

Score Fusion of SIFT & SURF Descriptors for Face Recognition Using Wavelet Transforms

Musa M.Ameen

Ishik University/Computer Engineering Department, Erbil, 44001, Iraq Email: musa.ameen@ishik.edu.iq

Alaa Eleyan

Avrasya University/Electrical and Electronics Engineering Department, Trabzon, 61000, Turkey Email: aeleyan@avrasya.edu.tr

Received: 09 June 2017; Accepted: 12 September 2017; Published: 08 October 2017

Abstract—Automatic face recognition is a major research area in computer vision which aims to recognize human human intervention. face without Significant developments in this field have shown that in many face recognition applications the automated techniques outperform humans. The conventional Scale-Invariant Feature Transform (SIFT) and Speeded-Up Robust Features (SURF) are used in face recognition where they provide high performances. However, this performance can be improved further by transforming the input into different domains before applying SIFT and SURF algorithms. Hence, we apply Discrete Wavelet Transform (DWT) or Gabor Wavelet Transform (GWT) at the input face images, which provides denser and extra information to be used by the conventional SIFT or SURF algorithms. Matching scores of SIFT or SURF from each subimage is fused before making final decision. Simulations show that the proposed approaches based on wavelet transforms using SIFT or SURF provides very high performance compared to the conventional algorithms.

Index Terms—Speeded-Up Robust Features, Scale-Invariant Feature Transform, Discrete Wavelet Transform, Gabor Wavelet Transform.

I. INTRODUCTION

Face recognition is one of the most common biometric systems. Due to its higher acceptability rate, researchers have developed various algorithms for face recognition purpose. The process of recognition using these algorithms has been described as a difficult task because of the similarity nature or shapes of human faces [1]. Despite the difficulties encountered in designing these systems, several reasons contributed to the enormous attention in automatic digital image processing and video processing in different types of applications, which include wide availability of powerful and low-cost desktop and embedded computing systems. Also, it has been described as one of the best applications of image processing and analysis [2]. Different statistical methods and algorithms such as Principal Component Analysis or Eigenface (PCA) [3], Local Binary Pattern (LBP) [4], Independent Component Analysis (ICA) [5], and triplet half band wavelet filter bank (TWFB) [6] algorithms have been developed for face recognition purposes. In [7] Speed-Up Robust Feature (SURF) and Linear Discriminant Analysis (LDA) are used to improve the quality parameters of face recognition and optimizing the result. Due to continuous research, a significant improvement in recognition performance is obtained over years [8],[10]. Characteristic faces are more easily recognized than typical faces. Low frequency bands contain information that determines the sex of the specific subjects, while recognition of individuals depends on the high frequency features. The global description is determined by the low frequency, while the finer descriptions high frequency modules give to the finer information required for the identification procedure [11],[13].

The core task of this paper work is to investigate how the recognition performance can be enhanced and speeded up. Therefore, image transformation approach is used as a pre-processing stage before the feature extraction stage.

The rest of the paper is organized as follows; Section 2 describes materials, while section 3 explains the methods. Section 4 shows the results. Finally, section 5 includes the discussion and conclusion.

II. MATERIALS

The basics of feature extraction is the dimensionality reduction, by choosing some dominant or distinct features that can best represents the face image with less distortion to the original image. Appropriate algorithms are used to extract the salient features from the relevant patterns. The face representation is done in two ways: the first way is the appearance (holistic) texture features and is applied to the whole face image; the second way is the component based which utilizes the linear relationships between the facial features such as eyes, mouth, and nose. The unpopular component (feature) based approaches utilize some special facial points, and characterize them by applying a bank of filters which extract the typical texture around them [10]. The holistic approaches attract more attention than the component based methods. In this paper, two of the popular holistic or appearance based methods are studied briefly to extract features from face images.

A. Scale-Invariant Feature Transform

Scale-Invariant Feature Transform (SIFT) was developed by D. Lowe [14]. SIFT can detect and extract distinctive features from different face images to achieve robust and stable matching between different face images of the same subject (person) with various facial expressions, face poses, and the features extracted from face images are scale, illumination and rotation invariance. Fig 1. shows four important stages involved for detecting keypoints in the SIFT algorithm.



Fig.1. SIFT features extraction process.

