (ICPR 2012)	94.73	59.79	68.84	55.13	69.62
PCANet _{MP,5x5} [5] (TIP, 2015)	99.25	100	94.46	93.16	96.72
PCANet _{FERET,5x5} [5] (TIP, 2015)	99.16	100	94.04	92.31	96.38
* DFD [21] (TPAMI, 2014)	99.40	100	91.80	92.30	95.88
* Cbfd [22] (TPAMI, 2015)	99.80	100	93.50	93.20	96.63
2-FGFC_{13x13}	99.41	100	95.98	94.02	97.35
2-FGFC _{17x17}	99.41	99.48	95.57	93.59	97.01
2-FGFC _{21x21}	99.16	99.48	95.15	93.59	96.85

4.5. Performance Evaluation on FERET II

The performance of the 2-FGFC descriptor against pose variations is examined on FERET II with respect to varying Gabor offspring filters of 13×13 to 21×21 , where the global histogram features are WPCA-ed onto 300 dimensions. **Table 3** displays that the 2-

FGFC descriptor is relatively sensitive against awful pose angles of $\pm 40^\circ$. Some FERET II exemplars are shown in **Fig. 3** for reference.



Fig. 3. Some exemplars extracted from FERET II.

Table 3. Performance summary for 2-FGFC, in terms of rank-1 recognition rate (%), on **FERET II**.

DESCR.	BC +40°	BD +25°	BE +15°	BF -15°	BG -20°	BH -40°	MEAN
2-FGFC _{13x13}	79.50	99.50	100	100	96.00	65.00	90.00
2-FGFC _{17x17}	88.00	100	100	100	98.50	75.00	93.58
2-FGFC_{21x21}	84.00	100	100	100	98.50	82.00	94.08

4.6. Performance Evaluation on AR

Table 4 lists the rank-1 recognition rate (%) for the PCA whitened 2-FGFC descriptor of 180 dimensions, with respect to the Gabor offspring filters of 13×13 , 17×17 and 21×21 . The 2-FGFC descriptor, on the whole, shows remarkable robustness, particularly to illumination, and also the sun-glasses and scarf disguises. It is noteworthy that our protocol (refer to Section 4.1) only includes 2 frontal faces with neutral expression per subject as references.

Table 4. Performance summary for 2-FGFC, in terms of rank-1 recognition rate (%), on **AR**.

DESCR.	EXPR.	ILLM.	OCCL.	MEAN
2-FGFC _{13x13}	98.82	100	99.24	99.35
2-FGFC _{17x17}	98.65	100	99.92	99.52
2-FGFC_{21x21}	98.82	100	100	99.61

5. CONCLUSIONS

To yield offspring filters of diversified traits, this paper proposes to self-cross convolve the pre-selected Gabor filters by \mathcal{M} -fold (\mathcal{M} -FGFC). This paper also outlines a viable instance to summarize the standard 40 multi-scale, multi-orientation Gabor filters, termed as SGFE, into the condensed Gabor filter ensemble (CGFE) of only 8 filters. The demodulated Gabor phase features are leveraged by an average pooling operator followed by whitening PCA to obtain the final representation. The proposed CGFE is empirically uncovered to be parallel to SCFE. In addition to that, the 2-FGFC descriptor is revealed to exhibit the state of the art recognition performance on the FERET I (frontal), FERET II (non-frontal) and AR datasets, in terms of rank-1 recognition rate (%). It is also evidenced that the 2-FGFC performance surpasses other representative face descriptors, including the most-reputed learning-based LBP variants: DFD and Cbfd. For future work, the \mathcal{M} -FGFC descriptor, for $\mathcal{M} \geq 2$, will be scrutinized on other face datasets, e.g., FRGC, LFW, CMU-PIE, etc, against various circumstances.

6. ACKNOWLEDGEMENT

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by Ministry of Science, ICT and Future Planning (2013006574) and Institute of BioMed-IT, Energy-IT and SmartIT Technology (BEST), a Brain Korea 21 Plus Program, Yonsei University.

7. REFERENCES

- [1] C. Liu, and H. Wechsler, "Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition," *IEEE Trans. on Image Processing*, vol. 11, no. 4, pp. 467-476, Apr. 2002.
- [2] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *International Journal of Computer Vision*, vol. 43, no. 1, pp. 29-44, June 2001.
- [3] C. Schmid, "Constructing models for content-based image retrieval," *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 39-45, 2001.
- [4] M. Varma, and A. Zisserman, "A statistical approach to texture classification from single images," *International Journal on Computer Vision*, vol. 62, pp. 61-81, 2005.
- [5] T. H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: A simple deep learning baseline for image classification?," *IEEE Trans. on Image Processing*, DOI 10.1109/TIP.2015.2475625.
- [6] J. Kannala, and E. Rahtu, "BSIF: Binarized statistical image features," *21st International Conference on Pattern Recognition (ICPR)*, pp. 1363-1366, Nov. 2012.
- [7] R. Jenssen, and T. Eltoft, "ICA filter bank for segmentation of textured images," *4-th International Symposium on Independent Component Analysis and Blind Signal Separation*, pp. 827-832, Apr. 2003.
- [8] X. Liu, and D. Wang, "Texture classification using spectral histograms," *IEEE Trans. on Image Processing*, vol. 12, no. 6, June 2003.
- [9] X. Liu, and D. Wang, "Image and texture segmentation using local spectral histograms," *IEEE Trans. on Image Processing*, vol. 15, no. 10, Oct. 2006.
- [10] W. Zhang, S. Shan, W. Gao, and H. Zhang, "Local Gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition," *Proc. 10th IEEE Int'l Conf. on Computer Vision*, pp. 786-791, 2005.
- [11] B. Zhang, S. Shan, X. Chen, W. Gao, "Histogram of Gabor phase patterns (HGPP): A novel object representation approach for face recognition," *IEEE Trans. on Image Processing*, vol. 16, no. 1, Jan. 2007.
- [12] S. Xie, S. Shan, X. Chen, and J. Chen, "Fusing local patterns of Gabor magnitude and phase for face recognition," *IEEE Trans. on Image Processing*, vol. 19, no. 5. pp. 1349-1361. May 2010.
- [13] Z. Lei, S. Liao, M. Pietikainen, S. Z. Li, "Face recognition by exploring information jointly in space, scale and orientation," *IEEE Trans. on Image Processing*, vol. 20, no. 1, Jan. 2011.
- [14] S. U. Hussain, T. Napoleon, and F. Jurie, "Face recognition using local quantized patterns," *British Conference on Machine Vision*, Sept. 2012.
- [15] X. Tan, and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. on Image Processing*, vol. 19, no. 6, pp. 1635-1650, Jun. 2010.
- [16] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1148-1161, Nov. 1993.
- [17] S. Marčelja, "Mathematical description of the responses of simple cortical cells," *Journal of the Optical Society of America*, vol. 70, no. 11, pp. 1297-1300, 1980.
- [18] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *Journal of the Optical Society of America A*, vol. 2, no. 7, pp. 1160-1169, July 1985.
- [19] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090-1101, Oct. 2000.
- [20] A.M. Martinez and R. Benavente, "The AR Face Database." *CVC Technical Report #24*, June 1998.
- [21] Z. Lei, M. Pietikainen, and S. Z. Li, "Learning Discriminant Face Descriptor," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 289-302, Feb. 2014.
- [22] J. Lu, V. E. Liong, X. Zhou and J. Zhou, "Learning Compact Binary Face Descriptor for Face Recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, no. 1, pp. 1, DOI 10.1109/TPAMI.2015.2408359.