The Symbiotic Relationship of Parts and Monolithic Face Representations in Verification

Simon Lucey

Advanced Multimedia Processing Laboratory, Department of Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh PA 15213, USA
slucey@ieee.org

Abstract

Techniques that treat the face holistically as a vector of pixel values, which we refer to as a monolithic representation, are still widely considered state of the art for the task of face verification in literature. Recently good performance has be attained in the task of face verification, using Gaussian mixture models (GMMs), via estimating a parts (i.e. image patch) face model; where the shape (i.e. spatial) information is largely ignored. In this paper we postulate that the characteristics current algorithms employ for verifying a face using monolithic and parts representations differ and are in many ways symbiotic; lending themselves to synergetic combination and improved verification performance. Results are presented on the BANCA database that demonstrate excellent verification performance in the presence of many common real world variabilities (e.g. camera degradation, minor pose variabilities, some changes in background and lighting).

1. Introduction

From its inception by Sirovich and Kirby [1] and subsequent *Eigenface* work by Turk and Pentland [2] the monolithic representation of the face has been considered state of the art; even in the most recent open evaluations (e.g. FERET 2000 [3], AVBPA 2003 [4]) monolithic representations have obtained the best performance. We use the term *monolithic* in this paper to describe the holistic vectorized representation of the face based purely on pixel values within an image array.

Many researchers have attempted to venture away from the canonical monolithic representation. Previous work by Brunelli et. al [5], Moghaddam et. al [6] and Wiskott et. al [7] cited good recognition performance by representing the face as a set of salient parts/regions (eg. eyes, nose, mouth). Images in the gallery set were used to create modular templates for comparison with salient regions of the probe images. Both Brunelli and Moghaddam noted superior performance by analyzing the image in a modular manner, rather than holistically as long as the salient regions had been localized to a satisfactory accuracy. Wiskott employed a technique of face recognition that obtained a similarity measure from the Gabor filter response at each node. The nodes of the grid normally, but not necessary, correspond to so called fiducial points. Martínez [8] demonstrated a technique to model the uncertainty associated with the localization of these salient regions during the estimation of the modular templates. However, all these techniques essentially compare "points" (i.e. the distance from a probe's eye image to a eye template) not distributions.

Recently work has been conducted [9, 10] demonstrating that good verification performance can be attained by relaxing many of the spatial constraints in the canonical monolithic face representation. The technique departs from the traditional idea of comparing points on the face, placing greater emphasis on comparing distribution of the parts themselves (i.e. the position of the parts has no bearing on the final similarity measure). Through this technique very good verification performance can be attained from collapsing spatial constraints to generate a distribution of parts.

In this paper we attempt to demonstrate that monolithic and parts face representations are in many ways symbiotic, and can be combined in a synergetic manner leading to improved face verification. We demonstrate that the symbiotic relationship is useful under matched and mismatched gallery and probe conditions. We employ the relatively new BANCA [11] database to demonstrate this ability as it contains facial variabilities that are encountered in real world environments (e.g. camera degradation, minor pose variabilities, some changes in background and lighting).

2. Monolithic representation

In a monolithic representation one attempts to preserve all dependencies in the pixel representation of the face, treating the face as a vector of pixel values. In this paper we will be restricting ourselves to monolithic techniques based

around a nearest neighbor classifier (i.e. the comparison of individual probe images with gallery images, rather than the generation of a single template/model from the ensemble of gallery images) as this technique has been demonstrated to work well when only a few gallery images of a client are available [3, 4]. Most of these techniques have largely differed from one another in a) the subspace chosen to represent the vector [6, 12, 13, 14] (e.g. original pixel space, principal component analysis (PCA), linear discriminant analysis (LDA), independent component analysis (ICA), etc.), and b) the metric of similarity [6, 12, 13, 14] (e.g. Euclidean, Mahalanobis, cosine, Bayesian, support vector machine). Other monolithic techniques that generate client specific templates (e.g. client specific support vector machines (SVMs) [15], multilayer perceptron layers (MLPs) [15], correlation filters [16]) are outside the scope of this paper.

