Fast and Scalable Enrollment for Face Identification based on Partial Least Squares

Gérson de Paulo Carlos, Helio Pedrini, William Robson Schwartz

***Abstract*— Face recognition has received increased attention due to its application in biometrics and surveillance systems, emerging two main tasks, verification and identification of faces. The first one aims at accepting or rejecting an identity assigned to a correct face, whereas the second aims at, given an unknown probe face, finding the best identity to it from a gallery of known faces. For the face identification problem, discriminative approaches such as the one-against-all method have achieved higher accuracy than descriptive approaches such as eigenfaces. However, such methods have scalability issues when new subjects are enrolled in the gallery once it is necessary to rebuild all discriminative models to take into account the new individuals. This work describes and evaluates a novel method for making the process of gallery maintenance more efficient. This method employs an association between the one-against-some classification scheme, which differently from the one-against-all approach that considers a random subset of subjects as counterexamples, and the use of a priority queue to provide a scalable approach to enrolling new subjects to the gallery. Experimental results obtained by applying the proposed method on publicly available face data sets demonstrate its advantage when compared to the one-against-all approach.**

I. Introduction

Face recognition is a very active research topic due to its applications in areas such as surveillance, biometrics and human computer interaction. Face verification and identification figure among the main tasks performed by face recognition. While the former is responsible for accepting or denying the identity claimed by an individual given a pair of samples, the latter focuses on matching a sample of an unknown person to a gallery of known subjects.

The identification task presents particular interest in surveillance applications that perform face recognition in monitored areas, in which the identity of individuals needs to be determined to provide, for example, non-intrusive monitoring of circulation on restricted areas. Due to the dynamic nature of these environments, in which new subjects are incrementally added, the identification system not only needs to be accurate, but also it is important to provide efficient and robust enrollment mechanisms.

Due to its ability of generating discriminative subspaces and working with few high dimensional input samples, the statistical method *Partial Least Squares* (PLS) [1] has been successfully employed to the problem of face recognition in the past few years for both verification and identification tasks [2]–[7]. Even though the one-against-all classification scheme combined with PLS has provided significant improvements on the recognition rates for the identification task [2], [4], [6], it presents the drawback of not being scalable to the enrollment of new subjects since all existing PLS models representing the subjects in the gallery need to be rebuilt, leading to a high computational cost proportional to the gallery size.

To handle the enrollment of new subjects to the gallery while maintaining the generation of highly discriminative subspaces with PLS, we employ a classification scheme called one-against-some, which does not consider all remaining subjects as counterexamples, but only a subset of them. The one-against-some approach maintains a trade-off between the discriminatory power achieved by the one-against-all, in which all remaining subjects are set as counterexamples, and the scalability when new subjects are enrolled in the gallery since the PLS models used to represent subjects already enrolled do not need to be rebuilt.

The method proposed in this work creates few PLS models for each subject considering a random subset of the remaining subjects already enrolled as counterexamples. When a new subject is added to the gallery, new PLS models are built without the need for rebuilding previously constructed models since not all subjects are required to be used as counterexamples. In addition, we propose the use of a priority queue to maintain a low number of projections when a probe sample is presented to the system. Such combination allows a fast enrollment maintaining the accurate results achieved by the one-against-all approach.

Experimental results obtained using two publicly available data sets to perform face identification, FRGC [8] and PubFig83 [9], show that even though the computational cost for the enrollment becomes constant (different from the quadratic rate obtained when the one-against-all is employed), the recognition rates achieved and the number of projections performed are comparable to ones achieved by the one-against-all approach.

II. Related Work

This section briefly reviews and discusses some concepts and references related to the topic investigated in our work.

A. Face Recognition

Significant advances have been achieved over the past decade related to the face recognition problem. Comprehensive surveys on face recognition, including identification and verification tasks, can be found in the literature [10]–[14].

G.P. Carlos and H. Pedrini are with Institute of Computing, University of Campinas, Campinas-SP, Brazil, 13083-852

W.R. Schwartz is with Department of Computer Science, Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brazil, 31270-010

Face recognition techniques can be classified into two main categories, holistic and local approaches. Holistic approaches [15]–[17] extract information from the entire face image to perform the recognition. To reduce the high-dimensionality of data, face images are projected onto a lower dimensional space. On the other hand, local approaches [18]–[22] extract information from local facial features and discriminate faces through the comparison and combination of local statistics.

Even though several proposed approaches have achieved high recognition accuracy rates under controlled conditions, several factors demonstrate the face recognition problem to be more complex due to pose, occlusion, illumination, facial expression, scalability, and other real conditions. To perform recognition under fairly uncontrolled conditions, the works in [23], [24] focus on illumination normalization, approaches developed in [7], [25] deal with pose variations. In addition, face recognition via sparse representation-based classification has become popular [26], providing improved results.

Another important issue in considering face identification systems is their scalability since the problem can involve the recognition of a large number of individuals. Thus, search techniques for matching probe samples to the face gallery must be efficient [27], [28]. Furthermore, the need of dynamically rebuilding the gallery models whenever a new subject is added can compromise the system performance. The latter problem is the main focus of this work.