In the initial stage, a difference of Gaussian (DoG) [8] was used to detect specific features and points which are orientation and scale invariance. In the stage of localizing keypoints, they are filtered with a predefined model which is based on their stability. A few orientations are given to the results using local image gradient. In the final stage, around each key point region at different selected scales measurements applied on the image gradients.

B. Speeded-Up Robust Features

Speeded-Up Robust Features (SURF) was created by Bay et al. [15]. SURF algorithm is a robust keypoint detector of local features in a face image. It is a developed version of SIFT and Hessian blob detectors integer approximation to the determinant is calculated with integral images.

In SIFT, Lowe approximated Laplacian of Gaussian (LoG) using DoG for scale-space step. SURF goes a slightly more than Laplacian of Gaussian using Box Filter. Fig. 2. shows approximation demonstration. This approximation's biggest advantage is that; it simply uses integral images to calculate the convolution with box filters. Also for different scales, it can be done in parallel. The determinant of Hessian matrix is a major component of SURF for both position and scale.



Fig.2. The box filters of approximations of Gaussian second order partial derivative.

Orientation assignment achieved using wavelet responses in vertical and horizontal direction for a neighborhood of size 6 multiplied by the scale in which keypoint is detected. Suitable Gaussian weights are also performed on it. The calculation of the sum of all responses within a sliding orientation window of 60 ° estimates the main orientation. The exciting part is that simply integral images can be used to find out wavelet response at any scale. Rotation invariance is not requiring in many applications, so finding this orientation is not needed, by this speed of process increases. SURF delivers an extra method called Upright-SURF or U-SURF which increases speed and is strong up to ± 15 °. Fig. 3(a) and Fig. 3(b) shows distinctive SURF and SIFT keypoints of a face image respectively.



Fig.3. Keypoints detected in a face image using (a) SURF (b) SIFT.

C. Wavelet Transforms

2D-DWT and GWT are mostly used as tunable filters suitable for detecting and extracting orientation information from the image. Apart from orientation, invariant to illumination property makes them appropriate to capture phase information of the pixels. Additionally, it is also an effective method to capture the texture of images [16]. A Gabor wavelet filter is a Gaussian kernel function modulated by a sinusoidal plane wave as in (1).

$$\psi_g(u,v) = \exp\left(-\pi^2 \left(\frac{(u'-f)^2}{\alpha^2} + \frac{v'^2}{\beta^2}\right)\right)$$
$$u' = u\cos\theta + v\sin\theta,$$
$$v' = v\cos\theta - u\sin\theta.$$
(1)

where f is the dominant frequency of the sinusoidal plane wave, α is the sharpness of the Gaussian along the major axis parallel to the wave, θ is the anticlockwise rotation of the Gaussian and the envelope wave, and β is the sharpness of the Gaussian minor axis perpendicular to the wave. $\gamma = f/\alpha$ and $\eta = f/\beta$ are used to keep frequency and sharpness ratio in a constant state [17].

Fig. 4(b) and Fig. 4(c) show the magnitude and phase of the Gabor wavelets for 1 scale and 8 angels, respectively. The wavelet at all levels is a Gaussian bandpass filter. Gabor wavelets have various features and properties that could be used in different ways and applications. One of the most distinctive and important features is directional selectivity. With this feature, one can orient Gabor wavelets in any desired direction. Fig. 4(d) and Fig. 4(e) show the transformation results after applying the magnitude and phase of Gabor wavelets on the face image, respectively.

The 2D-DWT of a signal is performed by repeating the 2D analysis filter bank on the lowpass sub image. Here, in the processing of each scale, four subimages are used

instead of one. There are three wavelets which are associated with the 2D wavelet transform [18]. Repetition of the filtering and decimation process on low-pass outputs made multiple levels (scales). a 1-scale transformation of input face images using SIFT and SURF. Keypoint detection and description are performed on the output subband images using SIFT and SIFT defined as (DWT-SIFT, DWT-SURF).



Fig. 4. (a) The original image, (b) The magnitude and (c) the phase of the Gabor kernels at 1 scale and 8 angles, (d) The magnitude and (e) phase results of convolved face image with Gabor kernels.

In Fig. 5 DWT transformation applied on a face image, outputs four different subband images, namely; approximate, horizontal, vertical and diagonal.