Recently Navarrete and Solar [12, 13] conducted a study into the respective merits of common subspaces and similarity measures used in nearest neighbor monolithic face recognition. A similar study was undertaken by Sadeghi et. al [14] on the BANCA database for the task of face verification. In both studies they found good performance could be obtained using differential techniques, first investigated by Moghaddam and Pentland [6], where a similarity measure, commonly based on Bayesian or SVM classifiers, is estimated from intra-personal and extra-personal differences of subjects in a development gallery. Differential similarity measures obtained using a Bayesian classifier [13] exhibited good performance when a dual PCA space was found for the intra-personal and extra-personal classes. When employing a SVM classifier [13, 14] good performance was attained when the images were left in their original pixel space. Solar stated that similar performance can be achieved using either a Bayesian or SVM similarity measure. A major drawback in all differential similarity measure techniques is their data hungry nature; as they require the estimation of intra-personal and extra-personal observation classes.

Equivalent and sometimes superior performance was also attained in both studies by Solar and Sadeghi by using a simple cosine similarity measure in LDA space. A major advantage of this measure, over differential measures, is that the cosine similarity measure is data independent. Due to its simplicity and good performance this will be the monolithic face verification algorithm used in this paper.

2.1. Eigen- and Fisher-faces

In this paper we will be evaluating two monolithic techniques. The first, which will be referred to as **PCA-MAH**, is the baseline Eigenface [2] technique which employs PCA to generate a subspace preserving the 50 most energy preserving modes. The Mahalanobis distance is then employed to gain a measure of similarity between the two observation vectors $\mathbf{a}$ and $\mathbf{b}$,

$$d_{MAH} = (\mathbf{a} - \mathbf{b})'\mathbf{D}^{-1}(\mathbf{a} - \mathbf{b}) \quad (1)$$

where $\mathbf{D}$ is a diagonal matrix of eigenvalues found from the PCA process. The second technique which we shall refer to as **LDA-COS**, is a variant on the Fisherface [17] technique which employs LDA to generate a subspace preserving the 50 most discriminant modes. As suggested by [12, 13, 14] good performance can be attained if we employ the cosine distance to gain a measure of similarity,

$$d_{COS} = \frac{\mathbf{a}'\mathbf{b}}{||\mathbf{a}||||\mathbf{b}||} \quad (2)$$

3. Parts representation

Recent work [9, 10] has demonstrated that Gaussian mixture models (GMMs) can be used to great success for the task of face recognition. In this type of approach the face is represented as a distribution (of many points), as opposed to the canonical single observation representation of the face commonly employed by monolithic representations. This "many point" representation can be attained by viewing the face as being composed of both *parts* and *shape* [18]. The *parts* are image patches containing information about the face in a local region. The *shape* component provides information describing where these patches are located globally within the face. By collapsing some of the *shape* structure of a face, accurate distributions can be estimated that generalize well to most permutations of faces whilst still providing distinction between faces.

Learning the face as a distribution (i.e. many observations), as opposed to a single observation, has many appealing properties for face classification tasks. First, the many observations (representing the face) can exist in a low dimensional space circumventing problems associated with the "curse of dimensionality" [19] when training a classifier with high dimensional observations. Second, by representing a face with many observation points one naturally has more observations (of a lower dimensionality) to aid in the estimation of a classifier. Through the use of GMMs to model the face distribution, it has been shown that good verification performance can be attained by throwing away most *shape* information. We refer to this type of face model as a full shape collapse GMM (**FSC-GMM**).

A major problem in applying a parts philosophy to face verification is the typically small gallery set of images available for a subject. The lack of training observations drastically affects the ability to estimate conditional distributions that are generalized to differing permutations of a subject's face yet still contain enough complexity to discriminate between subjects. A technique which we refer to as rele-

vance adaptation (RA) [9], based on Bayesian learning, is employed that is able to produce such distribution models. These distributions, specifically GMMs, provide a drastic improvement over techniques that do not employ adaptation.

3.1. Parts feature representation

An initial investigation into what features are most effective for the *parts* representation of frontal face image was conducted by Sanderson et. al [10] for the task of face verification. Sanderson's work is pertinent to our work as it was one of the first investigations for *parts* based face verification using GMMs; albeit using a maximum likelihood (ML) criterion. In this work a modified form of the 2D discrete cosine transform (2D-DCT) was recommended, in comparison to other representations like 2D-Gabor features, as an ideal way to gain a compact *parts* representation that provided good distinction between the faces of subjects and fast feature computation. A depiction of the feature extraction process can be seen in Figure 1.

