Face Recognition Based on similarity Feature-Based Selection and Classification Algorithms and Wrapper Model

Chi-Kien Tran

***Abstract*—Similarity feature-based selection and classification (SFSC) algorithms, introduced by Tran *et al*. in 2013, have been used as a tool to reduce storage cost and increase performance of face recognition systems. However, these still exist a problem when automatically selecting a suitable threshold. This paper introduces a new approach, which combines SFSC algorithms, and a wrapper model, to automatically select a suitable threshold and improve face recognition accuracy. The training face image set (which is split into two separated subsets including a training subset and a wrapper subset) is utilized as data input for the similarity feature-based selection algorithm in combination with the wrapper model to identify a best feature set. The obtained feature set will be used for classification. The experiments were conducted on the histogram-based feature and two databases, ORL database of faces and Georgia Tech face database. The results demonstrated that the proposed algorithm not only allowed for automatic detection of the suitable feature set, but also achieved a better recognition accuracy compared to conventional algorithms.**

***Index Terms*—Face recognition, similarity feature, feature selection, filter model, wrapper model.**

I. Introduction

Feature selection, as an essential task in a face recognition system, could be considered the next step after the feature extraction process [1]. A good dimensionality reduction method can decrease the dimension of feature space, increase recognition accuracy, while maintaining the lowest level of classification errors.

A feature selection method selects the best subset of the input feature set that properly describes the given problem with a minimum reduction in performance. Feature selection methods broadly fall into three models: filter, wrapper, and embedded [2]-[4]. The filter model evaluates features without involving any learning algorithm. The wrapper model requires a learning algorithm and uses its performance to evaluate the goodness of features. The embedded model incorporates feature selection as part of the learning process, and uses the objective function of the learning model to guide the process of searching for relevant features such as decision trees or artificial neural networks.

In a recent study, two feature selection and classification algorithms based on a filter model that named similarity feature-based selection and classification algorithms (SFSC) have been initially proposed by Tran *et al*. [5]. The goal of the algorithms is to retain similarity features of the training images in a class in order to minimize within-class differences, while maximizing between-class differences and to use this feature set for classification. They have been proven an efficient tool for improving the performance of face recognition systems using local binary patterns (LBP), local ternary patterns, and local directional pattern (LDP) features [6], [7]. However, SFSC algorithms still have a limitation as the value of threshold parameter is not automatically set, meaning that user needs to test many different values of threshold to find the best similarity feature set. To overcome this limitation, we propose a novel approach based on wrapper model, WSFSC, to find the optimal similarity feature set. Firstly, a face image set is divided into three subsets: training images, wrapper images, and testing images. Secondly, similarity feature-based selection algorithm (SFS) is conducted on two subsets (training images, wrapper images) to find the optimal similarity feature set. Finally, the best similarity feature set is used for classification. The experiments on the ORL database of faces (ORL) [8] and Georgia Tech face database (GTFD) [9], [10] showed that the proposed method was effective for performance improvement of face recognition system.

The remaining part of the paper is organized as follows. Section II describes the similarity feature-based selection and classification algorithm. Section III presents SFSC-Wrapper model. The experimental results on the face databases and discussion are presented in Sections IV. Finally, Section V draws the conclusion remarks.

II. Similarity feature-Based Selection and Classification Algorithms

The similarity feature-based selection and classification algorithms (SFSC) is an effective algorithm to decrease the dimensions of feature space and improve face recognition rates [5]-[7]. For the similarity feature-based selection algorithm (SFS), its fundamental aim is to retain the similarity features of the training images in a class to minimize within-class differences. First, the variance of features is computed. Next, the obtained values of previous step are normalized. Finally, the features greater than threshold ε (threshold value is set by user) are removed. For the similarity

Manuscript received November 20, 2019; revised April 19, 2019.
Chi-Kien Tran is with the Faculty of Information Technology, Hanoi University of Industry, No. 298, Cau Dien Street, Bac Tu Liem district, Hanoi, Vietnam (e-mail: chikien.tran@haui.edu.vn).

doi: 10.18178/ijmlc.2019.9.3.810

feature-based classification (SFC), the classification is implemented based on a nearest neighbor classifier (NN). An input image is classified by calculating the average distance of the feature pairs which have the same coordinates and the value of similarity feature is different from -1. Details of these algorithms can be found in [5]-[7].

