

# Face Recognition Based on Improved SIFT Algorithm

\*EHSAN SADEGHIPOUR

\*Department of Electrical and Computer Engineering  
Hormozgan University  
Bandar Abbas, Iran

NASROLLAH SAHRAGARD

Department of Electrical and Computer Engineering  
Hormozgan University  
Bandar Abbas, Iran

**Abstract**—People are usually identified by their faces. Developments in the past few decades have enabled human to automatically do the identification process. Now, face recognition process employs the advanced statistical science and matching methods. Improvements and innovations in face recognition technology during 10 to 15 past years have propelled it to the current status. Due to the wide application of face recognition algorithms in many practical systems, including security control systems, human-computer interaction systems, etc., algorithms with high success rate are highly interested in research areas in recent years. Most of suggested algorithms are about correctly identifying face photos and assigning them to a person in the database. This study focuses on face recognition based on improved SIFT algorithm. Results indicate the superiority of the proposed algorithm over the SIFT. To evaluate the proposed algorithm, it is applied on ORL database and then compared to other face detection algorithms including Gabor, GPCA, GLDA, LBP, GLDP, KGWRCM, and SIFT. The results obtained from various tests show that the proposed algorithm reveals accuracy of 98.75% and run time of 4.3 seconds which is shorter. The new improved algorithm is more efficient and more accurate than other algorithms.

**Keywords**—face detection; improved SIFT descriptor; KGWRCM; GPCA; GLDA

## I. INTRODUCTION

Faces play the major role in identifying people and showing the feelings in the image. The human ability to recognize faces is remarkable. This ability stands against changes in visual conditions such as facial expressions, age, as well as changes in glasses, beard or hair style. Face recognition is a key subject in applications such as security systems, credit card control, culprit identification, mugshot matching (photos taken from the face), surveillance and control of access, personal identification, as well as applications for forensic pathology and constitution enforcement. In this case, a database containing facial photos for a group of people is employed.

Then, based on unrecognized photo, we want to answer the question: "does the photo belongs to a person in the database?" Many algorithms are proposed to solve this problem [2]. In this work, a new method is proposed based on SIFT [1]. Then, it is compared to Eigen-faces and Fisher-faces algorithms [3],[4],[5]. The results obtained by the proposed algorithm show improvements compared to other algorithms.

## II. LITERATURE

Eigen-faces algorithm [6] is based on reducing the size of Principal Component Analysis (PCA). The basic idea is that each image is considered as a vector in a high-dimensional space. As the input image, the target image is first transformed using Eigen vectors matrix or vectors comprising the subspace. Then, it is compared with existing data in the dimension-reduced space and the most similar image is selected as the identified one. First, the data covariance matrix and then the vectors matrix and eigenvalues are calculated. Eigen vectors matrix is composed of orthogonal vectors forming the feature subspace. Data will be independent by transforming to this subspace. An input image can be displayed as a linear combination of Eigen-faces by transforming to the new Eigen-faces space. After transforming the input image to the new Eigen-faces space, the most similar image is selected as the identified image using the nearest neighbor. To cluster the approach, two types of nearest neighbor are considered. The first input image is compared with all images in the database. Then, the nearest cluster center is obtained. Calculations are done using each cluster (the face of a person) and the cluster is selected for comparison.

Fisher-faces approach [3],[4] is based on the expansion of linear resolution analysis (LDA) in face recognition. The objective is generating a space with minimum intra-class dispersion and maximum inter-class dispersion, finding vectors providing the best classes for separation, as well as trying to maximize the difference between the classes and minimize the number of classes. Here, the mechanism is similar to that of PCA, i.e. an image is considered as point in a high-dimensional space. Then, LDA is applied on new basic vectors. This method is called Fisher-faces. Accordingly, the images are used for matching process. Again, the two nearest neighbors can be employed for the discussed process.

## III. SIFT ALGORITHM

Scale Invariant Feature Transform (SIFT) is an algorithm employed in machine vision to extract specific features of images for applications such as matching various view of an object or scene (for binocular vision) and identifying objects [6]. The obtained features are invariant to scale and rotation, and partially invariant to change in lighting. Such features are extracted in four steps.

In first step, using a DOG filters, all points in the image at all scales are searched to find features invariant to the orientation and scale. In the next step, an extended model is used for any potential point to determine the appropriate points based on different sustainability criteria. In third step, based on the local gradient of the image at target points, one orientation is assigned to each feature point. Finally, the information contained in the gradient function around the feature points is encoded to be used as the point's characteristic for future works such as feature matching. The name of descriptor is taken from its mechanism, i.e. converting the image data algorithms into scale-invariant coordinates in accordance with local features. Each feature is considered as a vector in a 128-dimensional space identified over the key point neighborhood.