B. Feature Descriptors

A variety of feature descriptors have been employed in face recognition. Scale-invariant feature transform (SIFT) [29] and histogram of oriented gradients (HOG) [30] are powerful descriptors used to extract facial features. Local binary patterns (LBP) [31] describe texture through histograms of labels assigned to the image pixels by thresholding a neighborhood of each pixel with the center value and considering the result as a binary number. LBP features are invariant to monotonic intensity changes. Gabor filters [32] extract facial features characterized by spatial locality, spatial frequency, and orientation selectivity to overcome image variabilities due to illumination and facial expression changes. Variations or combinations based on LBP and Gabor descriptors have also been proposed [31], [33] for face recognition.

More recently, methods for combining several features have been proposed for face recognition [34], [35]. Global features, where the entire image is used to construct the feature vector, are combined with local features extracted from regions of the image. Since certain global and local features are complementary to each other, results from such combination can outperform the application of individual features.

C. Partial Least Squares

The construction of person-specific subspaces for human face recognition has been explored to capture variations of a same person in order to produce more effective representations of individuals under varying illumination or pose. Some subspace methods found in the literature include Eigenface [15], Fisherface [16], Tensorface [36] and Bayesian algorithms [37]. These methods seek to model face variations under illumination changes through a set of training face set, such that new face samples can be projected into low dimensional spaces and compared against a number of images.

Although face recognition performance can be improved in lower dimensional subspaces, most of these methods do not work well in the case of unconstrained face images, that is, they typically capture either just structure that is common to all faces or just structure that is discriminative between two sets of faces. Despite such difficulties, the recent application of the subspace method based on Partial Least Squares has achieved improved results in face recognition [2]–[7].

Partial Least Squares (PLS) is a statistical method used to find relations between observed variables through the estimation of a low dimensional latent space that maximizes the separation between samples with different characteristics [38]. PLS estimates latent variables as linear combinations of the original variables in a matrix $\boldsymbol{X}$, composed of variables used to describe samples, and a matrix $\boldsymbol{Y}$ containing a set of response variables.

Let a problem with n samples described by d variables each, stored in a mean-centered matrix $\boldsymbol{X}_{n\times d}$, associated to k response variables, stored in a mean-centered matrix $\boldsymbol{Y}_{n\times k}$. PLS estimates a p-dimensional space by decomposing $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively, into $\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^T + \boldsymbol{E}$ and $\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{Q}^T + \boldsymbol{F}$, where $\boldsymbol{T}_{n\times p}$ and $\boldsymbol{U}_{n\times p}$ are matrices composed of the latent variables, matrices $\boldsymbol{P}_{d\times p}$ and $\boldsymbol{Q}_{k\times p}$ represent the loadings, and matrices $\boldsymbol{E}_{n\times d}$ and $\boldsymbol{F}_{n\times k}$ are the residuals.

A common method for iteratively extracting the latent variables is the non-linear iterative partial least squares (NIPALS) [1], so that matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ are decomposed by subtracting their rank-one approximations as $\boldsymbol{X}_{i+1} = \boldsymbol{X}_i - \boldsymbol{t}_i\boldsymbol{p}_i^T$ and $\boldsymbol{Y}_{i+1} = \boldsymbol{Y}_i - \boldsymbol{t}_i\boldsymbol{q}_i^T$, respectively, where $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$ are the data representation for the i-th iteration, where $\boldsymbol{t}_i$ represent the i-th columns of matrices $\boldsymbol{T}$, $\boldsymbol{X}_1 = \boldsymbol{X}$ and $\boldsymbol{Y}_1 = \boldsymbol{Y}$, and $\boldsymbol{p}_i$ and $\boldsymbol{q}_i$ denote the i-th columns of the matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$, respectively. After the extraction of p projection vectors, the p-dimensional representation of $\boldsymbol{X}_{n\times d}$ is given by $\boldsymbol{T}_{n\times p}$, which is used to extract the regression coefficients $\boldsymbol{\beta}_{d\times k}$ by $\boldsymbol{\beta} = \boldsymbol{W}(\boldsymbol{P}^T\boldsymbol{W})^{-1}\boldsymbol{T}^T\boldsymbol{Y}$. Finally, the regression responses, Y_v, for a feature vector $\boldsymbol{v}_{d\times 1}$ is obtained by $Y_v = \overline{Y} + \boldsymbol{\beta}^T\boldsymbol{v}S$, where $\overline{Y}_{1\times k}$ is the sample mean of each variable of $\boldsymbol{Y}$ and $S_{1\times k}$ is the standard deviation of the variables in $\boldsymbol{Y}$.

D. Classification Schemes

The process of matching a probe sample to a gallery of faces can follow some strategies, including the one-against-all scheme [39], where all training models need to be rebuilt when a new subject s_i is presented to the system, since all the remaining samples are used as counterexamples (negative

