

Markov Network-based Unified Classifier for Face Identification

Wonjun Hwang^{1,2}

Kyungshik Roh¹

Junmo Kim²

¹Advanced Media Lab., Samsung Advanced Institute of Technology, Gyeonggi, Korea

²Dept. of EE, Korea Advanced Institute of Science and Technology, Daejeon, Korea

{wj.hwang, kyung.roh}@samsung.com junmo@ee.kaist.ac.kr

Abstract

We propose a novel unifying framework using a Markov network to learn the relationship between multiple classifiers in face recognition. We assume that we have several complementary classifiers and assign observation nodes to the features of a query image and hidden nodes to the features of gallery images. We connect each hidden node to its corresponding observation node and to the hidden nodes of other neighboring classifiers. For each observation-hidden node pair, we collect a set of gallery candidates that are most similar to the observation instance, and the relationship between the hidden nodes is captured in terms of the similarity matrix between the collected gallery images. Posterior probabilities in the hidden nodes are computed by the belief-propagation algorithm. The novelty of the proposed framework is the method that takes into account the classifier dependency using the results of each neighboring classifier. We present extensive results on two different evaluation protocols, known and unknown image variation tests, using three different databases, which shows that the proposed framework always leads to good accuracy in face recognition.

1. Introduction

When we use a face image as a query, we can retrieve several desired face images from a large image database. We calculate many similarities of the query image and the gallery images in the database, and the retrieved gallery images are ranked by similar orders. It is a one-to-many identification problem [8] and has many applications such as searching similar face images in a database and face tagging in images and videos. However, the traditional face recognition algorithms [10][14] have been developed for one-to-one verification [8], in particular, a biometric task. Recent successful face recognition methods have attempted to merge several classifiers using multiple feature sets of different characteristics, as in component-based methods, which extract features from separate spatial re-

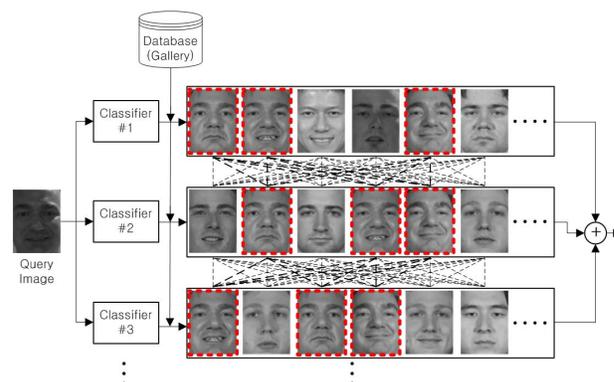


Figure 1. Top ranked gallery images of each independent classifier. The same persons are shown in red dashed boxes. True positive examples are mostly shown in the top ranks, but there are some false positives. The result of a classifier could be adjusted correctly using the results of the neighbor classifiers.

gions [14][6][5], and heterogeneous feature-based methods [21][9], which merge different domain features. These methods used the classifiers not only based on the different feature sets but also trained independently, and the similarity scores are merged with the predefined parameters. The parameter comes from the training database and it is not the best choice when the input image has different conditions. Note that these methods lead to good accuracy in face verification, but there is no specific framework for the one-to-many identification problem.

In this paper, we design a novel recognition framework for the one-to-many identification issue, and the simple concept is illustrated in Figure 1. First, we assume that we have multiple classifiers that have complementary characteristics. We can unify the multiple classifiers based not on the predefined weight values but on a Markov network, as summarized in Figure 2. For this purpose, we assign one node of a Markov network to each classifier. Nodes are connected by lines, which represents the statistical dependencies. For an observation node, we extract a feature from a query image using the corresponding classifier. At its paired hidden node, we retrieve the first n similar gallery

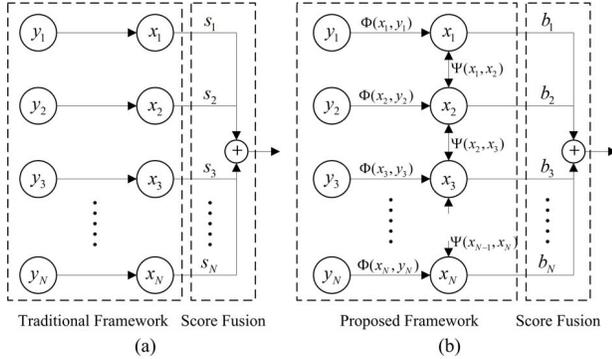


Figure 2. N multiple classifiers are deployed under (a) the traditional recognition framework and (b) the proposed recognition framework using a Markov network. The circles represent nodes, and the lines are the statistical dependency between the nodes. Observation node y is assigned by the feature of a query image and hidden node x is assigned by the feature of a gallery image. Φ and Ψ are the compatibility functions, which will be described in Section 4, that model the relationships between nodes. s and b are the score variables.

samples from the database, and their orders are made by the similarity scores for the query face image. The multiple classifiers have their own lists of retrieved gallery images, which are not identical in general, thereby complementing the neighbor classifiers. Because the hidden nodes are connected by the network lines, the relationship of the connected nodes is learned by the similarity scores between the neighbor classifiers, and the scores are calculated by concatenating the two gallery features of the neighbor classifiers. The posterior probability at each hidden node is easily computed by the belief-propagation algorithm. Finally, we have marginal probability for a score value at each classifier. We also analyze the generalizability of the proposed method using different multiple classifiers such as the Random Sampled Gabor (RSG) method, which consists of the simple and weak classifiers, and the Extended Curvature Gabor (ECG) method [4], which consists of more complex and stronger classifiers.