#### IV. THE PROPOSED ALGORITHM

Image matching is a basic process in image processing and plays an important role in photogrammetry and remote sensing processes. SIFT is an efficient matching algorithm which has successfully been used to match the remote sensing images automatically. However, feature extraction operator is still a major weakness of the algorithm. Despite the high computational complexity, it can be partially controlled in feature extraction from remote sensing images. To alleviate the computational complexity while improving the efficiency, this study limit the feature extraction operators in SIFT as well as key-points extraction operators to a certain number to develop an efficient way to match images. The main feature of this method is taking the advantages of a specific strategy to select the best key points. After extracting the key points and making their descriptors, matching process starts considering the Euclidean distance between the descriptors as a multilevel process using a one-to-many procedure as done. Then, using least squares model, the accuracy of the initial matching-pairs are checked to find the compatible matching. Finally, the two images are matched by determining the parameters of a segment transformation function. Practical results obtained by applying the method on different face images indicate the great efficiency of the proposed method in comparison with the standard SIFT algorithm. In the proposed method, a consistent feature is considered to be compared with another feature for which the distance is less than a certain deductions (in this project 0.79) of a distance to the next nearest feature. This will reduce the number of incorrect matches. For incorrect matches, there will be a number of other near features due to the high-dimensional existing features. On the other hand, if there is a correct match, it is unlikely to find other features due to the distinctive nature of the SIFT descriptor.

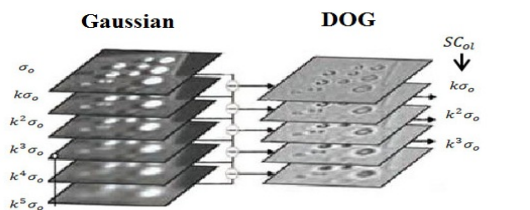


Fig. 1. The structure space is composed of a set of Gaussian images and DOG (Difference of Gaussian)

In key point extraction, the first step is to create a scale space. The structure space is a display of image structures at different scales. The structure space is composed of a set of Gaussian images and DOG (Difference of Gaussian) stored in a structure called the octave as Figure 1 shows.

To create the Gaussian images in scale space for any octave, the image must be convolved with the kernel Gaussian with varying kernel. Therefore, the Gaussian image  $L(x, y, \sigma)$  is created by convolving the image  $I(x, y)$  with Gaussian kernel  $G(x, y, \sigma)$  through equation (1):

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y) \quad (1)$$

Where  $\otimes$  denotes the convolution operator. In Eq. 1,  $\sigma$  represents the scale with the initial value of 1.6 repeated increasingly using  $K$  at different levels of the octaves. Then, DoG images are calculated based on the difference between the two adjacent Gaussian images in the scale space (scale of the smaller Gaussian image is considered as the scale of the DoG image). The octave Gaussian image with the value twice the initial scale is selected, the dimension is halved by resampling, and it is considered as the primary image for the next octave. In order to establish a stable condition, the space scale of each pixel in the intermediate levels of DoG images in each octave is compared with the 8 neighboring pixels for the target DoG image and 9 neighboring pixels for each upper and lower DoG images. If it is an extreme (maximum or minimum), it will be selected as the key point. The number of scale levels in each octave (i.e. the number of searched to extract key points) is a parameter affecting the number of extracted key points is. According to Lowe, the number of Gaussian image in each octave and the number of DoG images is considered 6 and 5, respectively. In the proposed algorithm,  $\sigma$  is calculated as follows Eq. (2):

$$\sigma = k^{(k_1 + 2 * k_2)} * 1.6 \quad (2)$$

This change makes the number of Gaussian images per octave 4 instead of 6, thus decreasing the SIFT run time. Besides, the proposed method considers the number of key points fixed as 50 points for each image with a rating higher than others. This ultimately increases the speed of the algorithm. The number of octaves in the scale space also varies depending on the image size which is considered 3 after testing different numbers.

Scale factor of each scale levels is calculated through Eq. (3) as:

$$SC_{0l} = \sigma \cdot 2^{(o-1 + \frac{l}{LN})} = \sigma \cdot K^{LN(o-1) + l} \quad (3)$$

$$o = 1, 2, \dots, ON; l = 1, 2, \dots, LN; k = 2^{1/LN}$$

Where  $SC_{0l}$  is the scale factor for level  $l$  of octave  $o$ . Finally, using the Euclidean distance, key points between two images are selected.