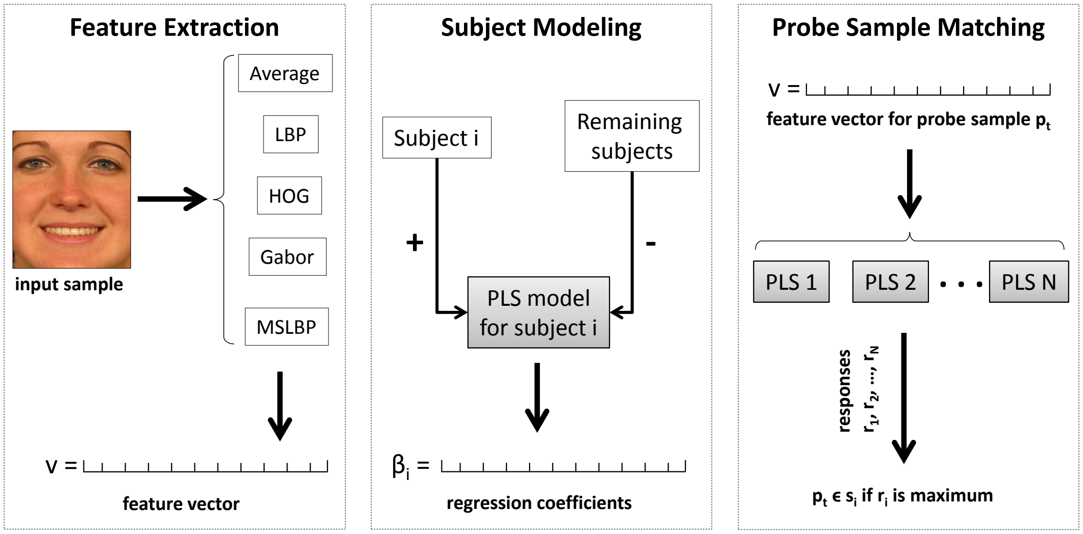


Fig. 1. Diagram showing the steps performed in the one-against-all approach.

class) for s_i. Although this scheme can achieve higher recognition rates, it is not efficient in terms of training time.

Another strategy, known as pairwise classification [40], transforms a multi-class problem into a series of two-class problems. Instead of constructing one binary classifier for each class c_i as in the one-against-all, so that positive training samples are those that belong to class c_i and the negative training samples are formed by all other classes, the pairwise classification converts an n-class problem into $n(n-1)/2$ binary problems, one for each pair of classes.

A classification scheme, called one-against-some [41], has been recently proposed in the context of person re-identification, where only a subset of individuals are considered as negative samples, instead of remaining subjects as in the one-against-all classification scheme. The work in [41] presented the main concepts of the one-against-some classification scheme without focusing on the computational time, which is considered in this work with the employment of a priority queue.

III. Methodology

In this section, we first describe the face identification approach proposed by Schwartz et al. [2], which performs face identification based on a one-against-all classification scheme. Then, based on the method in [2], we discuss the employed one-against-some classification scheme. Finally, we employ a priority queue intended to speed-up the matching process for probe samples when the one-against-some approach is used.

A. Face Identification based on One-Against-All Approach

The face identification method proposed by Schwartz et al. [2] is structured in three main steps: feature extraction, individual modeling for subjects in the gallery and execution of regressions to perform the matching of probe samples to subjects in the gallery. These steps are briefly described as follows and illustrated in Figure 1. For more details, we refer the reader to [2].

During the feature extraction, after cropping the face region, it is scaled to a given size and each sample is decomposed into a set of overlapping blocks from which feature descriptors are extracted. A combination of different descriptors is considered, which includes information regarding to shape extracted by histograms of oriented gradients (HOG) [30], textural information with the use of the original local binary patterns (LBP) [20] and one of its extensions called multi-scale local binary patterns (MSLBP) [31], color information captured by the average of pixel colors and salient visual properties extracted using Gabor filters [32]. After the feature extraction for all blocks, the descriptors are concatenated in a feature vector $\boldsymbol{v}$, used to describe the face.

The procedure to estimate PLS models for each subject in the gallery $g = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N\}$, in which $\boldsymbol{s}_i$ denotes samples of the i-th subject represented by their feature vectors, is the following. In order to increase the discriminability between classes, PLS models based on the one-against-all classification scheme is considered. Therefore, when the i-th subject is being modeled, the samples of the remaining subjects, $g \setminus \boldsymbol{s}_i$, are used as counterexamples. The PLS estimates the discrimination ability of the descriptors in the feature vector and returns regression coefficients $\boldsymbol{\beta}_i$. This process is executed for each subject in the gallery. Therefore, at the end, N PLS models will be estimated.

After the construction of all PLS models, their regression coefficients are used to perform the matching between probe samples and subjects in the gallery. When a probe sample is presented to the identification system, its feature vector is used to evaluate the regression response for each PLS model. The subject associated with the highest regression response (a high regression response indicates that feature vector of the probe and subject's samples are similar) is considered to be the best match for the probe sample.