Here, a face image is first encoded by a local descriptor. Then a histogram-feature vector of the encoded image is constructed. Let f_{mn} be a feature matrix of a subject (m individuals, n features). Each row is a histogram-feature vector, which depicts a face image. In order to retain the similarity features, a threshold value ε is set by a user. Algorithms 1 is used to select similarity features. Its details are described as follows.

Algorithm 1 Similarity feature-based selection (SFS)
Input: The feature matrix f_{mn} of a class, the threshold value ε.
Output: The similarity feature matrix *X*.
Procedure:
For $j = 1$ To n Step 1
  Compute the mean of the *j*th feature $\psi_j = \frac{1}{m}\sum_{i=1}^{m} f_{ij}$;
  Compute the variance of the *j*th feature $V_j = \frac{1}{m-1}\sum_{i=1}^{m} (f_{ij} - \psi_j)^2$;
End For
Find the maximum variance value $\lambda = \max\{V\}$;
For *i* = 1 To *m* Step 1
  For j = 1 To *n* Step 1
    If $V_j / \lambda \le \varepsilon$ Then
      $X_{ij} = f_{ij}$;
    Else
      X_{ij} = -1;
    End If
  End For
End For
Return *X*;

Algorithm 2 is a similarity feature-based classification algorithm. It calculates the distance between the feature vector of a testing image Y and the similarity feature vectors produced from Algorithm 1. The calculation is based on the distance of the feature pairs which have the same coordinates and the value of similarity feature is different from -1. The result is divided by the number of similarity features in the similarity feature vector. The measurement, *L*, can be similarity measures such as Manhattan distance, Euclidean distance, and Chi-square statistics.

Algorithm 2 Similarity feature-based classification (SFC)
Input: The feature vector of a testing image *Y*, the similarity feature matrix M_{mn} of *m* training images of classes, and the *L* measurement.
Output: A label of class that is nearest to *Y*.
Procedure:
For $k = 1$ to m
Compute the distance between *Y* and M_{kj}: $d_k(Y, M_{kj}) = \frac{1}{q^k} L(Y, M_{kj})$,
where q^k is the number of similarity features of the *k*th training image.
End For
Find the label which corresponds to the minimum distance between *Y* and M_{kj}:

$$s = \text{argmin}_k(d_k);$$

Return *s*;

III. Similarity Feature-Based Selection Algorithm Using Wrapper Model

In this section, a new feature selection algorithm, which is called wrapper model based similarity feature selection (WSFS), is introduced. It is a combination of the SFS algorithm and a wrapper model to identify a suitable training feature set. In this algorithm, (1) the training image set is divided into two subsets: the training set and the wrapper set; (2) these sets are extracted features; (3) the SFS and SFC algorithms using these two sets are implemented with different threshold values. Each threshold value will produce a training feature set. The training feature set, which achieves the highest accuracy, will be chosen as the output of the algorithm.

In the proposed algorithm (Algorithm 3), *min_ε* is the minimum threshold value, *max_ε* is the maximum threshold value, *step_value* is a step value of the loop and it also is utilized to change the value of the threshold. The value of these variables are set up by the user. The highest value of *max_ε* should be 1. The value of *step_value* should be smaller than 0.1.

The structure of the SFSC-Wrapper model is illustrated in Fig. 1 and Algorithm 3 is described as follows.

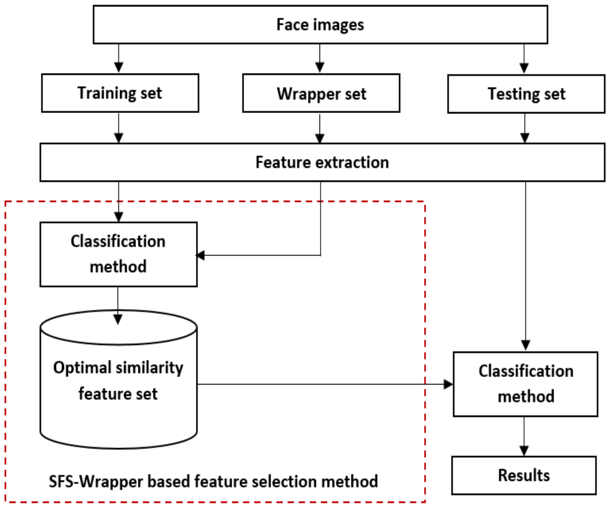

Fig. 1. The structure of the proposed method.