Our main contribution is the novel recognition framework that successfully organizes the classifier relationship using the similarity between the top ranked gallery images of the corresponding classifiers. Its own node characteristics are iteratively propagated to other neighbor nodes. As a result, all nodes are correlated to others, which improves the recognition results in a one-to-many identification task. Extensive experiments are also performed on cross-variations of well-known evaluation data sets to allow us to identify the performance changes due to the known and unknown image conditions changes.

The rest of this paper is organized as follows: Related works are introduced in Section 2. The proposed method is presented in Section 3 and 4, and the experimental results

are given in Section 5. We offer a conclusion in Section 6.

2. Related Work

The performances of recent face recognition algorithms have gradually advanced, which is largely due to techniques that merge several classifiers of different characteristics [21][20][9]. They have used many features such as Gabor, Local Binary Pattern (LBP), Fourier coefficient, and Discrete Cosine Transform (DCT) features. In this paper, we use Gabor feature-based classifiers as representative classifiers, but the proposed framework is applicable to any other classifiers.

When surveying the literature on face recognition using a Markov network, Huang et al. [3] proposed a hybrid face recognition method that combines holistic and feature analysis-based approaches using a Markov random field (MRF) model. They divided a face image into small patches, and the relationships between the patches were captured by the MRF model. The final recognition was based on majority voting using the results from the patches. On the other hand, Rodrigues et al. [17] estimated the eigenspace parameters of facial images with different poses using the MRF model over the parameter space to model a smoothness assumption. They reduced the image reconstruction error and improved the performance in pose estimation. Our method mainly differs in the purpose of using the Markov network. For example, in Huang's method, the Markov network was used to represent the relationship between the image patches and the patch IDs, and they adopted the delta-based function as the hidden-state compatibility function, which means that it will check whether the information of the neighbor blocks is same for the voting scheme for the verification task. On the other hand, we use the Markov network to apply the relationship between the multiple classifiers to the face recognition scheme and propose a similarity-based compatibility function between the neighbor classifiers for the identification task.

3. Review of the Markov Random Field

Freeman et al. [2] proposed a learning-based method for low-level vision problems known as the example-based super resolution. They generated a synthetic world of scenes and corresponding rendered images, and they modeled their relationship using a Markov network. However, in this paper, given a query image, y , we would like to identify the most similar gallery image, x , from the enrolled image set in a one-to-many face identification scenario. This problem can be cast as a maximization of the posterior probability, $P(x|y) = \frac{P(x,y)}{P(y)} = cP(x,y)$. Under the Markov assumption, the conditional probability and joint probability of N

nodes can be written as

$$P(\mathbf{x}|\mathbf{y}) = c \prod_{(i,j)} \psi_{ij}(\mathbf{x}_i, \mathbf{x}_j) \prod_i \phi_i(\mathbf{x}_i, \mathbf{y}_i), \quad (1)$$

where c is a normalization constant, ψ and ϕ are the compatibility functions, and the products are done through over all the neighboring pairs of nodes, i and j . We make use of the belief-propagation algorithm, which is known as the fast converging method.

The belief-propagation algorithm iteratively updates the message \mathbf{m}_{ij} from the node i to the node j , and the equation is as follows:

$$\mathbf{m}_{ij}(\mathbf{x}_j) = \sum_{\mathbf{x}_i} \psi_{ij}(\mathbf{x}_i, \mathbf{x}_j) \prod_{k \neq j} \mathbf{m}_{ki}(\mathbf{x}_i) \phi_i(\mathbf{x}_i, \mathbf{y}_i), \quad (2)$$

where $\mathbf{m}_{ij}(\mathbf{x}_j)$ is an element of the vector \mathbf{m}_{ij} corresponding to the gallery candidate \mathbf{x}_j . The marginal probability \mathbf{b}_i for gallery \mathbf{x}_i at node i is derived by

$$\mathbf{b}_i(\mathbf{x}_i) = \prod_k \mathbf{m}_{ki}(\mathbf{x}_i) \phi_i(\mathbf{x}_i, \mathbf{y}_i). \quad (3)$$

4. Proposed Unified Recognition Framework

4.1. Node

In [3], an image was divided into several patches and nodes were assigned to these patches under the Markov assumption, but in this paper, we assign the nodes to the feature vectors of the multiple classifiers. For example, the query feature and the gallery feature are assigned to the observation node, \mathbf{y} , and the hidden node, \mathbf{x} , respectively, and we design the parallel network without loops for simplicity, as shown in Figure 2 (b). In this respect, each classifier is influenced by two neighbor classifiers under the Markov assumption, and we can finally have N observation and hidden nodes pairs. A pair of the observation and the hidden nodes might be assumed as traditional face recognition, but the hidden nodes are connected by the network lines. At each hidden node, we collect n gallery candidates that are most similar to the observation feature. The hidden nodes have their own sets of gallery candidates retrieved from the database, and the n candidates of each node are not the same as the others because we assume that the multiple classifiers have different characteristics. For example, when assuming that there are two classifiers, we can have two sets of the n first ranked gallery candidates, and the relationship between hidden node x_1 and x_2 can be evaluated for pairs of realizations, $(\mathbf{x}_{1p}, \mathbf{x}_{2q})$, $1 \leq p \leq n, 1 \leq q \leq n$. In detail, x_{11} , the first ranked element of the gallery candidates, is compared with all the n candidates of the neighborhood classifier, $\{\mathbf{x}_{21}, \mathbf{x}_{22}, \dots, \mathbf{x}_{2n}\}$.