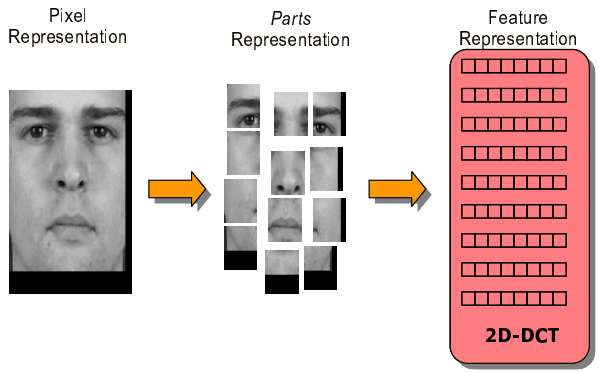

Figure 1: Graphical depiction of the parts and feature representations of a face. Note: even though overlapping blocks are not depicted in practice the overlapping of blocks leads to greater performance.

Additional information about the generation of the feature representations can be obtained from [9, 10].

3.2. Gaussian mixture models

A GMM models the probability distribution of a d dimensional statistical variable $\mathbf{o}$ as the sum of M multivariate Gaussian functions,

$$f(\mathbf{o}|\boldsymbol{\lambda}) = \sum_{m=1}^{M} w_m \mathcal{N}(\mathbf{o}; \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) \tag{3}$$

where $\mathcal{N}(\mathbf{o}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the evaluation of a normal distribution for observation $\mathbf{o}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The weighting of each mixture component is denoted by w_m and must sum to unity across all mixture components. The parameters of the model $\boldsymbol{\lambda} = \{w_m, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m\}_{m=1}^{M}$ can be estimated using the Expectation Maximization (EM) algorithm [20] based on either a maximum a posteriori (MAP) or ML criterion. In the ML case, K-means clustering [19] was used to provide initial estimates of these parameters. In our work the covariance matrices in $\boldsymbol{\lambda}$ are assumed to be diagonal such that $\boldsymbol{\Sigma} = diag\{\boldsymbol{\sigma}^2\}$, as substantial benefit can be attained by reducing the number of parameters needing to be estimated.

To evaluate a sequence of observations, generated from a claimant's probe image, we obtain the average log-likelihood,

$$\mathcal{L}(\mathbf{O}|\boldsymbol{\lambda}_c) = \frac{1}{R} \sum_{r=1}^{R} \log f(\mathbf{o}_r|\boldsymbol{\lambda}_c) \tag{4}$$

Given the average log-likelihood, for the client and world models, one can then calculate the log-likelihood ratio,

$$\Lambda(\mathbf{O}) = \mathcal{L}(\mathbf{O}|\boldsymbol{\lambda}_c) - \mathcal{L}(\mathbf{O}|\boldsymbol{\lambda}_w) \tag{5}$$

For our work we found good performance could be attained if we employed GMMs with 512 mixture components.

3.3. Relevance adaptation

Given a world model $\boldsymbol{\lambda}_w = \{w_{w_m}, \boldsymbol{\mu}_{w_m}, \boldsymbol{\Sigma}_{w_m}\}_{m=1}^{M}$ and training observations from a single client, $\mathbf{O} = [\mathbf{o}_1, \ldots, \mathbf{o}_R]$, using the iterative EM-algorithm one can obtain update equations that incorporate the a priori knowledge in the world model, to maximize the parametric representation of a GMM. A world model is simply a single model trained from a large number of subject faces representative of the population of subject faces expected during verification, and usually has been estimated from a training set independent of the clients to be adapted. This world model is typically trained using the ML criterion (i.e. no informative prior). We refer to the adaptation of the world model $\boldsymbol{\lambda}_w$ to produce a client model $\boldsymbol{\lambda}_c$ as relevance adaptation (RA). For RA this results in the following update equations,

$$w_{c_m} = \left[(1 - \alpha_m^w) w_{w_m} + \alpha_m^w \frac{\sum_{r=1}^{R} \gamma_m(\mathbf{o}_r)}{\sum_{m=1}^{M} \sum_{r=1}^{R} \gamma_m(\mathbf{o}_r)} \right] \beta \tag{6}$$