Algorithm 3 Wrapper model based similarity feature selection (WSFS)
Input: The training feature set T_{mn}, the wrapper feature set W_{mn}, the minimum threshold value *min_ε*, and the maximum threshold value *max_ε*.
Output: The selected feature matrix *S*.
Procedure:
Initialize *T_opt* = NULL; *max_rate* = 0;
For *ε* = *min_ε* To *max_ε* Step *step_value*
  Select the similarity feature set which corresponds to threshold *ε*:
    SF = SFS(*T*, *ε*);
  Compute the accuracy: *r* = SFC(*SF*, *W*);
  If *max_rate* < *r* Then
    max_rate = *r*;
    T_opt = *SF*;
  End If
End For
Return *T_opt*;

IV. Experimental Results and Discussion

A. Data Input

In the experiment, we choose two well-known databases to test our proposed algorithm. Each has its different emphasis, as shown in Fig. 2 and Fig. 3. The ORL Database of Faces (ORL) [8] consists of 400 images of 40 different persons. The images mainly vary in pose and scale. The size of each image is 112 × 92 pixels (see Fig. 2). The Georgia Tech face database (GTFD) [9] with the background removed is made up of 15 face images per 50 subjects. Each face is characterized by different facial expressions, illumination, rotation and size. In this study, the images were converted to gray-scale, resized to 132 × 102 pixels (see Fig. 3).

Fig. 2. Example images of two persons in the ORL database of faces.

Fig. 3. Example of images of four persons in the GTFD face database.

B. Experimental Settings

In order to compare the efficiency of proposed and conventional methods, LBP [11] descriptor was used to represent the face image and histogram-based feature was extracted from obtained images. The chi-square distance [6], [11], [12] was chosen for nearest neighbor classifier. The conventional method (CM) is a face recognition method using nearest neighbor classifier. The proposed method uses the same classification approach. Face image sets were randomly divided into three separated subsets (training subset, wrapper subset and testing subset). Each experimental plan uses X training images, Y wrapper images, and Z testing images. CM uses two subsets (training subset and testing subset) for experimental plans. The value of threshold varies from the maximum threshold value to the minimum threshold value. For the proposed method, the range of threshold is from 1 to 0.2, each step with a value of –0.05 for all the experiments. Each image of the ORL and GTFD databases was divided into 5 × 5 and 4 × 3 blocks, respectively. The normal algorithm uses the training subset for training and testing subset for testing.

The accuracy of our method is calculated as a percentage of correct classifications, which is computed as follows:

$$\text{Accuracy}(\%) = \frac{\#\text{of correct classification}}{\#\text{of total testing images}} \times 100. \quad (1)$$

C. Experimental Results on the Georgia Tech Face Database

The experimental results conducted on the GTFD database are displayed in Table I-Table IV. These tables are split into groups. Each group (three columns) is an experimental plan. In these groups, the first column lists the results of the conventional method (CM). The second column demonstrates the results of the proposed method (WSFSC). The third column shows the value of corresponding thresholds used in the proposed method. The improvement results are displayed in bold and the poor results are shown in (.).

The average accuracies are listed in three final rows (S1, S2, and S3) of the table. S1 is the average accuracies of conventional and proposed methods. S2 shows the average accuracy of positive tests, which achieve accuracy improvement and S3 is the average accuracy of negative tests, which reduce accuracy.