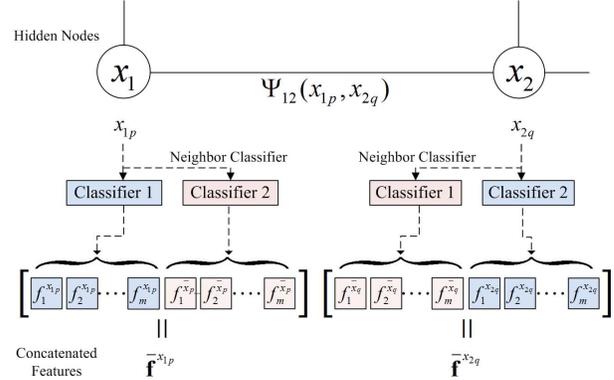


Figure 3. The similarity between instances \mathbf{x}_{1p} and \mathbf{x}_{2q} of different hidden nodes \mathbf{x}_1 and \mathbf{x}_2 is computed by comparing two concatenated features, $\tilde{\mathbf{f}}^{\mathbf{x}_{1p}} = [\mathbf{f}^{\mathbf{x}_{1p}} \mathbf{f}^{\mathbf{x}_{1p}}]$ and $\tilde{\mathbf{f}}^{\mathbf{x}_{2q}} = [\mathbf{f}^{\mathbf{x}_{2q}} \mathbf{f}^{\mathbf{x}_{2q}}]$, where \mathbf{x}_{1p} and \mathbf{x}_{2q} are from the retrieved gallery sets at the first and second hidden nodes, respectively. We generate the neighbor features, $\mathbf{f}^{\mathbf{x}_{1p}}$ and $\mathbf{f}^{\mathbf{x}_{2q}}$, using the neighbor classifier (red), and we add them to the main features, $\mathbf{f}^{\mathbf{x}_{1p}}$ and $\mathbf{f}^{\mathbf{x}_{2q}}$, (blue), in order. $f_k^{\mathbf{x}_{1p}}$ is the k th element of the feature vector, $\mathbf{f}^{\mathbf{x}_{1p}}$.

4.2. Compatibility Function

For a given observation, \mathbf{y}_i , the query-gallery compatibility function, $\Phi(\mathbf{x}_i, \mathbf{y}_i)$, is evaluated for n gallery candidates of \mathbf{x}_i , $\{\mathbf{x}_{i1}, \dots, \mathbf{x}_{in}\}$, thus generating a vector in $\mathbb{R}^{n \times 1}$. The gallery-gallery compatibility function, $\Psi(\mathbf{x}_i, \mathbf{x}_j)$, is evaluated for $n \times n$ pairs of $(\mathbf{x}_i, \mathbf{x}_j)$, where \mathbf{x}_i takes value on $\{\mathbf{x}_{i1}, \dots, \mathbf{x}_{in}\}$ and \mathbf{x}_j on $\{\mathbf{x}_{j1}, \dots, \mathbf{x}_{jn}\}$, thus generating a matrix in $\mathbb{R}^{n \times n}$.

Using the normalized correlation between two feature vectors as a measure of similarity, the n most similar gallery images are retrieved at each hidden node. We define the compatibility function between the hidden nodes i and j as,

$$\psi_{ij}(\mathbf{x}_i, \mathbf{x}_j) = \exp(-|s_{ij}(\mathbf{x}_i, \mathbf{x}_j) - 1|^2 / 2\sigma^2), \quad (4)$$

where σ is a noise parameter. We evaluate the compatibility function for each pair of the instances, that is, we compute $\psi_{ij}(\mathbf{x}_{ip}, \mathbf{x}_{jq})$ for $1 \leq p \leq n$ and $1 \leq q \leq n$. Now we need to define the similarity measure, $s_{ij}(\mathbf{x}_{ip}, \mathbf{x}_{jq})$. However, \mathbf{x}_{ip} and \mathbf{x}_{jq} are from different face classifiers, and we cannot directly compare their corresponding features, $\mathbf{f}^{\mathbf{x}_{ip}}$ and $\mathbf{f}^{\mathbf{x}_{jq}}$. To address this problem, we propose the concatenated features, $\tilde{\mathbf{f}}^{\mathbf{x}_{ip}}$ and $\tilde{\mathbf{f}}^{\mathbf{x}_{jq}}$, which are an augmented version of $\mathbf{f}^{\mathbf{x}_{ip}}$ and $\mathbf{f}^{\mathbf{x}_{jq}}$. Figure 3 illustrates how to construct the concatenated features. The similarity between i and j nodes is measured by the normalized correlation coefficient between the concatenated features:

$$s_{ij}(\mathbf{x}_{ip}, \mathbf{x}_{jq}) = \langle \tilde{\mathbf{f}}^{\mathbf{x}_{ip}}, \tilde{\mathbf{f}}^{\mathbf{x}_{jq}} \rangle / \|\tilde{\mathbf{f}}^{\mathbf{x}_{ip}}\| \|\tilde{\mathbf{f}}^{\mathbf{x}_{jq}}\|. \quad (5)$$