Fig. 5. The box filters of approximations of Gaussian second order partial derivative.

III. METHODS

The proposed approach contains details of the stages taken in carrying out the simulations. All the images are transformed using DWT or GWT. We proposed two approaches using SIFT and SURF.

In first approach, SURF or SIFT was used as a feature extraction algorithm, but before extracting features, input face images are transformed using DWT. DWT generates four different subband images namely: approximate, vertical, horizontal and diagonal. Fig. 6 and Fig. 7 shows



Fig.6. The block diagram of proposed approach for DWT-SIFT.



Fig.7. The block diagram of proposed approach for DWT-SURF.

All keypoint features that are extracted from SURF will be stored. Then, each corresponding feature of keypoints will be compared using kNN to get a score (that defines the number of matched keypoints). Then, summation of scores are stored. At last decision, will be made based on the highest score, which will define if a subject belongs to a class or not. In 2-scales transformation, after applying 1-scale transformation, DWT was applied as a second scale on approximate subimage, which produces four subband images. Scores of all eight subband images will be fused and decision will be made based on results.

In second approach, SURF or SIFT was used as a feature extraction algorithm, but before extracting features input face images were transformed using GWT. GWT outputs eight different subband images in each scale. Fig. 8 shows 1-scale transformation of input

images, and features are extracted from output subband images using SURF or SIFT defined as (DWT-SURF, DWT-SIFT).



Fig.8. The block diagram of proposed approach for GWT-SIFT and GWT-SURF.

All keypoint features that are extracted from SURF or SIFT will be stored. Then, each corresponding feature of keypoints will be compared using kNN to get a score (that defines the number of matched keypoints). Then, summation of scores is stored. At last, decision will be made based on the highest score, which will define if a subject belongs to a class or not.

IV. RESULTS

The proposed approach is tested on two different face databases: ORL [18] and PUT [19] face databases. The Olivetti Research Laboratory (ORL) face database is tested to assess our proposed approach in the existence of head poses and variations in time since images were taken between April 1992 and April 1994. There are 40 different subjects (persons), 10 images per subject, a total of 400 face images. For most of the subjects, the face images were recorded at light variance, time variance, face details (glasses / no glasses), face expressions (open / closed eyes, smiling / not smiling) and head poses (rotation and tilting up to 20 °). Most of the face images were recorded against a dark regular background. The Poznan University of Technology (PUT) face database is used to test partially controlled illumination conditions over a homogenous background and they have different head pose variations and in some images subjects wear glasses. There are 100 distinct subjects, 10 images per each subject, a total of 1000 face images. For most of the subjects, the images were recorded with different face expressions, illumination and head poses. The database supplies additional information including: rectangles containing face, nose, eyes, and mouth.

For the experiments on each dataset, 5 randomly chosen face images are considered as the gallery (train) set and the remaining 5 face images are considered as the probe (test) set. There was no overlapping between images from gallery and probe sets. Subjects in both databases have 10 face images with different conditions such as different illumination, pose, expression, etc. Each of the images in probe set is compared against the images in the gallery set, and the results and scores are fused before the final decision is made.

The proposed approaches were tested against conventional algorithms. We applied 10-fold scenario in our experiments where, the program where ran for 10 times with randomly selected gallery and probe sets. Score of matching SIFT and SURF descriptors from different subbands were fused before making a final decision at each experiment.

Proposed approaches were applied two wavelets transforms namely DWT and GWT. With DWT, different types of filters were applied (db1, db2, db3, db4, db5, haar...) before picking the one with the best performance for the rest of experiments. GWT subband images are complex-valued. SIFT and SURF were applied on the magnitude and phase parts of subband images of GWT.

A. Experiments conducted on ORL database

At first stage, SURF was applied and the average result of recognition rate was 90.09% using a different number of individuals that varies between 5 to 40 subjects. Performance and recognition rate of all algorithms that have been tested in these experiments were decreased when the number of subjects increased. Using different transformation filters, 1-scale transformation was applied on face images. The performance of the proposed approach was not good enough because, after the transformation of face images, SURF didn't extract enough features to describe them. The approximate subband images had zero SURF points for any image in the gallery or probe sets. So, the remaining subband images (horizontal, vertical and diagonal) of DWT were used in the experiments.