The main drawback of the one-against-all classification

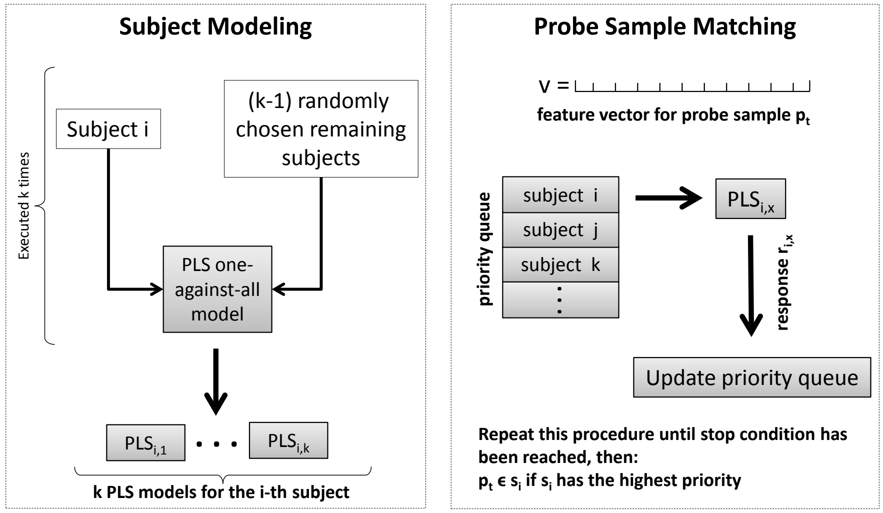


Fig. 2. Diagram illustrating the steps performed in the one-against-some approach. First step - Subject Modeling - shows that k different PLS models for each subject in the gallery are built. The selection of $(k-1)$ subjects is performed in a such way that the distribution throughout the models is uniform. Second step - Probe Sample Matching - describes the use of the priority queue when a probe sample is presented. Once the queue is initialized, the probe sample is projected onto the model $\text{PLS}_{i,x}$ associated with the subject s_i with the highest priority. Then, if the response obtained, $r_{i,x}$, is the current minimum for subject s_i, $r_{i,x}$ is set as the priority of s_i. This step is repeated until a stopping condition has been reached.

scheme is its high computational cost to enroll new subjects, once all existing PLS models need to be rebuilt to consider samples of the new subject as counterexamples. To avoid the reconstruction of all PLS models, we employ the classification scheme called one-against-some. This approach builds PLS models considering only a subset of subjects as counterexamples, which avoids the reconstruction of the PLS models already built.

B. One-Against-Some

Previous works employing one-against-all approach have accomplished excellent results in terms of accuracy [2], [4] when estimating best match for N known subjects. However, they have failed making the scheme scalable whenever a new subject is presented to the system, as all models require to be rebuilt.

To overcome the scalability problem when new subjects are enrolled, a similar method to one-against-all is employed, called one-against-some. This approach requires only some of the remaining subjects to be added as negative samples. Due to this change, more PLS models are created aiming at representing every subject accurately, resulting therefore, in more than one regression response per subject.

The intuition to choose the one-against-some approach comes from the fact that while the addition of all remaining subjects as counterexamples might present redundant information to build the PLS model, the application of the pairwise approach may not be enough to emphasize the most discriminative feature descriptors. This way, the one-against-some presents a trade-off between computational cost to add a new subject and high discriminability obtained when all remaining subjects are considered.

As it will be shown in the experiments, the accuracy is improved when more than one PLS model is estimated for each subject (considering subjects randomly chosen as counterexamples), which increases the number of projections required when probe samples are presented. However, the use of a priority queue to control which models should be considered first, makes this problem negligible.

Figure 2 illustrates the proposed approach, which is a variation of the one-against-all (Figure 1) with modifications in the last two steps: subject modeling and probe sample matching.

The subject modeling describes the creation of k PLS models for an i-th subject, using $k-1$ randomly selected between all subjects but the i-th individual as negative class. This step results in kN models, where N is the number of known subjects by the time of enrollment and k denotes the number of subjects used as counterexamples (the value of k is estimated in Section IV-B).

The probe sample matching step aims at finding the best matching using a priority queue to reduce the number of projections needed since testing all models would cost $N \times k$ projections opposed as N projections required by the one-against-all technique. The next section discusses the use of the priority queue.

When a new subject is enrolled in the gallery, k new PLS models are computed, in which each model contains samples of this new subject as the positive class and a subset containing $k-1$ individuals randomly chosen from subjects already enrolled in the gallery as negative class. Since only a subset of individuals is considered, there is no need to rebuild the existing models, which would be necessary for the one-against-all scheme once it requires the use of all remaining subjects as counterexamples.

C. Priority Queue

After the modeling step has been executed, each subject possesses k PLS models as positive samples. Thus, a naive search for the best matching subject for a probe sample would cost $N \times k$ projections, in contrast to the N projections required by the one-against-all classification scheme. In order to reduce that number of projections, a better technique for evaluating the probe samples is required.

Aiming at avoiding unnecessary projections (those PLS models associated with subjects unlikely to be the correct match), a priority queue-based approach is proposed. This way, the search relies on testing subjects with high chance of being the correct match. Thus, projections onto models associated with subjects with low priority might be avoided, reducing the computational cost.

The proposed queue associates a priority value to each subject. When a probe sample is presented to the system, the priorities are initialized to be the same for all subjects. A PLS model containing the subject with highest priority (a random subject when the process starts) in the positive class is chosen and the probe sample is projected onto it. The priority of the subject is updated according to its regression response (a high response increases the priority). Then, another PLS model associated with the subject presenting the highest priority is considered for projection. This process is repeated until a stop condition has been reached. Algorithm 1 describes the search process using the priority queue.

Algorithm 1: Querying and priority queue

Initialization:
- initialize queue with null values;
- choose randomly a subject of gallery;
- select and mark a PLS model containing this subject as positive class;
- project the probe sample onto the model;
- update the queue with the regression response;

Main loop:

while *there are unevaluated subjects* **do**
- select the subject s_i with the highest priority;
- **if** *there are unmarked models for* s_i **then**
  - select an unmarked model, $\text{PLS}_{i,x}$;
- **else**
  - break;
- project probe face onto $\text{PLS}_{i,x}$, mark it, and retrieve response $r_{i,x}$;
- update the priority of s_i if it is greater than $r_{i,x}$;

- the subject with the highest priority is nominated as the best match.