The query-gallery compatibility function and the similarity function are derived in a relatively straightforward way by

$$\phi_i(\mathbf{x}_i, \mathbf{y}_i) = \exp(-|s_i(\mathbf{x}_i, \mathbf{y}_i) - 1|^2 / 2\sigma^2), \quad (6)$$

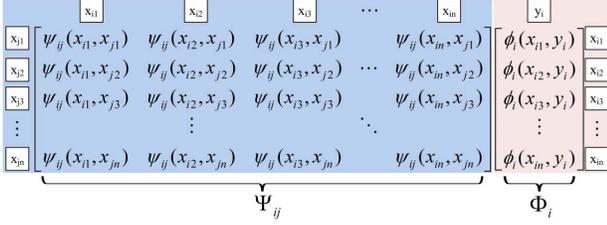


Figure 4. Simple concept of the message passing method. Given a query image at node i , n gallery images that are most similar to the query image, \mathbf{y}_i , are retrieved. The query-gallery compatibility function at node i is evaluated and represented by a column vector Φ_i . The gallery-gallery compatibility function between the gallery node i and the neighbor node j is computed for each pair of the candidates and is represented by a matrix, Ψ_{ij} .

$$s_i(\mathbf{x}_{ik}, \mathbf{y}_i) = \langle \mathbf{f}^{\mathbf{x}_{ik}}, \mathbf{f}^{\mathbf{y}_i} \rangle / \|\mathbf{f}^{\mathbf{x}_{ik}}\| \|\mathbf{f}^{\mathbf{y}_i}\|, \quad (7)$$

where $1 \leq k \leq n$.

4.3. Belief-Propagation

Belief propagation is a message passing algorithm. From Equations (2) and (3), we derive the following Equations for the proposed model of Figure 2 (b):

$$m_{(i-1)i}(\mathbf{x}_i) = \sum_{\mathbf{x}_{(i-1)}} \psi_{(i-1)i}(\mathbf{x}_{(i-1)}, \mathbf{x}_i) m_{(i-2)(i-1)}(\mathbf{x}_{(i-1)}) \phi_{(i-1)}(\mathbf{x}_{(i-1)}, \mathbf{y}_{(i-1)}), \quad (8)$$

where $\mathbf{x}_{(i-1)}$ in the summation runs through \mathbf{x}_i , that is, $\{\mathbf{x}_{(i-1)1}, \dots, \mathbf{x}_{(i-1)n}\}$. We evaluate Equation (8) for n values of \mathbf{x}_i , that is, $\{\mathbf{x}_{i1}, \dots, \mathbf{x}_{in}\}$.

$$m_{(i+1)i}(\mathbf{x}_i) = \sum_{\mathbf{x}_{(i+1)}} \psi_{(i+1)i}(\mathbf{x}_{(i+1)}, \mathbf{x}_i) m_{(i+2)(i+1)}(\mathbf{x}_{(i+1)}) \phi_{(i+1)}(\mathbf{x}_{(i+1)}, \mathbf{y}_{(i+1)}). \quad (9)$$

The marginal probability is

$$b_i(\mathbf{x}_i) = m_{(i-1)i}(\mathbf{x}_i) m_{(i+1)i}(\mathbf{x}_i) \phi_i(\mathbf{x}_i, \mathbf{y}_i). \quad (10)$$

As the posterior probability $p(\mathbf{x}_i | \mathbf{y}_1, \dots, \mathbf{y}_N)$ is proportional to $b_i(\mathbf{x}_i)$ at the hidden node \mathbf{x}_i , we have a chance to use these N marginal probabilities such as $b_1(\mathbf{x}_1), b_2(\mathbf{x}_2), \dots, b_N(\mathbf{x}_N)$. For simplicity, the final similarity score is the sum of the marginal probabilities such as $S = \sum_{i=1}^N b_i$.

For further discussion, let's illustrate Equation (8) using Figure 4, where the message that comes from the neighbor node i to the current node j corresponds to the column

Table 1. Known variation evaluation protocol using FRGC ver 2.0 database.

Protocol		Database	Individual	Image
Training		FRGC Training	222	12,766
Controlled vs	Query	FRGC Controlled	465	16,028
Controlled (CvC)	Gallery	As Above	465	16,028
Uncontrolled vs	Query	FRGC Uncontrolled	465	8,014
Controlled (UvC)	Gallery	FRGC Controlled	465	16,028

Table 2. Unknown variation evaluation protocol using FRGC ver 2.0, XM2VTS, and BANCA databases.

