QUERY-DRIVEN LOCALLY ADAPTIVE FISHER FACES AND EXPERT-MODEL FOR FACE RECOGNITION

Yun Fu [1*], *Junsong Yuan* [2], *Zhu Li* [3], *Thomas S. Huang* [1], *and Ying Wu* [2]

[1]Beckman Institute,University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[2]Dept. of EECS, Northwestern University, Evanston, IL 60208, USA
[3]Multimedia Research Lab (MRL), Motorola Labs, Schaumburg, IL 60196, USA

ABSTRACT

We present a novel expert-model of Query-Driven Locally Adaptive (QDLA) Fisher faces for robust face recognition. For each query face, the proposed method first fits local Fisher models with different appearances. A hybrid expert model then integrates these local models and combines the classification results based on the estimated error rate for each local model. This approach addresses the large size recognition problem, where many local variations can not be adequately handled by a single global model in a single appearance space. To speed up the query process, Locality Sensitive Hash(LSH) is applied for fast nearest neighbor search. Experiments demonstrate the approach to be effective, robust, and fast for large size, multi-class, and multi-variance data sets.

Index Terms— Expert model, face recognition, query, nearest neighbor, Fisher face, locality sensitive hash

1. INTRODUCTION

The difficulty of face recognition [4] stems from the difficulty in statistical modeling of face images under pose, scale, expression, occlusion, and illumination variations. Linear models [5] and Gaussian-assumption [6] based methods have been found to be inadequate in handling these issues under the practical circumstances. Recent advances in machine learning and pattern recognition has focused on non-linear methods (e.g. kernel based method [11] or manifold learning) and graph modeling methods (e.g. GE [12] or LEA [2]). By applying a non-linear kernel mapping or a locally preserved graph embedding, it improves the discriminating power of the decision boundaries. However, a serious numerical problem arises because the covariance modeling in kernel method or the neighborhood graph in the graph modeling method is typically of $n \times n$ dimension, where n is the number of labeled training sample. When n is very large, the solution can be unstable and the calculation of the $n \times n$ matrix is computationally impractical, thus more efficient method is desired.

*This research was funded in Part by the U.S. Government VACE program, and in part by the NSF Grant CCF 04-26627. The views and conclusions are those of the authors, not of the US Government or its Agencies.

Furthermore, a single global model that tries to capture all the variations in data has several limitations also. Firstly, to fit the large training data, a kernel with mapping to a much higher dimensional space is required, which sometimes results in a complex decision boundary with poor generalization ability and often involves time-consuming optimization or computation of pair-wise distances. Secondly, for the graph embedded linearization, the larger the data set, the poorer the learning performance, since the large data set defeats the discriminating power of locality preserved graph modeling.

To handle these problems, we developed the framework of an expert-model of Query-Driven Locally Adaptive (QDLA) Fisher faces method that can work with multiple global appearance models for face recognition. The query-driven local adaptation in a single model is achieved as follows: first a local neighborhood of known labeled faces is identified with the querying (unknown) face [13], then a local linear (non-linear) Fisher face model is created for the labeled data in the neighborhood, and the error rate is estimated for this local model. Depending on the number and quality of labeled samples in the neighborhood, a linear (non-linear) model for classification is created, as well as the resulting classification error rate. The expert-model is achieved by building multiple global appearance models according to different facial features and voting the final recognition result with corresponding QDLA models of each global model based on the error rate. The LSH [1, 3] based fast querying strategy is also introduced to deal with the high dimensional NN search problem. We build multiple appearance models of different area and scale of faces to improve the face recognition performance. Considering the locally adaptive classification and error estimation in each model, we combine the recognition results to improve the overall performance.

This paper addresses the problem of numerical difficulty in large size recognition, where many local variations can not be adequately handled by a single global model. By localizing the modeling, the classification error rate estimation is also localized and thus appears to be more robust and flexible for model selection among different model candidates. The proposed expert-model is a general framework for classifica-

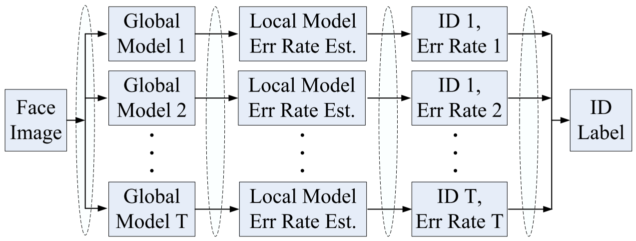

Fig. 1. Framework of expert-model of QDLA-Fisher faces for face recognition.

tion and can be very applicable to large size, multi-class, and multi-variance face recognition data sets.