Using the priority queue without a stop condition would cause it to test every PLS model. Therefore, a stopping condition was incorporated to the system: once all models having the subject with the highest priority as positive class have been considered, the search stops if at least one model associated with each subject has been used to project the probe sample.

Algorithm 1 shows that priorities are defined as the minimum response obtained. That approach works because regressions responses are positive for the correct class and negative for incorrect classes. Therefore, even the minimum response will be positive for the desired class. This way, after evaluating all models containing a given subject in the positive class, if this subject is still with the highest priority, it is likely to be the best match for the probe sample.

IV. Experimental Results

This section evaluates several aspects of the proposed method. First, Section IV-A describes the data sets used in the experimental validation. Then, the estimation of parameter k and the effectiveness of using a priority queue are presented in Sections IV-B and IV-C, respectively. Section IV-D shows results regarding the incremental enrollment of subjects in the gallery. Finally, Section IV-E shows the recognition rates achieved in the considered data sets and other results from the literature are shown as reference.

A. Datasets

The proposed method is evaluated on two data sets used for face recognition: FRGC version 1 [8] and PubFig83 [9]. These data sets present different characteristics, while FRGC presents a challenging experiment (Experiment 4), which considers few images acquired under controlled conditions and frontal pose for the gallery and images acquired under uncontrolled conditions for the probe, PubFig83 is composed of several uncontrolled images with pose and expression variations.

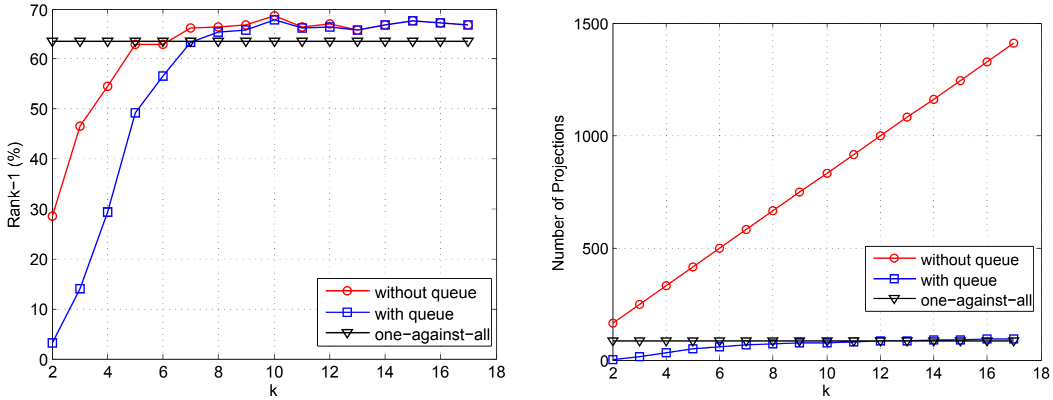
For the FRCG data set, we follow the feature extraction procedure and the evaluation protocols used by Schwartz et al. [2]. The FRGC version 1 for 2D still images considers three experiments, each one with 152 subjects in the gallery: Experiment 1 considers a single controlled sample to build the gallery and controlled probe images; Experiment 2 considers a gallery with four controlled still images per subjects; and Experiment 4 considers a gallery with a single controlled sample per subject to build the gallery and multiple uncontrolled probe images.

For the PubFig83 data set, composed of 83 different subjects, we follow the evaluation protocol defined by Pinto et al. [9], in which 90 samples per subject are used to build the gallery. Due to the large number of samples used in the gallery, we employed the feature extraction procedure defined in [42], which considers fewer descriptors per sample compared to the work in [2]. The samples for the PubFig83 data set were rescaled to 100×100 pixels and the number of descriptors per sample is 6039. We executed the method ten times, each time with a different split between samples used to build the gallery and probe samples and we report the average rank-1 recognition rates.

B. Number of Counterexamples

As described in Section III, the one-against-some approach depends on the parameter k, which defines the number of random subjects that will be added as counterexamples during the PLS model estimation.

To estimate the value of k that will be used in all remaining experiments, we have considered a subset (with fewer samples) of the PubFig83 data set and executed the



(a) Rank-1 recognition rates

(b) Average number of projections

Fig. 3. Estimation of parameter k and employment of the priority queue considering classification schemes one-against-all and one-against-some. Note that the recognition rates and the number of projections achieved by the one-against-all approach is independent of k.

face identification for multiple values of k. The results are shown in Figure 3(a).

We can see that the rank-1 recognition rates increase up to $k = 10$, which will be used in the remaining experiments for all data sets. It is important to note that the recognition rates achieved by the one-against-some approach for $k = 10$ are very similar to those achieved by the one-against-all. In addition, the results obtained with the one-against-some for smaller values of k are very poor, which reflects the lack of discriminability obtained by the PLS models when few subjects are used as counterexamples.

C. Priority Queue

We also evaluate the reduction in computational cost achieved through the priority queue approach proposed in Section III-C. Figure 3(b) compares the average number of projections required to evaluate one probe sample by the one-against-all and the one-against-some approaches; for the latter, the results are shown with and without the use of the priority queue. As in previous works [2], [42], the number of projections was chosen as metric due to its independence regarding implementation and hardware configuration.