$$\boldsymbol{\mu}_{c_m} = (1 - \alpha_m^{\mu}) \boldsymbol{\mu}_{w_m} + \alpha_m^{\mu} \frac{\sum_{r=1}^{R} \gamma_m(\mathbf{o}_r) \mathbf{o}_r}{\sum_{r=1}^{R} \gamma_m(\mathbf{o}_r)} \tag{7}$$

$$\boldsymbol{\sigma}_{c_m}^2 = (1 - \alpha_m^{\sigma}) \left(\boldsymbol{\sigma}_{w_m}^2 + \boldsymbol{\mu}_{w_m}^2 \right) + \alpha_m^{\sigma} \frac{\sum_{r=1}^{R} \gamma_m(\mathbf{o}_r) \mathbf{o}_r}{\sum_{r=1}^{R} \gamma_m(\mathbf{o}_r)} - \boldsymbol{\mu}_{c_m}^2 \tag{8}$$

where $\gamma_m(\mathbf{o})$ is the occupation probability for mixture m and α_m^ρ is a weight used to tune the relative importance of the prior and is calculated via a relevance factor τ^ρ in,

$$\alpha_m^\rho = \frac{\sum_{r=1}^{R} \gamma_m(\mathbf{o}_r)}{\tau^\rho + \sum_{r=1}^{R} \gamma_m(\mathbf{o}_r)} \tag{9}$$

Different relevance factors can be estimated for the weights, means and variances respectively (i.e. $\rho \in \{w, \mu, \sigma\}$). We have found effective performance can be attained by using a single relevance factor ($\tau = \tau^w = \tau^\mu = \tau^\sigma$). Additional information on RA can be found in [9].

4. BANCA database and evaluation

The English portion of the BANCA database was employed for these experiments containing 52 subjects; evenly divided into two sets $[g1, g2]$ of 26 as per the BANCA protocol [11]. Inside those sets there are an equal number of sexes (i.e. male=13, female=13). The $g1$ and $g2$ sets are used for the development and evaluation sets in our experiments. The development set is used to obtain any data-dependent aspects of the verification system (e.g. thresholds etc.). The evaluation set is where the performance rates for the verification system are obtained.

If the $g1$ set is used as the development set then the $g2$ set is used for the evaluation set; and vice versa. This is done to avoid any methodological flaw, as it is essential that the development set is composed of a distinct subject population as the one of the evaluation set. We will report results in this paper using both the $g1$ and $g2$ sets so as to gain a gauge for the statistical significance of our results. For the experiments in this correspondence we have employed images, examples of which can be seen in Figure 2, taken under three conditions,

Controlled: using a high end digital camera with controlled lighting and homogeneous background.

Degraded: using a low end (i.e. web) digital camera with uncontrolled indoor lighting and inhomogeneous background.

Adverse: using a high end digital camera with a combination of uncontrolled indoor and outdoor lighting and inhomogeneous background.

For our experiments the faces are cropped and normalized for scale and rotation so that their eyes line up in the same position, resulting in a 55×51 gray scale image. Examples can be seen in Figure 3. A lighting normalization method, proposed by Gross and Brajovic [21], is then employed that compensates for illumination variation in the face images by making an estimate of the illumination field.

Several protocols [11] have been devised for the BANCA database. We have restricted our investigation to protocols where the gallery images are always taken under *controlled* conditions, as this is a highly likely scenario in practice. The four protocols, listed below, we employ only differ in the probe images they use.

Mc: "matched controlled" protocol, probe taken under *controlled* conditions.

Ud: "unmatched degraded" protocol, probe taken under *degraded* conditions.

Ua: "unmatched adverse" protocol, probe taken under *adverse* conditions.

P: "pooled" protocol, probe taken under *controlled*, *degraded* and *adverse* conditions.

Further specifications of the protocols can be found in [11].

Figure 2: Example of images of a subject taken across (left to right) *controlled*, *degraded* and *adverse* conditions.

Figure 3: Example of cropped images of a subject taken across (left to right) *controlled*, *degraded* and *adverse* conditions.

Verification is performed by accepting a claimant when $ms \geq Th$ and rejecting him/her when $ms < Th$, where ms is a match score (i.e. a similarity measure d or log-likelihood ratio Λ) from the classifier and Th is a given threshold. Verification performance is evaluated using two measures; being false rejection rate (FRR), where a true client is rejected against their own claim, and false acceptance rate (FAR), where an impostor is accepted as the falsely claimed client. The FAR and FRR measures increase or decrease in contrast to each other based on the threshold Th. A common technique for evaluating overall performance is the Half Total Error Rate (HTER) which is defined as,

$$\text{HTER} = \left[\text{FAR}(Th^*) + \text{FAR}(Th^*)\right]/2 \tag{10}$$

where the threshold Th^* is found a priori on the development set, and is selected to obtain Equal Error Rate (EER) performance (i.e. FAR = FRR).