Protocol		Database	Individual	Image
Training		FRGC Training	222	12,766
XM2VTS vs	Query	XM2VTS	295	2,360
XM2VTS (XvX)	Gallery	As Above	295	2,360
BANCA vs	Query	BANCA	82	6,540
BANCA (BvB)	Gallery	As Above	82	6,540

vector obtained by multiplying the matrix (blue area) and the column vector (red area). Intuitively speaking, the column vector (red area) contains the information to be propagated and the matrix (blue area) plays the role of a channel over which the information is propagated. Each element in the matrix (blue area) is a measure of similarity between a gallery image and a gallery image of another node. If the two gallery images are similar enough to result in a large term in the matrix, the corresponding term in the column vector (red area) receives more emphasis. For example, suppose we would like to compare the query image and gallery image 1. If gallery image 1 is somehow similar to another gallery image, say, image 2, then the idea is that we will use the similarity between image 2 and the query to update the similarity between image 1 and the query. If image 1 and image 2 are two images of the same person taken in different situations, the achieved effect is a comparison between the query and a cluster of images, in this case, a cluster of image 1 and image 2. In this sense, the proposed inference algorithm performs a kind of clustering-based face recognition.

5. Experimental Result and Discussion

To evaluate the performances of the proposed algorithm, we make two different experimental protocols using the FRGC ver 2.0 [15], XM2VTS [12], and BANCA [1] databases. The first one is the known variation test protocol, as described in Table 1, where we can evaluate the performance changes of the algorithm due to the trained image variations such as the controlled and uncontrolled changes of FRGC. In this protocol, the multiple classifiers are trained by the FRGC training image set, and the test sets consist of the FRGC test images whose conditions are similar to that of the training images. There are two test sets: controlled image versus controlled image (CvC) and uncontrolled image versus controlled image (UvC). This is

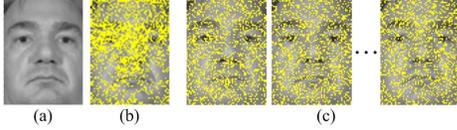


Figure 5. (a) Normalized face example. Yellow dots show the spatial localities of the 1,500 features selected by (b) the BGL method and (c) the samples of the proposed RSG classifiers.

the traditional evaluation protocol and is the same as the EXP1 and EXP4 of the original FRGC test protocol. In the other protocol, as described in Table 2, we train the multiple classifiers using only the FRGC training set, and the two test sets are composed of XM2VTS (XvX) and BANCA images (BvB), respectively, where we expect to observe performance changes according to the un-trained image variations. Note that three different database are collected not only for the different purposes but also under the different conditions. In the end, face recognition accuracy is calculated by the 1st rank of the Cumulative Match Characteristics (CMC) curve as a measurement of one-to-many face identification [8].

5.1. Gabor LDA-based Classifier

To evaluate the generalizability of the proposed recognition framework, we employ two different Gabor LDA-based classifiers such as the RSG classifiers, proposed in this paper, and the ECG classifiers [4], respectively.

In the RSG classifier, to build a single classifier, we first normalize an input image to a 60×80 resolution image based on the fixed eye distances (32 pixel), and then we apply a Gabor filter into a normalized image where Gabor parameters are from seven different frequencies, $\nu = \{-2, -1, \dots, 4\}$, eight orientations, $\mu = \{0, \dots, 7\}$, and $\sigma = \pi$. Now we have $7 \times 8 \times 60 \times 80$ Gabor features. For the feature selection, many works [23][19] have employed the boosting theory, but we randomly select 1,500 features, directly inspired by [22], and further reduce the dimension of the feature vector using LDA. Therefore, the final LDA output size is 221 due to the class number of the FRGC training set. Figure 5 illustrates the comparison of the spatial localities between the boosted Gabor LDA (BGL) and the random selected methods. Most features are selected around two eyes and a nose in the BGL method (Figure 5 (b)), but the selected features are well distributed by random sampling in the proposed method (Figure 5 (c)). In this paper, we build the ten RSG classifiers as the multiple classifiers ($N = 10$).

On the other hand, in the ECG classifier, the Gabor filter was extended to $7 \times 16 \times 60 \times 80$ features due to the addition of the curvature term. They selected 1,400 features from each different curved Gabor feature candidate using the AdaBoost method. They finally merged the multiple classifiers and achieved the best result among the Gabor-

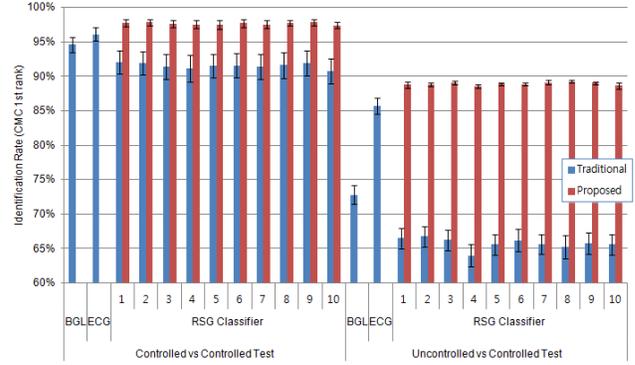


Figure 6. The average recognition rates of the single classifiers, for example, the BGL and ECG classifiers and ten RSG classifiers, are shown under the traditional and the proposed frameworks in the (a) CvC and (b) UvC tests. The bars are the standard deviations of three ROCs.