TABLE I: RECOGNITION ACCURACY CONDUCTED THE GTFD DATABASE WITH 3 TRAINING IMAGES

No	3-5-7			3-6-6			3-7-5			3-8-4		
	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε
1	57.71	57.71	1.00	50.66	**51.00**	0.25	55.60	**57.20**	0.30	51.50	51.50	0.35
2	58.28	**58.85**	0.25	48.00	**48.33**	0.55	52.80	**53.20**	0.20	52.00	52.00	0.85
3	52.00	52.00	0.80	52.00	**52.33**	0.25	54.40	**55.20**	0.35	52.50	**53.50**	0.25
4	50.00	**52.85**	0.20	45.66	**46.00**	0.50	57.20	(56.40)	0.45	46.50	**47.00**	0.50
5	56.57	**56.85**	0.20	52.33	**53.00**	0.65	54.00	54.00	0.20	54.50	**55.50**	0.55
6	51.71	**52.57**	0.40	53.00	**53.33**	0.85	52.40	52.40	0.30	52.50	**53.00**	0.85
7	53.71	**54.00**	0.25	58.33	**58.66**	0.95	58.00	58.00	0.95	59.00	**59.50**	0.90
8	55.71	55.71	1.00	52.66	**53.00**	0.95	59.60	**61.20**	0.20	58.50	(58.00)	0.65
9	50.28	**51.14**	0.45	53.66	53.66	1.00	58.00	**58.80**	0.95	52.50	**53.50**	0.55
10	52.57	**54.00**	0.25	52.66	52.66	1.00	52.40	52.40	0.40	53.00	53.00	1.00
S1	**53.85**	**54.57**	||	**51.90**	**52.20**	||	**55.44**	**55.88**	||	**53.25**	**53.65**	||
S2	||	**0.64**	||	||	**0.37**	||	||	**1.04**	||	||	**0.75**	||
S3	||	**(0.00)**	||	||	**(0.00)**	||	||	**(0.80)**	||	||	**(0.50)**	||

X-Y-Z: means X training images, Y wrapper images, and Z testing images. CM: means conventional method. WSFSC means proposed algorithms. ε: means the value of threshold.

TABLE II: RECOGNITION ACCURACY CONDUCTED THE GTFD DATABASE WITH 4 TRAINING IMAGES

No	4-4-7			4-5-6			4-6-5			4-7-4		
	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε
1	61.14	(60.00)	0.30	60.66	60.66	1.00	54.40	**56.00**	0.40	65.50	**66.50**	0.95
2	60.00	**61.14**	0.45	62.00	62.00	1.00	57.20	**57.60**	0.20	57.00	**58.50**	0.20
3	62.00	**62.57**	0.45	66.66	66.66	1.00	54.80	**55.60**	0.20	60.00	(59.50)	0.45
4	63.42	**63.71**	0.90	61.33	61.33	1.00	61.60	**62.00**	0.75	58.00	**58.50**	0.40
5	60.57	60.57	1.00	62.66	62.66	1.00	56.40	56.40	1.00	55.50	**57.00**	0.30
6	59.14	**59.71**	0.45	62.33	62.33	1.00	58.40	**59.20**	0.35	59.50	**60.00**	0.20
7	62.28	**62.57**	0.95	56.66	56.66	1.00	58.80	58.80	1.00	54.50	54.50	0.85
8	58.57	**60.57**	0.25	62.66	62.66	1.00	52.80	**53.60**	0.95	58.50	(58.00)	0.35
9	63.42	**64.28**	0.95	58.66	58.66	1.00	58.40	(58.00)	0.35	61.00	**63.50**	0.45
10	56.85	56.85	0.30	61.00	61.00	1.00	61.60	61.60	1.00	51.00	**52.50**	0.25
S1	**60.74**	**61.20**	||	**61.46**	**61.46**	||	**57.44**	**57.88**	||	**58.05**	**58.85**	||
S2	||	**0.81**	||	||	**0.00**	||	||	**0.80**	||	||	**1.28**	||
S3	||	**(1.14)**	||	||	**(0.00)**	||	||	**(0.40)**	||	||	**(0.50)**	||

TABLE III: RECOGNITION ACCURACY CONDUCTED THE GTFD DATABASE WITH 5 TRAINING IMAGES

No	5-4-6			5-5-5			5-6-4			5-7-3		
	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε
1	61.66	**62.00**	0.60	60.80	**61.20**	0.30	64.00	64.00	1.00	66.00	**69.33**	0.25
2	62.00	**62.33**	0.25	66.00	66.00	1.00	67.50	**68.50**	0.80	63.33	**64.00**	0.45
3	63.33	63.33	0.75	64.80	64.80	0.70	68.00	68.00	1.00	70.66	**73.33**	0.50
4	68.00	(67.66)	0.40	66.40	**66.80**	0.35	69.50	**70.50**	0.45	60.66	60.66	0.30
5	61.66	61.66	0.30	66.00	66.00	0.65	60.50	**61.50**	0.90	67.33	(66.00)	0.75
6	63.66	**64.33**	0.30	66.80	66.80	1.00	64.50	64.50	1.00	67.33	67.33	1.00
7	65.66	65.66	0.95	64.00	64.00	1.00	67.50	67.50	1.00	70.00	70.00	0.20
8	60.33	(59.66)	0.25	68.40	(68.00)	0.35	67.50	**68.00**	0.40	66.00	**66.66**	0.95
9	61.00	**61.33**	0.40	64.40	**64.80**	0.25	66.00	**66.50**	0.95	70.66	70.66	1.00
10	66.66	(66.33)	0.35	68.00	**68.40**	0.20	64.50	**65.50**	0.20	70.00	70.00	0.95
S1	**63.40**	**63.43**	||	**65.56**	**65.68**	||	**65.95**	**66.45**	||	**67.20**	**67.80**	||
S2	||	**0.41**	||	||	**0.40**	||	||	**0.83**	||	||	**1.83**	||
S3	||	**(0.44)**	||	||	**(0.40)**	||	||	**(0.00)**	||	||	**(1.33)**	||