Overall performance average of SIFT using ORL database was 83.15%. The performance of the proposed approach using 1-scale of DWT-SIFT lead to ~7% difference from SIFT algorithm. And the average difference between 1-scale and 2-scales of SIFT-DWT was ~2% using ORL face images. While applying a different number of subjects on the proposed approach, GWT-SURF outperforms and shows better results compared to DWT-SURF and SURF. The average difference in performance rate between GWT-SURF and SURF was ~4%.

The performance results of our proposed approaches were different compared to the previous results when GWT was used instead of DWT as a transformation step on the face images. In the 1-scale transformation, GWT outputs subband images in a complex form, which makes SURF nonfunctional in extracting features; however, our proposed approaches performed well using Magnitude, Phase, and both Magnitude and Phase (combined).

The performance of the proposed approaches using Magnitude and Phase of transformed images was ~5% higher than the conventional SURF algorithm. When the number of subjects increasing, the recognition performance of our proposed approaches decreases less compared to SURF algorithm. The performance of the proposed approaches was not as expected with DWT while using GWT produces higher recognition rate

that's ~12% higher than conventional SIFT algorithm.

The overall average recognition rate of SIFT was 83.15% however, the average recognition rate of SIFT after applying GWT (as a transformation layer) was 95.06%. Results for SIFT, SURF and proposed approaches using ORL database are tabulated in Table 1. and Table 2.

Table 1. The recognition rate of SURF and proposed approaches using ORL database.

# of Subjects	SURF	GWT _{Mag} - SURF	GWT _{phase} -SURF	GWT _(Mag+Phase) - SURF
10	93.00	90.20	92.60	94.20
15	90.53	91.73	93.60	95.20
20	89.30	90.60	92.50	93.70
25	88.72	91.28	92.88	94.16
30	89.13	88.73	92.33	93.33
35	89.54	89.31	92.00	93.60
40	88.50	89.70	91.60	93.40

Table 2. The recognition rate of SIFT and proposed approaches using ORL.

# of	SIFT	GWT _{Mag} -	GWT _{phase} -	GWT(Mag+Phase)-
Subjects		SIFT	SIFT	SIFT
10	79.20	90.20	14.60	96.00
15	83.06	91.73	6.93	96.27
20	85.30	90.60	5.80	94.20
25	82.40	91.28	4.72	94.08
30	83.93	88.73	4.40	94.27
35	84.91	89.31	3.89	94.63
40	85.20	89.70	3.30	94.60

It was observed that the phase of complex GWT subband images didn't work properly with SIFT however, magnitude and combination of magnitude and phase are quite more accurate compared to SIFT algorithm. The overall recognition rate using a different number of subjects for SIFT, SURF, and proposed approaches are shown in Fig. 9 and 10 respectively.



Fig.9. Overall recognition performance of SIFT, DWT-SIFT (1-scale), DWT-SIFT (2-scales), and GWT-SIFT on ORL database.



Fig.10. Overall recognition performance of SURF, DWT- SURF (1scale), DWT- SURF (2-scales), and GWT- SURF on ORL database.

B. Experiments conducted on PUT database

In this experiment 5 to 100 different subjects from the PUT database were tested, the resulting average recognition rate was 84.84% using SURF. The recognition performance of SURF algorithm decreases when the number of subjects increases. The same experiments and steps that were performed on ORL database, was performed on PUT database.

In 1-scale transformation, using different transformation filters, the performance of the algorithm was not good because, after transformation, SURF was not able to extract distinct features to describe face images. For cA (Approximate) subimages, there were no SURF points for all images in gallery and probe sets that were tested. Therefore, we only used cH (Horizontal), cV (Vertical) and cD (Diagonal) subimages of DWT.

In 2-scale transformation, the same filters were used as 1-scale transformation. For DWT transformation results, we can conclude that there is ~1% difference in recognition rate between 1-scale and 2-scale transformations. SIFT algorithm's performance rate varies when size of subject's decreases. For example, using 5 subjects results in the recognition rate of 98.80% while, 100 subjects result in 93.90%. The overall average recognition rate of SIFT using PUT database was 96.12%. Results for SIFT, SURF and proposed approaches using PUT database are tabulated in Table 3. and Table 4.