2. QDLA-FISHER FACES FOR EXPERT-MODEL FACE RECOGNITION

For the face recognition problem, assume we have a known set of face images $\mathcal{F} = \{f_k | k = 1...N\}$ with person ID label $\mathcal{L} = \{l_k | k = 1...M\}$. Now for an unknown face image f, we need to select a label l from $\mathcal{L}$ to identify f. Our QDLA-Fisher algorithm for a single appearance model is as:

- A local neighborhood of known labeled faces f_q is identified with the query (unknown) face f.
- Pick K nearest neighbors in the training data $\mathcal{F}$ for f_q.
- If the K nearest neighbors have the same label l_k, then f is labeled and exit; otherwise, depending on the number and quality of labeled samples in the neighborhood, a local linear (non-linear) Fisher model is created for f_q, along with the local-model classification error rate.
- Classify f with the local model.

The expert-model of QDLA-Fisher faces is summarized as:

- Build multiple appearance face models with variation in area and resolution of faces.
- Run QDLA-Fisher algorithm on each appearance model respectively.
- Vote the final recognition result with corresponding local models of each global model based on the error rate.

The framework of expert-model of query-driven QDLA-Fisher faces for face recognition is shown in Figure 1.

3. FAST NEAREST NEIGHBORS QUERYING

Suppose $\mathcal{S}$ is the training data set that is composed of finite points p (labeled) in the Euclidean space $\mathcal{R}^d$, let $n = |\mathcal{S}|$. In order to build the local Fisher model, the first step is to search for the nearest neighbors of the query face in the feature space. By considering only the local distribution of the high-dimensional data, such local model is more discriminative than the global model. However, the exhaustive linear

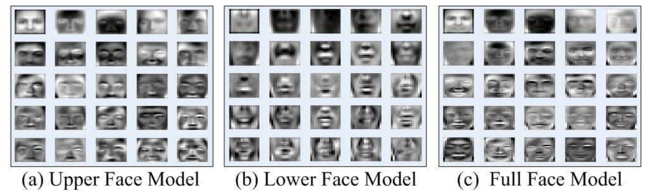

Fig. 2. Single QDLA-Fisher face models with different area and scale. (a) Upper face model (18×16). (b) Lower face model (14×18). (c) Full face model (21×28).

scan method to search for the nearest neighbors (NN) is computationally expensive for large database ($O(n)$). Note for every query face, we need to perform such NN query to build the local model. One speed-up solution is to take advantage of the data set spatial distribution in $\mathcal{R}^d$. From the database point of view, NN search can be treated as a querying problem. Many index structures have been well studied in the database community, such as kd-tree, R-tree, etc, in order to speed up the NN search. Nevertheless, almost all of such index based search methods suffer from the problem of curse of dimensionality. When compared with exhaustive linear scan search, such index-based search method normally can not perform better in the high-dimensional space. We use Locality Sensitive Hash (LSH) to speed up the NN query. The local Fisher model can then be built based on the NNs of the query.

To overcome the curse of dimensionality, the recently proposed method LSH provides a randomized solution for the high-dimensional NN search problem. Instead of searching for exact K-NN, LSH tends to search for the approximate NN which is defined as ϵ-Nearest Neighbor Search (NNS) — For a data set $\mathcal{S}$ which is composed of data points $p \in \mathcal{R}^d$, preprocess $\mathcal{S}$ to efficiently search for the approximate NNs of any given query q. That is, to find $p \in \mathcal{S}$, such that $d(q,p) \leq (1+\epsilon)d(q,\mathcal{S})$, where $d(q,\mathcal{S})$ denotes the distance of q to its closest neighbor in $\mathcal{S}$.

4. QDLA-FISHER MODEL

For global appearance based face models, we take the cropped whole face image as the learning feature. We apply three different global appearance models with known and labeled face data, that is, to have $\{f_k^{(j)} | k = 1...N, j = 1,2,3.\}$ projected from known face data $\mathcal{F}$. An example for the basis functions of upper, lower, and full FERET face models are visualized in Figure 2. The areas can be overlapped which capture the faces in different areas and different resolutions. In the recognition phase, for an unknown face f, and its projection in j^{th} global model space $x^{(j)}$, we apply LSH to identify its local neighborhood with $N(x^{(j)}) = \{f_i^{(j)} | \mathtt{s.t.} \|f_i^{(j)} - x^{(j)}\| < r^{(j)}\}$ and build local fisher models for each global appearance model, i.e. distribution of each class of faces identified in $N(x^{(j)})$.