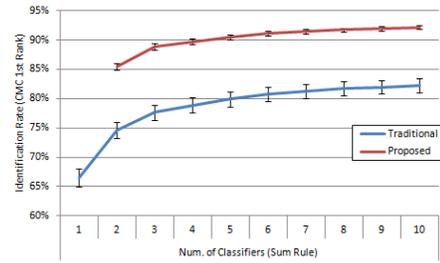


Figure 7. The average accuracy of combining different numbers of RSG classifiers under the traditional and proposed frameworks. The score fusion rule is the sum rule and the test protocol is the UvC test.

based recognition method. In this paper, we employ twelve ECG classifiers ($N = 12$) for the full performance without consideration of the computational complexity.

5.2. Experiment with the Known Variation Test

First, we evaluate the accuracy of each single RSG classifier. As shown in Figure 6, the average recognition rates of the RSG classifiers range from 63.22% (4th classifier) to 66.75% (2nd classifier), but the BGL classifier and the ECG classifiers achieve an average of 72.79% and 85.67%, respectively, in the UvC test. The proposed RSG classifier is not a good method from the viewpoint of a single classifier, but the benefit of the RSG method is its simplicity and that there is no limitation for generating multiple classifiers. Its characteristics are good for the proposed framework because we need more than one classifier. Note that all single RSG classifiers of the unified framework improve the average accuracy by approximately 12–13% with the UvC test compared to the original single RSG classifiers. Moreover, all single RSG classifiers of the proposed method show better accuracy than the BGL and ECG classifiers, and similar improvement is also found in the CvC test.

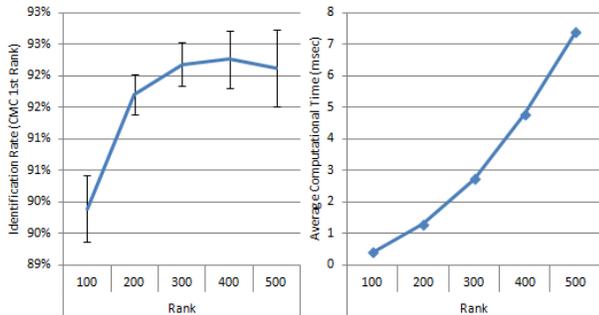


Figure 8. (a) The lines show the average recognition rates of the ten RSG classifiers as a function of the size of the retrieval images, n , in the UvC test. (b) The average computational times of the belief propagation module increases according to the number of ranks, n .

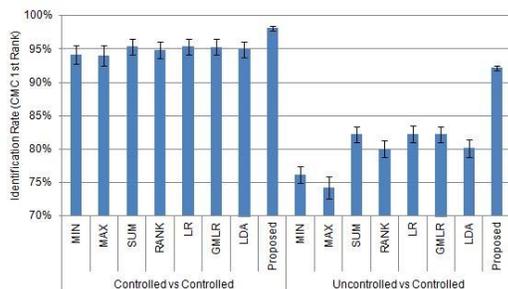


Figure 9. The performances of the proposed method are compared with well-known fusion methods that combined the ten proposed RSG classifiers. The proposed method is combined with the sum rule.

Now we will compare performance improvements according to the number of multiple classifiers. As shown in Figure 7, when the number of classifiers increases, the accuracy also increases in both the traditional and proposed methods. However, the proposed unified framework always leads to approximately 10% better recognition rates than the traditional one. In this paper, we use the sum rule [7] to combine RSG classifiers for simplicity, but more complex combination algorithms may further improve system performance.

The number, n , of candidate images at the compatibility matrix, Ψ , is an important factor for the performance of the proposed system because it is directly related to both accuracy and computational complexity. When the value of n increases, the size of the compatibility matrix, Ψ , also increases, which leads to heavy computational complexity. We evaluate the performance changes of the ten RSG classifiers, as shown in Figure 8. The average computational times of the belief propagation module are measured on a 2.66 GHz single processor PC. Considering both accuracy and complexity, we have selected $n = 300$ in this paper.

Figure 9 summarizes the performances of the well-known fusion methods for the ten RSG classifiers: the sim-

Table 3. The performances of the proposed method are compared with the Gabor-based classifiers such as the BGL and ECG classifiers. 1st rank of CMC is used for the identification measurement.

Test	Method	ROC1	ROC2	ROC3	Avg
CvC	BGL	95.86%	94.16%	93.76%	94.59%
	RSG (Sum)	96.69%	94.92%	94.39%	95.33%
	ECG (LR)	97.21%	95.69%	95.36%	96.09%
	Proposed (RSG)	98.53%	98.01%	97.77%	98.10%
	Proposed (ECG)	98.93%	98.44%	98.17%	98.52%
UvC	BGL	73.58%	71.24%	73.54%	72.79%
	RSG (Sum)	83.15%	80.85%	82.63%	82.21%
	ECG (LR)	86.86%	84.51%	85.64%	85.67%
	Proposed (RSG)	92.15%	91.85%	92.54%	92.18%
	Proposed (ECG)	94.06%	93.85%	94.42%	94.11%

ple ones (e.g., MIN, MAX, and Sum) [7], RANK based fusion [11], likelihood rate (LR) based fusion [16], Gaussian mixture model-based LR (GMLR) method [13], and fisher classifier (LDA) based score fusion method [18]. Their average recognition rates in the UvC test are 76.22%, 74.22%, 82.21%, 82.28%, 82.24%, and 80.15%, respectively, but the proposed method achieves an average recognition rate of 92.18%. This trend is similarly observed in the CvC test. In this respect, we can know that the proposed method achieves the best result among the published fusion methods.