Table I presents a comparison of the recognition accuracies between two methods using three training images. As can be seen from this table, for the first group (3-5-7), the proposed method produces seven tests with higher accuracies (No. 2, 4 ÷ 7, 9, and 10). Its improved average rate is 0.64%. For the second group (3-6-6), the proposed algorithm obtains eight higher results (No. 1 ÷ 8) than that of conventional algorithm. The improved average rate is 0.37%. For the third group (3-7-5), there are five instances (No. 1 ÷ 3, 8, and 9) in which the proposed algorithm achieve higher results and one instance with a poorer result than that of the conventional algorithm (No. 4). The improved average rate is 1.04%, while the reduced average rate is 0.8%. For the fourth group (3-8-4), there are six positive results (No. 3 ÷ 7, and 9) and one negative result (No. 8). The improved average rate is 0.75%, while the unimproved average rate stands at 0.5%.

TABLE IV: RECOGNITION ACCURACY CONDUCTED THE GTFD DATABASE WITH 6 TRAINING IMAGES

No	6-3-6			6-4-5			6-5-4			6-6-3		
	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε
1	71.00	(70.33)	0.80	68.00	**68.40**	0.65	67.00	**68.00**	0.30	70.00	**73.33**	0.20
2	68.00	**68.33**	0.90	66.00	**66.80**	0.35	65.50	65.50	1.00	62.66	62.66	0.25
3	64.33	64.33	1.00	68.40	68.40	1.00	70.50	**72.50**	0.50	68.00	**68.66**	0.65
4	67.66	**68.33**	0.20	71.20	71.20	0.95	67.00	67.00	1.00	72.66	72.66	0.60
5	69.00	69.00	1.00	65.20	(64.00)	0.45	68.50	68.50	0.75	61.33	61.33	0.40
6	71.00	**72.00**	0.30	63.60	**64.00**	0.55	70.00	**71.00**	0.95	66.66	(66.00)	0.95
7	66.00	66.00	0.95	66.40	**67.20**	0.30	70.50	**73.00**	0.30	69.33	69.33	0.55
8	68.33	68.33	0.95	62.00	**62.40**	0.40	69.00	69.00	0.95	72.66	**73.33**	0.55
9	68.66	(68.33)	0.95	64.80	**66.40**	0.50	70.00	**70.50**	0.55	66.66	66.66	0.85
10	69.66	69.66	1.00	70.00	70.00	1.00	68.00	68.00	1.00	56.00	56.00	0.60
S1	**68.36**	**68.46**	||	**66.76**	**66.88**	||	**68.60**	**69.30**	||	**66.60**	**67.00**	||
S2	||	**0.66**	||	||	**0.73**	||	||	**1.40**	||	||	**1.55**	||
S3	||	**(0.50)**	||	||	**(1.20)**	||	||	**(0.00)**	||	||	**(0.66)**	||