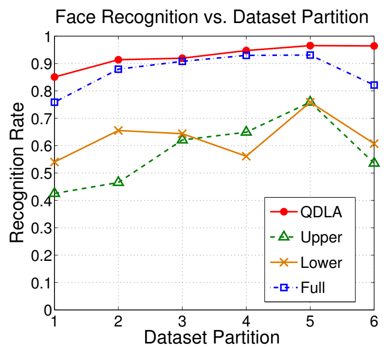

Fig. 3. Average ORL face recognition accuracy for 6 different dataset partition under the same model parameter settings.

The query face is then classified with these local models. Here $r^{(j)}$ denotes the local neighborhood radius. The Fisher model classification error rate e_j is also computed and recorded. After we do face recognition with each local model, the final recognition result is based on the error rate e_j of each local model. In the preferred embodiment, we select the minimum error rate of local model classification result as our final recognition result. If there is a tie in error rate, the preference is given in the order of full face model, upper face model and lower face model. We also define two parameters to tune the local model. One is the default local Fisher model classification error and the other is the minimum local sample ratio. During the query stage, if the ratio between the number of local samples within the range of local neighborhood radius and the training samples is lower than minimum local sample ratio, we consider the local Fisher model is insufficient for training. In this case, we set the local model to be a Fisher model with NN classifier and the error rate e_j to be the default Fisher model classification error.

5. EXPERIMENTAL RESULTS

We use two standard face data sets to demonstrate our algorithm. *1) FERET Database* [7]. The original database, containing 1209 subjects, was released in 2001 and consisted of 14051 greyscale images. The images, in a resolution of 256×384, were taken with head views ranging from frontal to left and right profiles. We select 2550 near frontal face images, crop and resize the images to the size of 21×28 according to the eye locations and pupils distance. *2) ORL Database* [8]. This database contains 40 subjects with 10 grayscale face images for each. The 400 images, in a resolution of 92×112, were taken at different times, varying lighting, facial expressions (open / closed eyes, smiling / not smiling) and accessories (glasses / no glasses), showing all frontal and slight tilt of the head. We crop and resize the images to the size of 21×28 according to the eye locations and pupils distance.

We train each single upper, lower, and full model of QDLA-Fisher with FERET images. The model sizes are defined by 18×16, 14×18, and 21×28, respectively. Figure 2 shows the basis functions of the QDLA-Fisher model created by FERET data. Since the ORL face database has specific labels for each subject and its 10 pose, we divide the data set into 6 partition cases with different training and testing samples (some difficult recognition cases). Table 5 shows the description of the data set partitions for the 6 recognition experiments. The model basis functions and model parameter $r^{(j)}$ are fixed for each single QDLA-Fisher model of case. The average face recognition accuracy for 6 different data set partition under the FERET basis is plotted in Figure 3. We can see that the expert-model of QDLA-Fisher outperforms all the single QDLA-Fisher model for under the 6 data set partitions. The full model of QDLA-Fisher model performs better than the upper and lower model, but worse than the expert-model of QDLA-Fisher. To further demonstrate the generalization and robustness properties of our method, we design the following more specific recognition experiments.

Cases	Sub. ID	Test ID	r^{Upper}	r^{Lower}	r^{Full}
1	6:10, 16:40	3, 7, 8	400	600	800
2	6:10, 16:40	3, 8	400	600	800
3	11:40	5, 8,10	400	600	800
4	11:40	5, 10	400	600	800
5	11:40	10	400	600	800
6	1:20, 31:40	5	400	600	800

Table 1. Description of the data set partitions for Fig. 3.