5. A question of resolution

Lemieux and Parizeau [22] recently conducted a study on parameters that affect eigenface recognition. In their study they found the eigenface recognition engine relied mostly on low frequency information for recognition. Lemieux and Parizeau reported that no loss in recognition rate was seen by downsampling (i.e. similar to blurring) the cropped and normalized face images from 200×220 to 12×13 pixel arrays. We achieved similar results on the Eigenface-based **LDA-COS** algorithm, whose results are depicted in Figure 4(a), where we test the performance of the algorithm after the probe images undergo blurring and sharpening operations. The blurring and sharpening operations employed a Gaussian kernel. The standard deviation of the kernel was chosen for the blurring operation so the blurred Mc probe images had a similar resolution to the normal Ud probe images. A similar selection criteria was employed for the sharpening operation so the sharpened Ud probe images had a similar resolution to the normal Mc probe images. We restricted ourselves to a comparison of the Mc and Ud protocols, as the main variation between the Mc and Ud protocols, apart from lighting, was predominantly the resolution of the images.

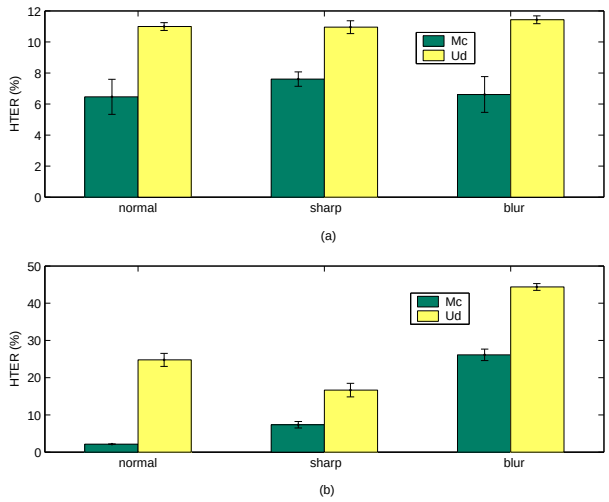

Figure 4: The effect, on verification performance, of sharpening and blurring probe images employing the (a) **LDA-COS** and (b) **FSC-GMM** algorithms. Note: HTER results presented are the average across $g1$ and $g2$ evaluations sets, with black lines denoting the deviation.

When a similar investigation is undertaken with the **FSC-GMM** algorithm, as depicted in Figure 4(b), the blurring and sharpening operations have a clear effect on performance. Initially one can see that the **FSC-GMM** algorithm receives far better performance than the **LDA-COS** algorithm on the normal Mc protocol. Performing a similar comparison on the Ud protocol we obtain the reverse situation where the **LDA-COS** algorithm receives far better performance than the **FSC-GMM** algorithm. Assuming that the Ud protocol probe images are missing much of the high frequency information contained in the Mc protocol probe images, this gives us an initial indication that the **FSC-GMM** is heavily dependent on high frequency information.

When the sharpening operation is performed, the **FSC-GMM** algorithm performance on the Ud protocol improves substantially, but little change is seen in the **LDA-COS** algorithm; giving an additional indication that the **FSC-GMM** algorithm is heavily dependent on high frequency information. We should also note that performance on the **FSC-GMM** algorithm for the Mc protocol degrades when the sharpening operation is performed. One hypothesis for this result is that the sharpening operation sharpens areas of the face that normally contain only low frequency information, hindering verification performance.

Finally we can see that the blurring operation has a definite effect on the **FSC-GMM** algorithm on protocol Mc, attaining similar verification performance to **FSC-GMM**'s Ud protocol in normal conditions. The blurring operation further degrades performance on the blurred Ud protocol. This final result gives additional support to our postulate that the **FSC-GMM** algorithm, employing a parts representation, requires high frequency information to gain good verification performance; where as the the **LDA-COS** can perform well with only low frequency information.