When compared with the performances of the other Gabor-based methods such as BGL and ECG classifiers, as shown in Table 3, the twelve ECG classifiers fused by the LR method show an average of 85.67% in the UvC test, which is better than the 72.79% average of the BGL method and the 82.21% average of the ten RSG classifiers fused by the sum rule. Note that the ECG classifiers achieved a 91.24% Verification Rate (VR) at a 0.1% False Acceptance Rate (FAR) in the EXP4 of FRGC ver 2.0 database [4], and we can know that the first rank of the CMC measurement is more strict than the VR of the FAR=0.1% measurement. On the other hand, the proposed recognition framework works successfully for the RSG classifiers and ECG classifiers. For example, using the proposed recognition framework, the recognition rate of the RSG classifiers is boosted from 82.21% to 92.18% and the recognition rate of the ECG classifiers is boosted from 85.67% to 94.11% in the UvC test. From this result, we can conclude that the proposed method efficiently combines multiple classifiers using the classifier relationship in the known variation test.

5.3. Experiment with the Unknown Variation Test

In this test protocol, we expect to observe performance changes according to the different un-trained image variations. Figure 10 summarizes the performance comparisons of various face recognition methods, including the LDA, Boosted Gabor (BG) [23], and BGL [19] methods. Moreover, we also measure the accuracies of the ten RSG clas-

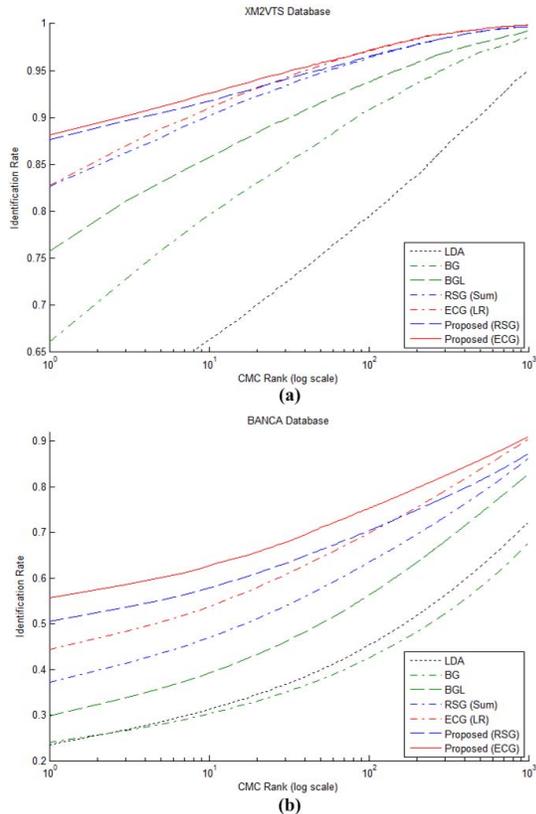


Figure 10. CMC curves corresponding to the well-known recognition methods and the proposed framework using RSG classifiers and ECG classifiers in the (a) XvX test and (b) BvB test, respectively. X-axis is log-scaled.

sifiers merged by sum rule, denoted by RSG (Sum), the twelve ECG classifiers merged by the LR method, denoted by ECG (LR), and the proposed frameworks with the RSG and ECG classifiers, respectively, in the XvX and BvB tests. In the experiments from the XvX test to BvB test, the traditional methods, for example, the LDA, BG, and BGL methods, show poorer accuracies than expected, because they have a single classifier, which is not robust to the situation changes. On the other hand, the multiple classifiers, for example, the RSG and ECG classifiers, show better results. In detail, the RSG and ECG classifiers show similar recognition rates, approximately 82%, in the XvX test, but the ECG classifiers achieve a 7% better recognition rate than the RSG classifiers in the BvB test. On the other hand, the proposed frameworks, the RSG and the ECG, always lead to approximately 5–11% better results than the other score fusion methods, for example, the RSG (Sum) and the ECG (LR). The best results are achieved by the proposed framework with twelve ECG classifiers, which score 88.15% and 55.7%, on the XvX and BvB tests, respectively. Note that compared to the performances of the proposed method in the known variation test protocol (i.e., 94–98%) the best ac-

curacies of the proposed method decrease in both the XvX and BvB tests. The difference in its performance between the known and unknown tests is caused by unexpected image variations. For example, XM2VTS images are collected in a studio to control for environmental conditions, which is similar to the FRGC database, and the performance of the proposed method falls slightly, but the BANCA images are collected from different places such as an office or a dining hall. They also use two different cameras, including a webcam, and some images are quite blurry. Such unexpected variations could have negative effects on the recognition system. Nevertheless, the overall performances of the proposed method are always better than the other methods in the unknown test protocols, which indicates that the superiority of the proposed method is well generalized over the different databases.

6. Conclusion

We propose a novel face recognition framework, particularly for the one-to-many identification task, based on multiple classifiers connected by a Markov network. The Markov network probabilistically models the relationships between a query and gallery images and between neighboring gallery images. From the viewpoint of an observation-hidden node pair, we retrieve the most similar gallery images from the database using a query image face model. The statistical dependency between the hidden nodes is calculated by the similarities between the retrieved gallery images. Hence, the resulting inference mechanism can be viewed as a kind of clustering-based face recognition. We prove the good performances of the proposed framework using RSG classifiers and ECG classifiers, respectively. Moreover, we have confirmed the generality of the proposed method with the known variation test and unknown variation test protocols which consist of three different databases: the FRGC, XM2VTS, and BANCA databases.