The recognition performances for ORL database with the expert-model of QDLA-Fisher faces based on FERET basis functions is shown in Table 2. The 10 images of each subject are randomly split with (6 labeled and 4 un-labeled faces) and (8 labeled and 2 un-labeled faces). We totally set up 10 sets of 160 and 80 recognition attempts test for the four cases. We set the local neighborhood radius $r^{(j)}$ to 420, 400, and 860 for single upper, lower, and full models of QDLA-Fisher respectively in the upper two cases of Table 2, and set $r^{(j)}$ to 640, 600, and 1000 in the lower two cases. The default NN classification error is set to 0.005 and the minimum local sample ratio is set to 0.2. The experiments show positive and encouraging results since the recognition and query modules were accurate, fast, and robust. We can observe that the recognition performance of QDLA-Fisher expert-model is better than any single QDLA-Fisher models and the global PCA Eigenface [9] recognition. It also has very low recognition rate variance of the random tests since the localization of models and adaptive multiple modes of classification make the error estimation robust for the results combining. The computation time needed for each recognition is as short as 0.23$\sim$0.24s, shown in Table 2, on a 2.0GHz Pentium CPU and 512MB RAM PC with un-optimized Matlab 6.0 implementation. The NN query is also computationally efficient, costing 0.005$\sim$ 0.02s for each query (depending on the value of radius $r^{(j)}$).

Table 2. Recognition accuracy (%) for ORL database with QDLA-Fisher faces of expert-model based on FERET basis functions. (Upper, Lower, and Full $\longleftrightarrow$ single models of QDLA-Fisher; PCA $\longleftrightarrow$ Eigenface recognition; QDLA $\longleftrightarrow$ expert-model of QDLA-Fisher; Time $\longleftrightarrow$ computation time needed for each recognition attempt; $n \longleftrightarrow$ # of subjects; $m \longleftrightarrow$ # of training images per subject; $t \longleftrightarrow$ # of testing images per subject; $s \longleftrightarrow$ minimum local sample ratio.)

Cases	$n = 30$	$m = 8$	$t = 2$	$s = 0.3$			$n = 30$	$m = 8$	$t = 2$	$s = 0.2$		
Test#.	Upper	Lower	Full	PCA	QDLA	Time(s)	Upper	Lower	Full	PCA	QDLA	Time(s)
1	0.8167	0.6667	0.9333	0.8333	0.9310	0.1325	0.8333	0.6833	0.9333	0.8333	0.9310	0.1187
2	0.8500	0.7000	0.9000	0.8500	0.9286	0.1315	0.8833	0.6833	0.9000	0.8500	0.9643	0.1096
3	0.9000	0.6667	0.9500	0.9333	0.9492	0.1169	0.8833	0.6667	0.9667	0.9333	0.9661	0.1198
4	0.8833	0.5833	0.9500	0.8500	0.9828	0.1253	0.8500	0.5833	0.9500	0.8500	0.9828	0.1125
5	0.8500	0.5333	0.9000	0.8167	0.9298	0.1268	0.8333	0.5167	0.9000	0.8167	0.9298	0.1013
6	0.7833	0.5833	0.9000	0.7667	0.8947	0.1190	0.7833	0.6000	0.9000	0.7667	0.8772	0.1086
7	0.8833	0.5833	0.9500	0.8500	0.9828	0.1219	0.8500	0.5833	0.9500	0.8500	0.9828	0.1013
8	0.8167	0.5500	0.9000	0.8000	0.8966	0.1224	0.8167	0.5500	0.9000	0.8000	0.8966	0.1172
9	0.8500	0.7000	0.9000	0.8500	0.9286	0.1219	0.8833	0.6833	0.9000	0.8500	0.9643	0.0979
10	0.7833	0.6167	0.9333	0.7833	0.9310	0.1237	0.8167	0.5833	0.9167	0.7833	0.9310	0.1096
Mean	0.8417	0.6183	0.9217	0.8333	**0.9355**	0.1242	0.8433	0.6133	0.9217	0.8333	**0.9426**	0.1097
Var.	0.0017	0.0037	0.0006	0.0022	0.0009	0.0000	0.0011	0.0038	0.0007	0.0022	0.0013	0.0001