To test a probe sample, the original one-against-all PLS approach requires one projection per subject in the gallery and the one-against-some classification scheme considering k subjects as counterexamples requires nk projections. To reduce the number of projections, a priority queue has been employed. According to the results shown in Figure 3(b), the one-against-some with priority queue achieves a number of projections very similar to the one-against-all approach for $k = 10$ without reduction in the rank-1 recognition rates, as seen in Figure 3(a). Therefore, the remaining experiments will also consider the use of the priority queue.

D. Incremental Enrollment

In this section, we evaluate the behavior of the proposed method when new subjects are incrementally added in the gallery and compare the results with the original one-against-all approach. To perform this evaluation, we build an initial gallery with few subjects (15 subjects in this experiment) and project probe samples to these models. Then, new subjects are incrementally added in the gallery until all subjects have been enrolled.

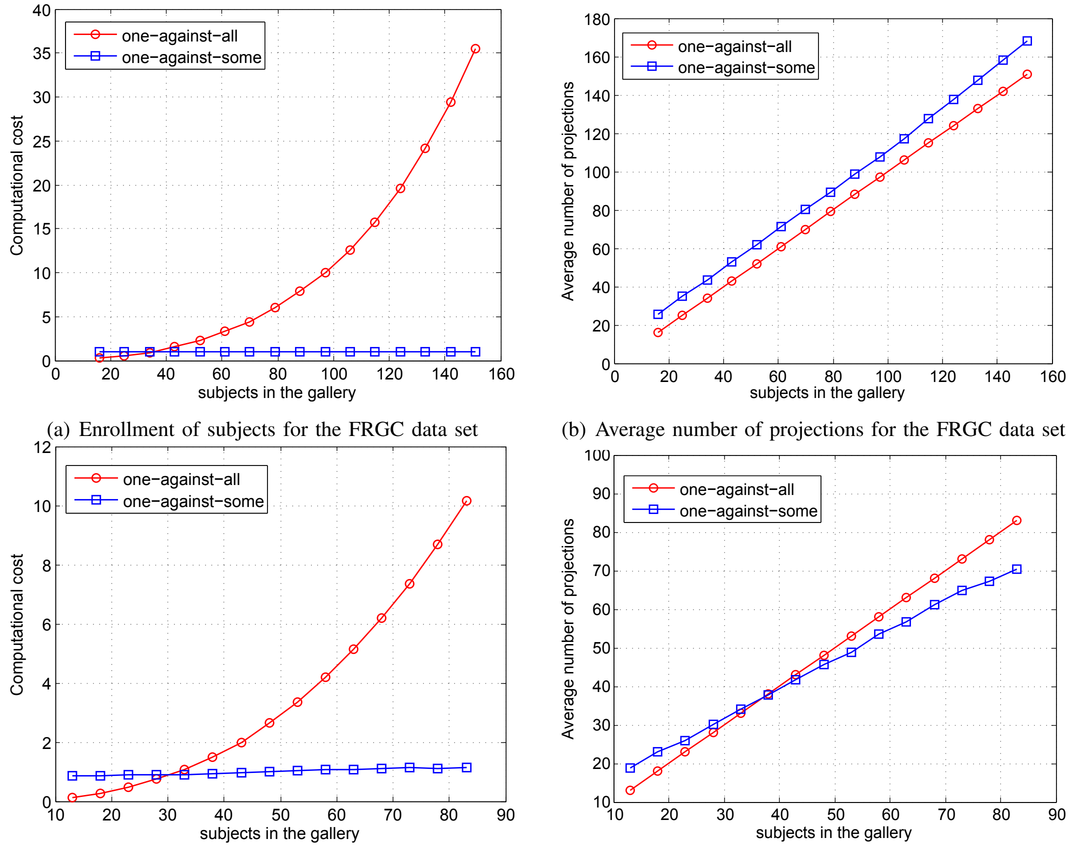
The first experiment evaluates the computational cost to add a new subject (the m-th) in the gallery considering that $m - 1$ subjects have already been added. The plots in the left column of Figure 4 show the results. To present relative results, the computational time is divided by the average time spent to add one subject when the one-against-some approach is considered. The results show that the one-against-some approach performs the addition of a subject in constant time, whereas the one-against-some increases in a quadratic rate as the number of subjects increases. Therefore, the scalability of the approach regarding the enrollment of subjects is demonstrated once the computational cost is constant and does not depend on the number of subjects in the gallery.

The second experiment, whose results are shown in the second column of Figure 4, evaluates the average number of projections required to test a probe sample. Even though the number of projections per probe sample performed by the one-against-some approach should be $n \times k$, the employment of the priority queue described in Section III-C allows the number of projections to be very similar to the results achieved by the one-against-all approach (n projections per probe sample).

According to the results, we can conclude that the approach described in this work is able to handle incremental enrollment of subjects in a scalable way and, even though more PLS models are built, the priority queue allows the number of projections to be comparable to the number achieved by the one-against-all. Results regarding the recognition rates are shown and discussed in the next section.

E. Recognition Rates

Since the goal of this work is to provide a scalable strategy for adding new subjects to the gallery, we compare the proposed method with the original one-against-all with respect to the recognition rates. Even though the aim of this work is not improve the best results found in the literature, we also show some state-of-the-art results in both data sets.