Table 3. The recognition rate of SURF and proposed approaches using PUT database.

# of	SURF	GWT _{Mag} -	GWT _{phase}	GWT(Mag+Phase)-
Subjects		SURF	-SURF	SURF
10	88.40	98.80	98.20	98.40
20	88.80	99.30	98.10	99.00
30	89.13	99.53	98.60	99.27
40	86.05	98.90	98.60	99.20
60	81.38	99.16	98.62	99.13
80	79.76	99.20	98.83	99.36
90	78.75	99.15	98.73	99.29
100	76.57	98.82	98.63	99.18

26

# of Subjects	SIFT	GWT _{Mag} - SIFT	GWT _{phase} - SIFT	GWT _(Mag+Phase) - SIFT
10	98.40	99.80	17.00	100.00
20	98.20	99.10	12.60	99.00
30	96.73	99.27	9.00	99.20
40	96.60	99.25	7.80	99.15
60	94.40	99.27	6.15	99.16
80	94.28	99.39	4.69	99.36
90	94.38	99.22	4.05	99.25
100	93.90	99.07	3.81	99.09

Table 4. The recognition rate of SIFT and proposed approaches using PUT.

The performance of our proposed approaches was completely different when GWT was applied as a transformation on the face images before extracting distinct features from them. In 1-scale transformation was used, GWT outputs vectors in complex. SURF was not functional with complex type, so Magnitude, Phase, and combination of both were used in our proposed approaches

The recognition performance of the proposed approaches using Phase of transformed images was ~3% higher than the conventional SURF algorithm. The recognition rate of our proposed approaches decreases slowly compared to SURF algorithm. The overall recognition performance rate using a different number of subjects for SURF and proposed approaches are shown in Fig. 11.



Fig.11. Overall recognition performance of SURF, DWT- SURF (1scale), DWT- SURF (2-scales), and GWT- SURF on PUT database.

The performance of the proposed approaches was not as expected using DWT, however with GWT there was ~13% difference in the recognition rate compared to conventional SIFT algorithm. The overall average recognition rate of SIFT was 83.15%. On the other hand, the overall recognition rate of SIFT after applying GWT as a transformation step was 95.06%.

We observed that using GWT algorithm with the combination of Magnitude and Phase were more accurate compared to SIFT algorithm. The overall recognition performance rate for different number of subjects for SIFT and proposed approaches can be observed in Fig. 12.



Fig.12. Overall recognition performance of SIFT, DWT-SIFT (1-scale), DWT-SIFT (2-scales), and GWT-SIFT on PUT database.

V. CONCLUSIONS

In this paper, SURF and SIFT were used to extract features from face images. Two approaches based on wavelet transforms were proposed to improve the recognition performances of SIFT and SURF descriptors. The first approach is based on DWT namely; DWT-SURF and DWT-SIFT. The second approach is based on GWT namely; GWT-SURF and GWT-SIFT. The DWT or GWT is applied to the image as a preprocessing stage before conventional SURF or SIFT is applied. SURF or SIFT is applied on the obtained subband images separately. The recorded scores from each subband image are then fused together to get the final score and make a more accurate decision. The results obtained show that the fusion of matching scores of SURF or SIFT descriptors on the multiresolution images substantially improved the recognition performance.

REFERENCES

- R. Gottumukkal, V. K. Asari, "An improved face recognition technique based on modular PCA approach," *Pattern Recognition Letters*, vol. 25(4), pp. 429-436, 2013.
- [2] Z. Liu, Z. Xu, L. Hu, "A comparative study of distance metrics used in face recognition," *Journal of Theoretical* & *Applied Information Technology*, vol. 5(2), pp. 324-332, 2009.
- [3] M. Turk, A. Pentland, "Eigenface for recognition," *Journal of Cognitive Neuroscience*, vol. 3(1), pp. 71-86, 1991.
- [4] M. Pietik änen, A. Hadid, G. Zhao, T. Ahonen, "Local binary patterns for still images," *Computer Vision Using Local Binary Patterns*, pp. 13-47, 2011.
- [5] K. Baek, B. A. Draper, J. R. Beveridge, K. She, "PCA vs. ICA: A comparison on the FERET data set," *Journal of Computer Information System*, vol. 15(3), pp. 824-827, 2002.
- [6] M. Abdul Muqeet, R. S. Holambe, "Enhancing face recognition performance using triplet half band wavelet filter bank", *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, vol. 8(12), pp. 62-70, 2016.