TABLE V: RECOGNITION ACCURACY CONDUCTED ON THE ORL DATABASE WITH 2 TRAINING IMAGES

No	2-2-6			2-3-5			2-4-4		
	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε
1	88.75	**90.41**	0.20	87.00	87.00	1.00	88.75	**90.00**	0.2
2	83.33	**83.75**	0.35	87.00	87.00	1.00	82.50	**83.12**	0.35
3	85.00	(84.16)	0.20	84.00	**84.50**	0.40	87.50	(86.87)	0.45
4	82.08	**82.50**	0.85	82.00	**82.50**	0.85	80.62	**81.25**	0.85
5	86.66	86.66	1.00	90.00	**90.50**	0.20	88.12	**88.75**	0.45
6	90.83	90.83	0.45	91.00	91.00	0.45	91.25	91.25	0.45
7	85.41	85.41	1.00	85.50	85.50	1.00	86.87	86.87	1
8	85.41	85.41	1.00	86.50	86.50	1.00	87.50	(86.50)	0.4
9	82.50	82.50	0.95	81.50	81.50	0.95	83.12	83.12	0.95
10	83.75	**84.58**	0.25	86.50	86.50	0.95	84.37	**85.62**	0.25
S1	**85.37**	**85.62**	||	**86.10**	**86.25**	||	**86.06**	**86.33**	||
S2	||	**0.83**	||	||	**0.50**	||	||	**0.87**	||
S3	||	**(0.83)**	||	||	**(0.00)**	||	||	**(0.81)**	||

Table II lists the results of the two methods on four training images. For the first plan (4-4-7), the proposed method produces seven tests with higher accuracies (No. 2 ÷4 and 6 ÷ 9) and one test with a poorer result (No. 1). The improved average rate is 0.81%, while the reduced average rate is 1.14%. For the second plan (4-5-6), there is no improvement from the proposed method. For the third plan (4-6-5), there are six instances (No. 1 ÷4, 6, and 8) in which the proposed algorithm produced higher results and one instance with poorer result than that of the conventional algorithm (No. 9). The improved average rate is 0.8%, while the reduced average rate is 0.4%. For the fourth plan (4-7-4), there are seven positive results (No. 1, 2, 4 ÷6, 9, and 10) and two negative results (No. 3 and 8). The improved average rate is 1.28%, while the unimproved average rate is 0.5%.

Table III displays the results of two methods using five training images. For the first plan (5-4-6), the proposed method has four improved tests (No. 1, 2, 6, and 9) and two reduced tests (No.4 and 10). The improved average rate is 0.41% and the unimproved average rate is 0.44%. For the second plan (5-5-5), the proposed algorithm obtained four good results (No. 1, 4, 9, and 10). The improved average rate is equal to the unimproved average rate (0.4%). For the third plan (5-6-4), there are six instances (No. 2, 4, 5, and 8 ÷10) in which the proposed algorithm produces higher results. The improved average rate is 0.83%. For the fourth plan (5-7-3), there are four positive results (No. 1 ÷ 3, and 8) and one negative result (No. 5). The improved average rate is 1.83%, while the unimproved average rate is 1.33%.

Table IV presents the results of the recognition accuracies between two methods using six training images. As can be seen from this table, the proposed method produces three tests with higher accuracies (No. 2, 4, and 6) with an improved average rate of 0.66% and two negative results (No. 1 and 9) with an reduced average rate of 0.5% (see group 1). For the second group (6-4-5), the proposed algorithm obtains six higher results (No. 1, 2, and 6 ÷9) and one lower result (No. 5) than that of the conventional algorithm. The improved average rate and unimproved average rate of proposed method are 0.73% and 1.2%, respectively. For the third group (6-5-4), there are five instances (No. 1, 3, 6, 7, and 9) in which the proposed algorithm produces higher results than that of the conventional algorithm. Its improved average rate is 1.4%. For the fourth group (6-6-3), there are three positive results (No. 1, 3, and 8) and one negative result (No. 6). The improved average rate is 1.55%, while the reduced average rate is 0.66%.

From the comparison of average results of two methods (see Table I-Table IV), it can be seen that the proposed method obtains higher recognition accuracies compared with the conventional method, except in the 4-5-6 case (see Table II).

D. Experimental Results on the ORL database of Faces

Tables V-VII list the experimental results conducted on the ORL database. These tables have the same structure as the tables in Section E. Table V displays the results of two methods using two training images. For the first plan (2-2-6), the proposed method has four improved tests (No. 1, 2, 4, and 10) and one poor test. The improved average rate is equal to the unimproved average rate (0.83%). For the second plan (2-3-5), the proposed algorithm obtains three good results (No. 3 ÷5) with an improved average rate of 0.5%. For the third plan (2-4-4), there are five positive results (No. 1, 2, 4, 5, and 10) and two negative results (No. 3 and 8). The improved average rate is 0.87%, while the unimproved average rate is 0.81%.