(a) Enrollment of subjects for the FRGC data set

(b) Average number of projections for the FRGC data set

(c) Enrollment of subjects for the PubFig83 data set

(d) Average number of projections for the PubFig data set

Fig. 4. Incremental enrollment of the subjects. (a) and (c) show the computational cost normalized by the average computational time to add a subject using the one-against-some approach; (b) and (d) show the average number of projections performed when a probe samples is presented to the system.

According to the results shown in Tables I and II, the recognition rates achieved with the proposed method and with the one-against-all approach are very similar. When compared to the state-of-the-art methods, the proposed approach presents lower recognition rates. That is because it is limited to the feature descriptors and the PLS approach defined by the one-against-all method in [2]. For example, the method proposed by Choi et al. [4] considers an enhanced set of feature descriptors and the work proposed by Chiachia et al. [6] considers an SVM classifier associated with biologically inspired feature descriptors.

We have shown in the previous section that the proposed method is able to add new subjects in the gallery at a constant cost, the average number of projections to test a probe sample is similar to the original one-against-all approach, and the results described in this section demonstrated that recognition rates are comparable to those achieved by the one-against-all approach. Therefore, the proposed method demonstrated to be scalable and can be employed as an alternative to the traditional one-against-all classification scheme.

V. Conclusions

This work proposed and evaluated a new approach to maintaining a gallery of faces in an efficient way. A one-against-some scheme is employed in the construction of models through random selection of a subset of individuals. The use of a priority queue is proposed to reduce the number of projections when a probe sample is presented to the system. Experimental results obtained on two publicly available face data sets demonstrated that the proposed method is effective and scalable to identify faces.

TABLE I

Performance comparison (rank-1 recognition rates in %) for the FRGC data set.

Method	Exp.1	Exp.2	Exp.4
LC_1C_2 [43]	-	-	75.00
ROCA [44]	-	96.40	75.50
CS-POP [4]	98.00	99.80	89.00
PLS one-against-all [2]	97.90	99.80	86.20
proposed PLS one-against-some	97.20	99.30	84.70

TABLE II

Performance comparison (rank-1 recognition rates $\pm$ std. err. (%)) for the PubFig83 data set.

Method	Recognition Rate
HT-L3-1st [9]	87.11 ± 0.56
PS-PLS [6]	88.75 ± 0.26
PLS one-against-all [2]	70.59 ± 0.36
proposed PLS one-against-some	73.27 ± 0.41

Acknowledgements

The authors are grateful to FAPESP, FAPEMIG, CNPq, and CAPES for the financial support.

References

[1] H. Wold, "Partial Least Squares," in *Encyclopedia of Statistical Sciences*. New York, NY, USA: Wiley, 1985, vol. 6, pp. 581–591.

[2] W. R. Schwartz, H. Guo, J. Choi, and L. S. Davis, "Face Identification Using Large Feature Sets," *IEEE Trans. on Image Processing*, vol. 21, no. 4, pp. 2245–2255, 2012.

[3] H. Guo, W. R. Schwartz, and L. S. Davis, "Face Verification using Large Feature Sets and One Shot Similarity," in *International Joint Conference on Biometrics*, 2011.

[4] J. Choi, H. Guo, W. R. Schwartz, and L. S. Davis, "A Complementary Local Feature Descriptor for Face Identification," in *IEEE Workshop on Applications of Computer Vision*, Breckenridge, CO, USA, Jan. 2012, pp. 121–128.

[5] A. Sharma and D. Jacobs, "Bypassing Synthesis: PLS for Face Recognition with Pose, Low-Resolution and Sketch," in *IEEE Intl. Conference on Computer Vision and Pattern Recognition*, 2011, pp. 593–600.

[6] G. Chiachia, N. Pinto, W. R. Schwartz, A. Rocha, A. X. Falcao, and D. Cox, "Person-Specific Subspace Analysis for Unconstrained Familiar Face Identification," in *British Machine Vision Conference*, 2012.

[7] A. Li, S. Shan, X. Chen, and W. Gao, "Cross-Pose Face Recognition based on Partial Least Squares," *Pattern Recognition Letters*, vol. 32, no. 15, pp. 1948–1955, 2011.

[8] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. C., K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the Face Recognition Grand Challenge," in *IEEE Intl. Conference on Computer Vision and Pattern Recognition*, 2005, pp. 947–954.

[9] N. Pinto, Z. Stone, T. Zickler, and D. D. Cox, "Scaling Up Biologically-Inspired Computer Vision: A Case Study in Unconstrained Face Recognition on Facebook," in *IEEE Intl. Conference on Computer Vision and Pattern Recognition*, 2011.

[10] X. Zhang and Y. Gao, "Face Recognition Across Pose: A Review," *Pattern Recognition*, vol. 42, no. 11, pp. 2876–2896, 2009.

[11] A. Tolba, A. El-Baz, and A. El-Harby, "Face Recognition: A Literature Review," *International Journal of Signal Processing*, vol. 2, pp. 88–103, 2006.

[12] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.

[13] R. Jafri and H. R. Arabnia, "A Survey of Face Recognition Techniques," *Journal of Information Processing Systems*, vol. 5, no. 2, pp. 41–68, 2009.