- [7] N. A. Singh, M. B. Kumar, M. C. Bala, "Face recognition system based on SURF and LDA technique", *International Journal of Intelligent Systems and Applications (IJISA)*, vol. 8(2), pp. 13-19, 2016.
- [8] B. A. Draper, K. Baek, M. S. Bartlett, J. R. Beveridge, "Recognizing faces with PCA and ICA," *Computer Vision and Image Understanding*, vol. 91(1), pp. 115-137, 2003.
 [9] A. Eleyan, H. Demirel, "PCA and LDA based neutral
- [9] A. Eleyan, H. Demirel, "PCA and LDA based neutral networks for human face recognition," *Face Recogniton*, pp. 94-106, Rijeka, Croatia, In Tech, 2007.
- [10] R. Steinhoff, B. Wongsaroj, C. Smith, A. Gallon, J. Franscis, "Improving eigenface face recognition by using image registration preprocessing methods," Dept. of Computer Science and Mathematics, Florida Memorial University, Florida, USA.
- [11] A. P. Ginsburg, "Visual information processing based on spatial filters constrained by biological data," *Air Force Aerospace Medical Research Lab*, vol. 1/2, pp. 78-129, 1978.
- [12] L. D. Harmon, "The recognition of faces," Scientific American, vol. 229(2), pp. 71-82, 1973.
- [13] J. Sergent, H. D. Ellis, M. A. Jeeves, F. Newcombe, A. Young, "Microgenesis of face perception," *Aspects of Face Processing*, pp. 17-33, 1986.
- [14] D. G. Lowe, "Distinctive image features from scaleinvariant keypoints," *International Journal of Computer Vision*, vol. 60(2), pp. 1-28, 2004.
- [15] H. Bay, T. Tuytelaars, L. Gool, "SURF: Speeded Up Robust Features," *Proc. 9th European Conference on Computer Vision*, pp. 404-417, May 2006.
- [16] B. K. Bairagi, A. Chatterjee, S. C. Das, B. Tudu, "Expressions invariant face recognition using SURF and gabor features," *IEEE 3rd International Conference in Emerging Applications of Information Technology*, pp. 170-173, 2012.
- [17] A. Eleyan, H. Özkaramanli, H. Demirel, "Complex wavelet transform-based face recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2008(1), pp. 1-13, 2008.

- [18] F. Samaria, A. Harter, "Parameterization of a stochastic model for human face identification," *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, pp. 138-142, December 1994.
- [19] A. Kasiński, A. Florek, A. Schmidt, "The PUT face database," *Image Processing & Communications*, vol. 13(3-4), pp. 59-64, 2008.
- [20] K. Ramesha, K. B. Raja, "Dual transform based feature extraction for face recognition," *IJCSI International Journal of Computer Science*, vol. 8(2), pp. 115-124, 2011.

Authors' Profiles



Musa M.Ameen received the B.Sc. degree in Computer Engineering from Ishik University, Iraq in 2013, and M.Sc. degree in Computer Engineering from Mevlana University, Turkey in 2016. Currently, he works as Lecturer at Ishik University, Iraq. His current research interests are biometrics, computer crime

vision, and signal processing.



Alaa Eleyan received the B.Sc. and M.Sc. degrees in Electrical & Electronics Engineering from Near East University, Northern Cyprus, in 2002 and 2004, respectively. In 2009, He finished his PhD degree in Electrical and Electronics Engineering from Eastern Mediterranean University,

Northern Cyprus. Currently, he works as an Assist. Professor at Avrasya University, Turkey. His current research interests are biometrics, computer vision and image processing. He has more than 50 published journal and conference papers in the computer vision and image processing research fields.

How to cite this paper: Musa M.Ameen, Alaa Eleyan," Score Fusion of SIFT & SURF Descriptors for Face Recognition Using Wavelet Transforms", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.9, No.10, pp. 22-28, 2017.DOI: 10.5815/ijigsp.2017.10.03