Table VI displays the results of two methods using three training images. For the first plan (3-2-5), the proposed method produces one improved case (No. 10) with an improved average rate of 0.5%. For the second plan (3-3-4), the proposed algorithm obtains one good result (No. 8). The improved average rate is 1.25%. For the third plan (3-4-3), the proposed method produces no improvement.

Table VII illustrates the results of two methods using four training images. For the first plan (4-2-4), the proposed method produces two improved tests (No. 3 and 5) and two reduced tests (No. 2 and 10) with an improved average rate of 0.62% and an unimproved average rate of 0.62%. For the second plan (4-3-3), the proposed algorithm obtains one good result (No. 2) and one poor result. The improved average rate is less than the unimproved average rate (0.83% and 1.66%, respectively). For the third plan (4-4-2), there is one case (No. 3), in which the proposed algorithm achieves higher result. The improved average rate is 1.25%.

Comparing the average results obtained from two methods (see Tables V-VII), we can see that the proposed method achieves six higher results, two unchanged results (see 3-4-3, Table VI and 4-2-4, Table VII), and one poorer result (see 4-3-3, Table VII) compared with the conventional method.

E. Discussion

As mentioned in Section B, the threshold value ranges from 1 to 0.2 with a step value of -0.05. With a threshold value of 1, the similarity feature set is also the original feature set. The selection of the starting threshold value of 1 ensures that the result of the original feature set is included in the comparison to select the final feature set. Moreover, when multiple equal-value positive results are produced, a result with high threshold value is awarded higher priority. This is because the closer the threshold value approaches (or reaches) 1, fewer (or no) features are omitted, thereby ensures a certain level of robustness for the final feature set.

TABLE VI: RECOGNITION ACCURACY CONDUCTED ON THE ORL DATABASE WITH 3 TRAINING IMAGES

No	3-2-5			3-3-4			3-4-3		
	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε
1	93.50	93.50	0.55	90.00	90.00	1.00	90.00	90.00	1.00
2	91.50	91.50	1.00	88.12	88.12	1.00	91.66	91.66	0.95
3	91.50	91.50	1.00	92.50	92.50	1.00	92.50	92.50	0.35
4	86.00	86.00	1.00	86.87	86.87	1.00	87.50	87.50	0.85
5	94.50	94.50	0.20	95.62	95.62	1.00	93.33	93.33	1.00
6	94.50	94.50	1.00	93.12	93.12	1.00	87.50	87.50	0.65
7	89.50	89.50	1.00	88.75	88.75	1.00	89.16	89.16	0.95
8	89.50	89.50	1.00	93.12	**94.37**	0.35	99.16	99.16	0.35
9	89.00	89.00	1.00	96.87	96.87	0.65	95.00	95.00	1.00
10	90.50	**91.00**	0.35	93.75	93.75	0.90	94.16	94.16	1.00
S1	**91.00**	**91.05**	**||**	**91.87**	**92.00**	**||**	**92.00**	**92.00**	**||**
S2	**||**	**0.50**	**||**	**||**	**1.25**	**||**	**||**	**0.00**	**||**
S3	**||**	**(0.00)**	**||**	**||**	**(0.00)**	**||**	**||**	**(0.00)**	**||**

TABLE VII: RECOGNITION ACCURACY CONDUCTED ON THE ORL DATABASE WITH 4 TRAINING IMAGES

No	4-2-4			4-3-3			4-4-2		
	CM	WSFSC	ε	CM	WSFSC	ε	CM	WSFSC	ε
1	96.25	96.25	1.00	95.83	95.83	1.00	97.50	97.50	0.80
2	93.75	(93.12)	0.45	95.00	**95.83**	0.85	93.75	93.75	1.00
3	95.00	**95.62**	0.25	96.66	96.66	0.25	95.00	**96.25**	0.30
4	90.62	90.62	1.00	90.00	(88.33)	0.20	90.00	90.00	0.30
5	96.25	**96.87**	0.95	95.83	95.83	1.00	97.50	97.50	0.40
6	97.50	97.50	1.00	98.33	98.33	1.00	98.75	98.75	0.80
7	94.37	94.37	1.00	97.50	97.50	0.60	95.00	95.00	1.00
8	93.12	93.12	1.00	94.16	94.16	1.00	93.75	93.75	0.50
9	95.00	95.00	1.00	97.50	97.50	0.60	92.50	92.50	1.00
10	97.50	(96.87)	0.55	95.83	95.83	0.90	98.75	98.75	0.35
S1	**94.93**	**94.93**	**||**	**95.66**	**95.58**	**||**	**95.25**	**95.37**	**||**
S2	**||**	**(0.62)**	**||**	**||**	**0.83**	**||**	**||**	**1.25**	**||**
S3	**||**	**(0.62)**	**||**	**||**	**(1.66)**	**||**	**||**	**(0.00)**	**||**