6. Combining representations

From Section 5 one can see that an argument can be made about the characteristics used for verification by the **LDA-COS** and **FSC-GMM** algorithms. Heuristically we can also claim that the monolithic representation employed in the **LDA-COS** algorithm preserves global, albeit low frequency, dependencies across the entire face. The **FSC-GMM** algorithm employing the parts representation, although only preserving local dependencies in the face, preserves much of the high frequency detail. Using this very preliminary evidence we shall now use the rest of this paper to investigate the effect of fusing these two representations through the combination of the **LDA-COS** and **FSC-GMM** algorithms.

We shall refer to this technique as the **COMB** algorithm, where we employ the *sum* rule for combining the match scores of multiple classifiers. Kittler et. al [23] demonstrated that the *sum* rule can obtain good performance in

classifier combination, when the two classifiers are diverse and produce match scores approximately representative of their true a posteriori probabilities. Another advantage of the *sum* rule is that it is completely data independent; as it requires no tuning set to effectively fuse match scores. The **COMB** match score is generate by,

$$ms = \text{logsig}(d_{COS}) + \text{logsig}(\Lambda) \tag{11}$$

where the logsig operation is used to try and make the match scores obtained from the **LDA-COS** and **FSC-GMM** algorithm more representative of their true a posteriori probabilities. The logsig operation can be defined as,

$$\text{logsig}(a) = \frac{1}{1 + \exp(-a)} \tag{12}$$

The employment of the logsig operation results in the synergetic combination of match scores from the **LDA-COS** and **FSC-GMM** algorithms.

7. Results

Table 1 contains a summary of results for each algorithm discussed in this paper across the Mc, Ud, Ua and P protocols. For completeness we have also included results from [14] using their two best algorithms on the BANCA database; which are both based on monolithic representations. The first algorithm, which we shall refer to as **LDA-NC**, is basically their implementation of our own **LDA-COS** algorithm which employs a LDA subspace and cosine distance. The second algorithm, which we refer to as **ORG-SVM**, employs the original pixel representation of the image in conjunction with a SVM classifier to calculate a differential similarity measure.

Algorithm/Protocol	Mc	Ud	Ua	P
LDA-NC [14]	4.93	15.99	20.24	14.79
ORG-SVM [14]	5.43	25.43	30.11	20.33
PCA-MAH	10.2	17.84	26.63	21.57
LDA-COS	6.46	10.99	20.39	14.96
FSC-GMM	2.14	24.78	17.06	21.97
COMB	**1.42**	**9.65**	**16.51**	**13.71**

Table 1: Comparison of algorithms, in Half Total Error Rate (HTER), on the BANCA database. Note: for a fair comparison with results in [14] performance was averaged across the $g1$ and $g2$ evaluation sets.

It is interesting to immediately note the performance improvement of the **FSC-GMM** algorithm, in Table 1, over all other algorithms using a monolithic representation for protocol Mc. This indicates that even in matched conditions, where the probe and gallery images stem from similar recording conditions, parts based representations are of some benefit due to their natural ability to generalize to unseen images and preserve high frequency detail. When we include additional information from the monolithic based **LDA-COS** algorithm, resulting in the **COMB** algorithm, one can see that this additional information is fused synergetically resulting in increased performance.

The parts based representations do not fare so well for the Ud protocol, which can largely be explained by the blurred nature of the Ud protocol's probe images as investigated in Section 5. They did however outperform all monolithic representation algorithms on the Ua protocol. One hypothesis for this performance, can be attributed to the natural pose variation present in the Ua protocol's probe images. Trying to verify a face in parts, instead of using a monolithic representation, allows one to freely match areas of the face that are not as perturbed by pose.

Inspecting Table 1 one can note the catastrophic performance of the **SVM-ORG** algorithm, employing a data-dependent differential similarity measure, across the Ud, Ua and P protocols. This is an interesting result, indicating that differential data dependent similarity measures are especially prone to changes in the data. The **LDA-NC** and **LDA-COS** algorithms achieve similar performance across all protocols; as expected. A difference did occur for the Ud protocol with substantially better performance being witnessed for the **LDA-COS** algorithm. This could be partially be attributed to the lighting normalization technique [21], which was not employed in [14]. The proposed **COMB** algorithm achieves superior performance to all other algorithms investigated, across all protocols. The superior performance of the **COMB** algorithm seems to be limited by the best performances of the **LDA-COS** and **FSC-GMM** algorithms respectively.