Cases	$n = 40$	$m = 6$	$t = 4$	$s = 0.2$			$n = 40$	$m = 8$	$t = 2$	$s = 0.2$		
Test#.	Upper	Lower	Full	PCA	QDLA	Time(s)	Upper	Lower	Full	PCA	QDLA	Time(s)
1	0.8438	0.5250	0.8750	0.8375	0.9456	0.2338	0.8375	0.4875	0.8750	0.8000	0.9333	0.2544
2	0.8000	0.5250	0.8688	0.8187	0.9392	0.2363	0.8875	0.6250	0.9125	0.8500	0.9605	0.2597
3	0.8688	0.4750	0.8500	0.7875	0.9007	0.2342	0.8625	0.4875	0.8875	0.7750	0.9221	0.2628
4	0.8562	0.5313	0.8750	0.8000	0.9145	0.2357	0.9125	0.5625	0.9500	0.8625	0.9744	0.2310
5	0.8125	0.5188	0.8250	0.7500	0.8919	0.2300	0.8000	0.5125	0.8250	0.8000	0.8919	0.2327
6	0.8187	0.5313	0.8562	0.8125	0.9067	0.2319	0.9375	0.5500	0.9125	0.8375	0.9733	0.2195
7	0.7937	0.5000	0.8625	0.7500	0.9139	0.2309	0.8625	0.5125	0.9375	0.8250	0.9615	0.2351
8	0.8125	0.5188	0.8250	0.7500	0.8919	0.2290	0.8750	0.5750	0.9500	0.8250	0.9620	0.2293
9	0.9063	0.5500	0.9125	0.8625	0.9608	0.2376	0.8875	0.5500	0.8875	0.8875	0.9342	0.2697
10	0.9063	0.5500	0.8938	0.8063	0.9346	0.2362	0.8125	0.5125	0.8875	0.8000	0.9221	0.2484
Mean	0.8419	0.5225	0.8644	0.7975	**0.9200**	0.2336	0.8675	0.5375	0.9025	0.8263	**0.9435**	0.2443
Var.	0.0017	0.0005	0.0008	0.0015	0.0006	0.0000	0.0018	0.0019	0.0015	0.0012	0.0007	0.0003

6. CONCLUSION

The solution resolves some numerical difficulties and derives better Fisher models by fitting only data within query dependent local neighborhood. An expert model with multiple global appearance models is built for better robustness in recognition. Simulation results demonstrate the effectiveness. We will investigate more advanced local metric modeling and judicious use of multiple models [10] for further researches.

7. REFERENCES

[1] M. Datar, N. Immorlica, P. Indyk, and V. Mirrokni, "Locality-Sensitive Hashing Scheme Based on *p*-Stable Distribution," *Proc. SCG'04*, pp.253-262, 2004.

[2] Y. Fu and T.S. Huang, *Locally Linear Embedded Eigenspace Analysis*, www.ifp.uiuc.edu/~yunfu2/papers/LEA-Yun05.pdf.

[3] A. Gionis, P. Indyk, R. Motwani, "Similarity Search in High Dimensions via Hashing," *Proc. VLDB'99*, pp.518-529, 1999.

[4] S.Z. Li and A.K. Jain, *Handbook of Face Recognition*, Springer-Verlag, New York, 2005.

[5] A.M. Martinez and A.C. Kak, "PCA versus LDA," *IEEE Trans. on PAMI*, vol. 23, no. 2, pp. 228-233, 2001.

[6] B. Moghaddam, T. Jebara, A. Pentland, "Bayesian Face Recognition,"*Pattern Recognition*, vol.33, no.11, pp.1771-1782, 2000.

[7] P.J. Phillips, H. Moon, P.J. Rauss, and S. Rizvi, "The FERET Evaluation Methodology for Face Recognition Algorithms," *IEEE Trans. on PAMI*, vol. 22, no. 10, pp. 1090-1104, 2000.

[8] F.Samaria, A.C.Harter, "Parameterisation of a Stochastic Model for Human Face Identification,"*IEEE WACV*,pp.138-142,1994.

[9] M.A. Turk and A.P. Pentland, "Face Recognition Using Eigenfaces," *IEEE Conf. on CVPR'91*, pp. 586-591, 1991.

[10] R. Xiao, W.J. Li, Y.D. Tian, and X.O. Tang, "Joint Boosting Feature Selection for Robust Face Recognition," *IEEE Conf. on CVPR'06*, pp. 1415-1422, 2006.

[11] M.H. Yang, "Face Recognition Using Kernel Methods," *Proc. of NIPS'01*, 2001.

[12] S.C. Yan, D. Xu, B.Y. Zhang and H.J. Zhang, "Graph Embedding: A General Framework for Dimensionality Reduction," *IEEE Conf. on CVPR'05*, pp. 830-837, 2005.

[13] H. Zhang, A.C. Berg, M. Maire, and J. Malik, "SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition," *IEEE Conf. on CVPR*, pp. 2126-2136, 2006.