A comparison of the results across the tables I ÷ VII reveals that for each empirical plan, the number of instances in which the WSFC algorithm (the proposed algorithm) improves the accuracy is higher than the number of instances in which it reduces accuracy. Therefore, the average accuracy it produces is higher than that of the conventional algorithm. It easily follows from that postulate that a better result is normally achieved if the number of wrapper images is greater than the number of test images. This is especially true if a large number of wrapper images are similar to images in the testing set.

Despite the apparent improvement in accuracy achieved by adopting the proposed algorithm, a number of limitations should be considered:

- When the number of wrapper images is too low, too high or considerably different from the number of images in the testing set, the obtained feature set is unsatisfactory (the overfitting phenomenon). This issue causes the output similarity feature set produced by the algorithm to be robust only for the wrapper set while unsatisfactory when the testing set is dissimilar to the wrapper set.
- As the range of threshold values changes (e.g., from 0.6 to 0.2 instead of from 1 to 0.2), the similarity feature set produced by algorithm will be affected. This causes the recognition accuracy to be changed accordingly.

By means of conclusion, the proposed algorithm allows for the reduction of the feature dimension and improves the face recognition performance. The algorithm will produce better results if the images in the wrapper set are relatively similar to those in the testing set.

V. CONCLUSION

In this paper, we developed a novel algorithm based on the similarity feature-based selection, classification algorithms and the wrapper model. The input face images are divided into three subsets (training subset, wrapper subset, and testing subset). The first two subsets are used to find an optimal feature set, which helps to reduce the size of training feature set and improve recognition accuracy. The experiments on the ORL and GTFD databases indicated that the proposed method was effective for reducing storage cost and increasing performance of face recognition systems.

REFERENCES

[1] I. Marqués, "Face recognition algorithms," PhD thesis, Universidad del País Vasco, 2010.

[2] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *The Journal of Machine Learning Research*, vol. 3, pp. 1157-1182, 2003.

[3] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, "Feature extraction: Foundations and applications," *Studies in Fuzziness and Soft Computing*, Springer-Verlag New York, Inc., 2006.

[4] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artificial Intelligence*, vol. 97, pp. 245-271, 1997.

[5] C. K. Tran, T. F. Lee, C. C. Tuan, C. H. Lu, and P. J. Chao, "Improving face recognition performance using similarity feature-based selection and classification algorithm," in *Proc. 2013 Second International Conference on Robot, Vision and Signal Processing*, 2013, pp. 56-60.

[6] C. K. Tran, T. F. Lee, L. Chang, and P. J. Chao, "Face description with local binary patterns and local ternary patterns: Improving face recognition performance using similarity feature-based selection and classification algorithm," in *Proc. 2014 International Symposium on Computer, Consumer and Control (IS3C)*, pp. 520-524, 2014.

[7] C. K. Tran, T. F. Lee, and P. J. Chao, "Improving face recognition performance using similarity feature-based selection and classication algorithm," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 6, 2015.

[8] A. T. L. Cambridge. The Database of Faces [Online]. Available: http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[9] Georgia Tech Face Database [Online]. Available: http://www.anefian.com/research/face_reco.htm
[10] L. Chen, H. Man, and A. V. Nefian, "Face recognition based on multi-class mapping of Fisher scores," *Pattern Recogn.*, vol. 38, pp. 799-811, 2005.
[11] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Proc. ECCV*, 2004, vol. 3021, pp. 469-481.
[12] C. K. Tran, C. D. Tseng *et al*., "Local intensity area descriptor for facial recognition in ideal and noise conditions," *Journal of Electronic Imagin*g, vol. 26, pp. 023011-1-023011-10, 2017.

Chi-Kien Tran received the Ph.D.degree from the College of Electrical Engineering and Computer Science, National Kaohsiung University of Applied Sciences,Taiwan, in 2017. Currently, he is a lecturer of the Faculty of Information Technology, Hanoi University of Industry, Hanoi, Vietnam. His current research interests include computer vision, pattern recognition, image analysis, and soft computing.