8. Summary and Conclusions

In this paper we have taken some initial steps into understanding what type of representations are appropriate for the task of face verification. We have investigated the benefits and drawbacks of employing a monolithic and parts representation for face verification through the **LDA-COS** and **FSC-GMM** algorithms respectively. We have postulated, based on empirical evidence and heuristics, that the two representations are symbiotic; in the sense that they try and verify using substantially different characteristics of the face. The **LDA-COS** algorithm attempts to preserve all low-frequency dependencies, global and local, in the face pixel representation. The **FSC-GMM** algorithm attempts to only preserve local dependencies but is then able to preserve far more high-frequency information. We have seen in our results that both representations are equally valid and important; and that through synergetic combination of these two algorithms superior verification performance can be at-

tained than has been seen in literature previously [14].

Acknowledgments

This research is supported in part by the Technology Support Working Group (TSWG).

References

[1] L. Sirovich and M. Kirby, "A low-dimensional procedure for the characterization of human faces," *J. Opt. Soc. Amer. A.*, vol. 4, no. 3, pp. 519–524, 1987.

[2] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, 1991.

[3] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face recognition algorithms," *IEEE Trans. PAMI*, vol. 22, pp. 1090–1104, October 2000.

[4] K. Messer, J. Kittler, M. Sadeghi, and et. al, "Face verification competition on the XM2VTS database," in *AVPBA*, pp. 964–974, 2003.

[5] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. PAMI*, vol. 15, no. 10, pp. 1042–1052, 1993.

[6] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object recognition," *IEEE Trans. PAMI*, vol. 19, no. 7, pp. 696–710, 1997.

[7] L. Wiskott, J. M. Fellous, N. Kruger, and C. Von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. PAMI*, vol. 19, pp. 775–779, July 1997.

[8] A. M. Martínez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class," *IEEE Trans. PAMI*, vol. 24, no. 6, pp. 748–763, 2002.

[9] A. author, "Anonymous author's previous related paper,"

[10] C. Sanderson and K. Paliwal, "Fast features for face authentication under illumination direction changes," *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2409–2419, 2003.

[11] E. Bailly-Bailliere, S. Bengio, K. Messer, and et. al, "The BANCA database and evaluation protocol," in *AVBPA*, pp. 625–638, 2003.

[12] P. Navarrete and J. Ruiz-del-Solar, "Analysis and comparison of eigenspace-based face recognition approaches," *Int. Journal of Pattern Recognition and Artificial Intelligence*, vol. 16, no. 7, pp. 817–830, 2002.

[13] J. Ruiz-del-Solar and P. Navarrete, "Towards a generalized eigenspace-based face recognition framework," in *4th Int. Workshop on Statistical Techniques in Pattern Recognition*, (Windsor, Canada), August 2002.

[14] M. Sadeghi, J. Kittler, A. Kostin, and K. Messer, "A comparative study of automatic face verification algorithms on the BANCA database," in *AVBPA*, pp. 35–43, 2003.

[15] F. Cardinaux and S. Marcel, "Face verification using MLP and SVM," in *XI Journees NeuroSciences et Sciences pour l'Ingenieur (NSI 2002)*, no. 21, (La Londe Les Maures, France), 15-19 September 2002.

[16] M. Savvides and B. Vijaya Kumar, "Quad phase minimum average correlation energy filters for reduced memory illumination tolerant face authentication," in *AVBPA*, pp. 19–26, 2003.

[17] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. PAMI*, vol. 19, no. 7, pp. 711–720, 1997.

[18] M. Weber, M. Welling, and P. Perona, "Towards automatic discovery of object categories," in *CVPR*, pp. 101–108, June 2000.

[19] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York, NY, USA: John Wiley and Sons, Inc., 2nd ed., 2001.

[20] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Royal Statistical Society*, vol. 39, pp. 1–38, 1977.

[21] R. Gross and V. Brajovic, "An image preprocessing algorithm for illumination invariant face recognition," in *AVBPA*, (Surrey, U.K.), pp. 10–17, 2003.

[22] A. Lemieux and M. Parizeau, "Experiments on Eigenfaces Robustness," in *ICPR'02*, 2002.

[23] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. PAMI*, vol. 20, pp. 226–239, March 1998.