[14] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D Face Recognition: A Survey," *Pattern Recognition Letters*, vol. 28, no. 14, pp. 1885–1906, 2007.

[15] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, vol. 3, pp. 71–86, 1991.

[16] K. Etemad and R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images," *Journal of Optical Society of America A*, vol. 14, no. 8, pp. 1724–1733, 1997.

[17] Y. Gao and Y. Qi, "Robust Visual Similarity Retrieval in Single Model Face Databases," *Pattern Recognition*, vol. 38, no. 7, pp. 1009–1020, Jul. 2005.

[18] A. Albiol, D. Monzo, A. Martin, J. Sastre, and A. Albiol, "Face Recognition Using HOG-EBGM," *Pattern Recognition Letters*, vol. 29, pp. 1537–1543, 2008.

[19] Z. Cao, Q. Yin, X. Tang, and J. Sun, "Face Recognition with Learning-Based Descriptor," in *IEEE Intl. Conference on Computer Vision and Pattern Recognition*, 2010, pp. 2707 –2714.

[20] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Recognition with Local Binary Patterns," in *European Conference on Computer Vision*, ser. 35, 2004, pp. 469–481.

[21] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition," in *IEEE Intl. Conference on Computer Vision*, 2005, pp. 786–791.

[22] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and Simile Classifiers for Face Verification," in *IEEE Intl. Conference on Computer Vision*, 2009.

[23] X. Tan and B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions," in *IEEE International Conference on Automatic Face and Gesture Recognition*, 2007, pp. 168–182.

[24] J. Holappa, T. Ahonen, and M. Pietikainen, "An Optimized Illumination Normalization Method for Face Recognition," in *IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2008.

[25] J. Wang, J. You, Q. Li, and Y. Xu, "Orthogonal Discriminant Vector for Face Recognition Across Pose," *Pattern Recognition*, vol. 45, no. 12, pp. 4069–4079, 2012.

[26] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 210–227, 2009.

[27] Q. Yuan, A. Thangali, and S. Sclaroff, "Face Identification by a Cascade of Rejection Classifiers," in *IEEE Computer Vision and Pattern Recognition Workshops*, 2005, pp. 152–159.

[28] Z. Zeng, T. Fang, S. Shah, and I. Kakadiaris, "Local Feature Hashing for Face Recognition," in *IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2009.

[29] J. Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade, and B.-L. Lu, "Person-Specific SIFT Features for Face Recognition," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007.

[30] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *IEEE Intl. Conference on Computer Vision and Pattern Recognition*, 2005, pp. 886–893.

[31] T. Ojala, M. Pietikäinen, and T. Maenpää, "Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[32] J. G. Daugman, "Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters," *Journal of the Optical Society of America A*, vol. 2, pp. 1160–1169, 1985.

[33] B. Zhang, S. Shan, X. Chen, and W. Gao, "Histogram of Gabor Phase Patterns (HGPP): A Novel Object Representation Approach for Face Recognition," *IEEE Trans. on Image Processing*, vol. 16, pp. 57–68, 2007.

[34] B. Yao, H. Ai, and S. Lao, "Person-Specific Face Recognition in Unconstrained Environments: A Combination of Offline and Online Learning," in *IEEE International Conference on Automatic Face and Gesture Recognition*, 2008, pp. 1–8.

[35] H. Zhou and A. H. Sadka, "Combining Perceptual Features With Diffusion Distance for Face Recognition," *IEEE Intl. Conference on Systems, Man and Cybernetics*, vol. 41, no. 5, pp. 577–588, Sep. 2011.

[36] M. Vasilescu and D. Terzopoulos, "Multilinear Image Analysis for Facial Recognition," in *IAPR International Conference on Pattern Recognition*, 2002, pp. 511–514.

[37] B. Moghaddam, "Bayesian Face Recognition," *Pattern Recognition*, vol. 13, no. 11, pp. 1771–1782, 2000.

[38] R. Rosipal and N. Kramer, "Overview and Recent Advances in Partial Least Squares," *Lecture Notes in Computer Science*, vol. 3940, pp. 34–51, 2006.

[39] R. Rifkin and A. Klautau, "In Defence of One-vs-All Classification," *Journal of Machine Learning Research*, vol. 5, pp. 101–141, 2004.

[40] S. hyeun Park and J. Fürnkranz, "Efficient Pairwise Classification," in *European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2007, pp. 658–665.

[41] W. R. Schwartz, "Scalable People Re-Identification Based on a One-Against-Some Classification Scheme," in *IEEE Intl. Conference on Image Processing*, 2012, (The preprint version was obtained with the author).

[42] W. R. Schwartz, H. Guo, and L. S. Davis, "A Robust and Scalable Approach to Face Identification," in *European Conference on Computer Vision*, ser. Lecture Notes in Computer Science, vol. 6316, 2010, pp. 476–489.

[43] P. Shih and C. Liu, "Evolving Effective Color Features for Improving FRGC Baseline Performance," in *IEEE Computer Vision and Pattern Recognition Workshops*, 2005, pp. 156–163.

[44] F. D. la Torre, R. Gross, S. Baker, and V. Kumar, "Representational Oriented Component Analysis (ROCA) for Face Recognition with One Sample Image per Training Class," in *IEEE Intl. Conference on Computer Vision and Pattern Recognition*, vol. 2, 2005, pp. 266 – 273.