V. COMPARING ALGORITHMS IN TERMS OF ACCURACY AND SPEED

ORL is considered as the input data for this study including 400 face images belonging to 40 different people taken from different angles as well as light intensity. For the first test, 100 images were randomly selected and 1 image of each person was considered as the training image (three persons were considered for training). As an example of the algorithm output, we assumed the three images in Fig. 2 as training images. Then, a randomly selected image from the three target persons was tested. A function removing all unnecessary part of the image such as background was called. Then, employing improved SIFT algorithms, key points in the training images using sift were extracted.

corresponding key points in training images, the Euclidean distance between the points was calculated and finally, the image with smaller total value was selected as the final image. Repeating the test on 100 randomly selected image resulted in 86 correct answer was. The second test was conducted to compare the proposed method with others. Here, 80% of all data was considered as training data and the remaining 20% as test data (i.e. 320 images for training and 80 ones for test). After applying Gabor, GPCA and GLDA, LBP, GLBP, KGWRCM [7] and SIFT on ORL data set, it was concluded that that the accuracy of the proposed algorithm is 6.25% more than SIFT with the runtime 1.9 seconds less than the original SIFT algorithm.

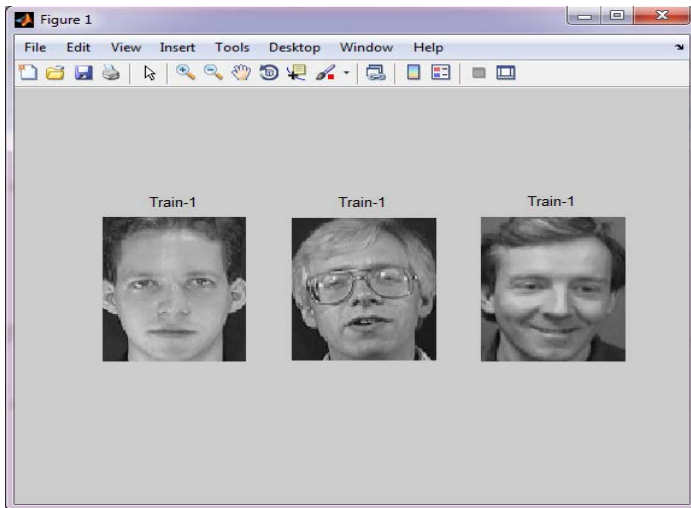


Fig. 2. As an example of the algorithm output

In the next step, an image from the three training persons was tested. Operations done on training images was repeated for the test image as shown in Figure 3. Fig. 3 shows the main test image, the test image after removing unnecessary areas, and the extracted key points.

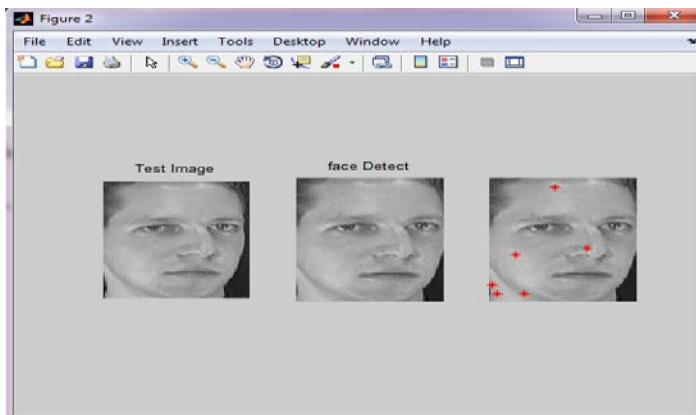


Fig. 3. Shows the main test image, the test image after removing unnecessary areas, and the extracted key points

Finally, the key points calculated for the test image were compared with the key points calculated for the training images. The comparison was so that after finding the

TABLE I. ACCURACY OF RECOGNITION ON ORL DATABASE

Algorithm	Accuracy
1)Gabor	81.34
2)GPCA	89.87
3)GLDA	97.50
4)LBP	89.61
5)GLBP	93.42
6)KGWRCM	99.32
7)SIFT	92.5
8)OUR METHOD	98.75

The graph comparing the proposed method with other methods is shown in Figure 4.