7. Acknowledgements

The authors would like to thank the anonymous reviewers for their constructive comments that greatly contributed to improving the final version of the paper. This work was partly supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2010-0028680).

References

- [1] E. Bailly-Bailliere, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariethoz, J. Matas, K. Messer, V. Popovici, F. Poree, and et al. The banca database and evaluation protocol. *Audio- and Video-based Biometric Person Authentication*, 2688/2003, 2003. 4

- [2] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low-level vision. *International Journal of Computer Vision*, 40(1):25–47, Jul. 2000. 2
- [3] R. Huang, V. Pavlovic, and D. N. Metaxas. A hybrid face recognition method using markov random fields. *International Conference on Pattern Recognition*, 3:157–160, Aug. 2004. 2, 3
- [4] W. Hwang, X. Huang, K. Noh, and J. Kim. Face recognition system using extended curvature gabor classifier bunch for low-resolution face image. *IEEE CVPR Workshops*, pages 15–22, Jun. 2011. 2, 5, 6
- [5] W. Hwang, H. Wang, H. Kim, S. Kee, and J. Kim. Face recognition system using multiple face model of hybrid fourier feature under uncontrolled illumination variation. *IEEE Trans. on Image Processing*, 20(4):1152–1165, Apr. 2011. 1
- [6] T. Kim, H. Kim, W. Hwang, S. Kee, and J. Lee. Component-based LDA face descriptor for image retrieval. *British Machine Vision Conference*, pages 507–516, Sep. 2002. 1
- [7] J. Kittler, M. Hatef, R. P. Duin, and J. G. Matas. On combining classifiers. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 3(20):226–239, Mar. 1998. 6
- [8] S. Z. Li and A. K. Jain. *Handbook of Face Recognition*. Springer, 2005. 1, 5
- [9] Z. Liu and C. Liu. Robust face recognition using color information. *Advances in Biometrics*, 5558/2009:122–131, 2009. 1, 2
- [10] A. Martinez and A. Kak. PCA versus LDA. *IEEE Trans. Pattern Recognition and Machine Intelligence*, 23(2):228–233, Feb. 2001. 1
- [11] K. McDonald and A. F. Smeaton. A comparison of score, rank and probability-based fusion methods for video shot retrieval. *International Conference on Image and Video Retrieval*, pages 61–70, 2005. 6
- [12] K. Messer, J. Kittler, M. Sadeghi, S. Marcel, C. Marcel, S. Bengio, F. Cardinaux, C. Sanderson, J. Czyz, L. Vandendorpe, and et al. Face verification competition on the xm2vts database. *Audio- and Video-based Biometric Person Authentication*, 2688/2003, 2003. 4
- [13] K. Nandakumar, Y. Chen, S. C. Dass, and A. K. Jain. Likelihood ratio based biometric score fusion. *IEEE Trans. On Pattern Recognition and Machine Intelligence*, 30(2):342–347, Feb. 2008. 6
- [14] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. *Proc. IEEE, Computer Vision and Pattern Recognition*, pages 84–91, Jun. 1994. 1
- [15] P. J. Phillips, P. J. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. *Proc. IEEE, Computer Vision and Pattern Recognition*, 1:947–954, San Diego, Jun. 2005. 4
- [16] S. Prabhakar and A. K. Jain. Decision-level fusion in fingerprint verification. *Pattern Recognition*, 4(35):861–874, Apr. 2002. 6
- [17] R. N. Rodrigues, G. N. Schroeder, J. J. Corso, and V. Govindaraju. Unconstrained face recognition using mrf priors and manifold traversing. *International Conference on Biometrics: Theory, Applications and Systems*, pages 1–6, Sep. 2009. 2
- [18] A. Ross and A. K. Jain. Information fusion in biometrics. *Pattern Recognition Letters*, 24:2115–2125, Sept. 2003. 6
- [19] S. Shan, P. Yang, X. Chen, and W. Gao. Adaboost gabor fisher classifier for face recognition. *Proceedings of International Workshop on Analysis and Modeling of Faces and Gestures*, pages 278–291, 2005. 5, 6
- [20] Y. Su, S. Shan, X. Chen, and W. Gao. Hierarchical ensemble of global and local classifiers for face recognition. *IEEE Trans. on Image Processing*, 18(8):1885–1896, August 2009. 2
- [21] X. Tan and B. Triggs. Fusing gabor and lbp feature set for kernel-based face recognition. *IEEE International Workshop on Analysis and Modeling of Face and Gestures*, pages 235–249, 2007. 1, 2
- [22] X. Wang and X. Tang. Random sampling for subspace face recognition. *International Journal of Computer Vision*, 70:91–104, 2006. 5
- [23] P. Yang, S. Shan, W. Gao, S. Z. Li, and D. Zhang. Face recognition using ada-boosted gabor features. *Proc. IEEE, Automatic Face and Gesture Recognition*, pages 356–361, May 2004. 5, 6