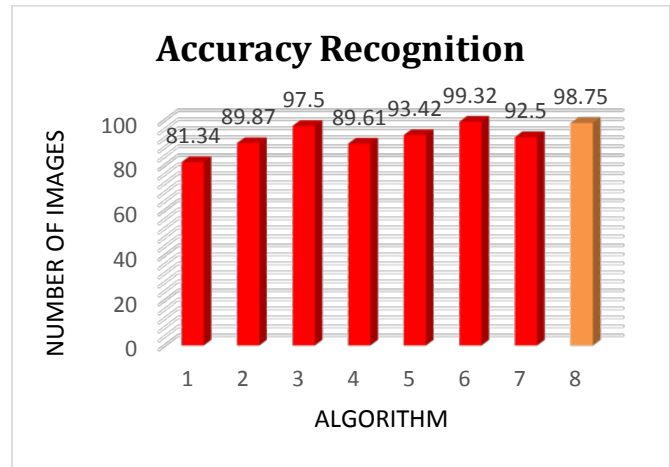


Fig. 4. Comparing the proposed method with other methods

The next comparison was conducted based on the run time indicating the significant advantage of the proposed algorithm compared to others shown in Table 2.

TABLE II. PROCESSING TIME FOR 300 IMAGES (IN SECONDS)

Algorithm	Accuracy
1)Gabor	1.8
2)GPCA	1.5
3)GLDA	7.2
4)LBP	6.9
5)GLBP	3.7
6)KGWRCM	4.9
7)SIFT	5.3
8)OUR METHOD	5.3

The graph comparing the run time of the proposed method with other methods is shown in Figure 5.

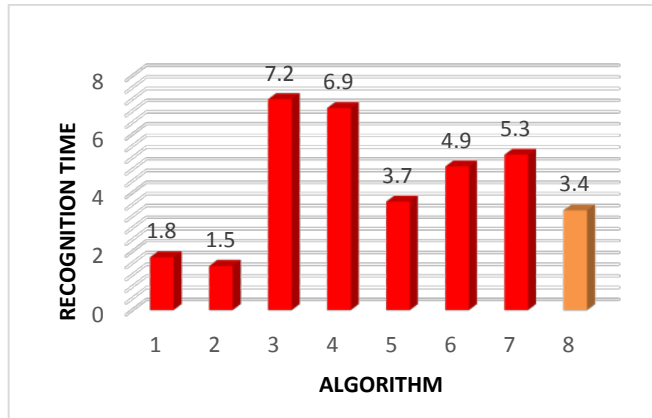


Fig. 5. The run time of the proposed method with other methods

## VI. CONCLUSIONS

People are usually identified by their faces. Developments in the past few decades has enabled human to automatically do the identification process. Early face recognition algorithm used simple geometric models. Now, face recognition process employs the advanced statistical science and matching methods. Improvements and innovations in face recognition technology during 10 to 15 past years have propelled it to the current status. Face recognition is applicable for both investigation and identification. This study provided a new face recognition method based on SIFTS features. Improving

SIFT algorithm, this study focused on face recognition. Results indicated the superiority of the proposed algorithm over the SIFT. To evaluate the proposed algorithm, it was applied on ORL database and then compared to other face detection algorithms including Gabor, GPCA, GLDA, LBP, GLDP, KGWRCM, and SIFT. The results obtained from various tests showed that the proposed algorithm, with 98.75% accuracy and run time of 4.3 seconds, is more efficient and accurate than other algorithms yet the run time was less than others.

## REFERENCES

- [1] David G. Lowe(2004), Distinctive image features from scale-invariant key points, International journal of computer vision 60.
- [2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld(2003). Face recognition: A literature survey. ACM Computing Surveys, 399-458 ,December.
- [3] N. Belhumeur Peter, Joao P. Hespanha, David J. Kriegman(1996).Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions On Pattern Analysis And Machine Intelligence,711-720, July.
- [4] Daniel L. Swets, John Weng (1996), Using discriminant eigen features for image retrieval, IEEE Transactions On Pattern Analysis And Machine Intelligence, 18(8):831-836, August.
- [5] Matthew A. Turk, Alex P. Pentland (1991).Face recognition using eigenfaces. In Proceedings CVPR 91., IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
- [6] Fisher A (1936), The use of multiple measures in taxonomic problems. Ann. Eugenics, 7:197-188.
- [7] Wei Xia, Shouyi Yin and Peng Ouyang (2013)A High Precision Feature Based on LBP and Gabor Theory forFace Recognition"*Sensors* 2013, 13, 4499-4513; doi:10.3390/s130404499,ISSN 1424-8220.