Face Recognition Systems Performance Modeling, Prediction, and Improvement

Qiang Ji

Department of Electrical, Computer and Systems Engineering
Rensselaer Polytechnic Institute
jiq@rpi.edu

NIST Biometrics Quality Workshop, November, 2007

Performance Modeling

- The performance of a face recognition system consists of *intrinsic* and *extrinsic* performance
- The intrinsic performance is determined by the intrinsic factors: face recognition algorithms, their parameters, and the gallery images
- The extrinsic performance is determined by extrinsic factors: conditions of face images including resolution, size, illumination, pose, occlusion, etc.

Research Goals

- Model the intrinsic performance for both online and offline intrinsic performance improvement using only gallery data

- Predict online the extrinsic performance on query images using minimal training data

Face Recognition Systems

- A face recognition system:
  - Gallery data set
  - A recognition algorithm and its parameters
  - Query images

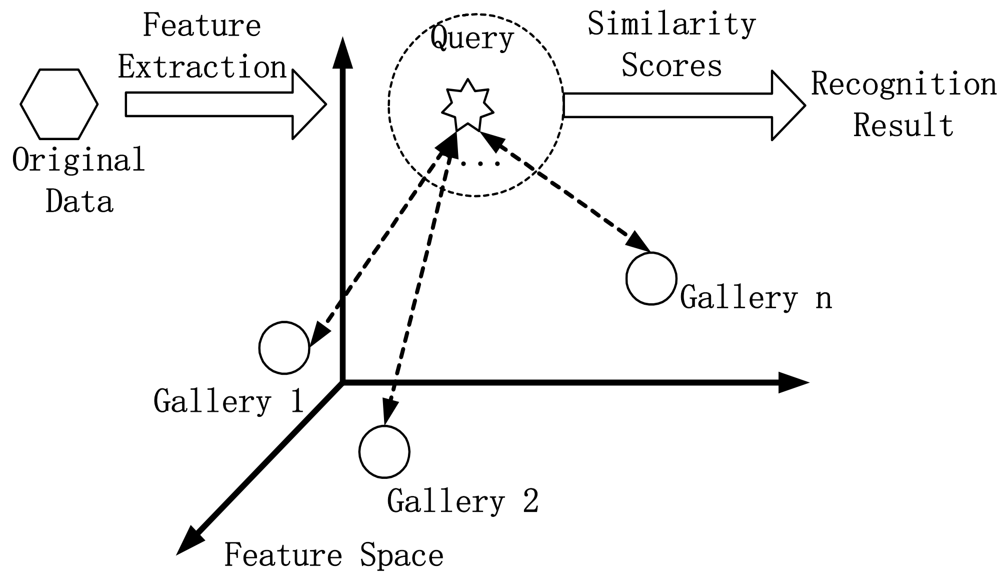


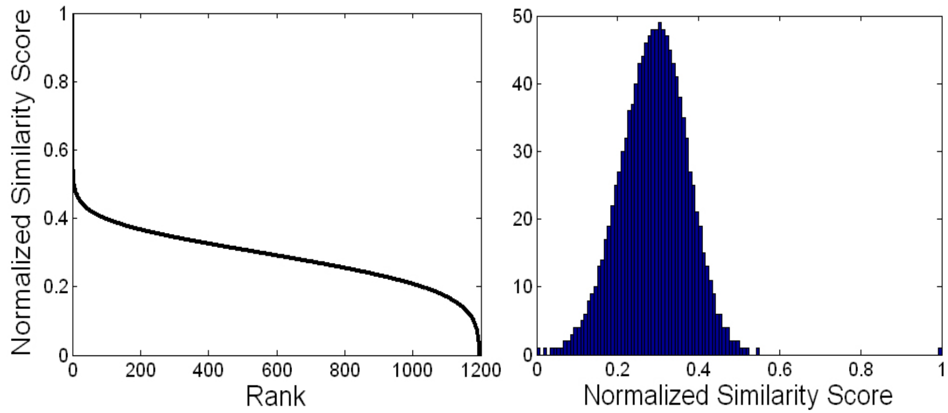
Illustration of face recognition systems

- Similarity scores encode information about both the intrinsic and extrinsic performance

- We want to discover both the intrinsic and extrinsic performance by analyzing similarity scores

Performance Metric

- Similarity scores $S(x_i, g_k)$ between a query image x_i and the k th rank gallery g_k data are sorted in descending order and normalized to [0 1]:



An example of normalized similarity scores for a single query data

- Matching score: the similarity score corresponding to matched gallery image (the largest score for rank 1 recognition)

  Non-matching scores: the remaining scores
- A performance metric f_i is defined for query data x_i

$$f_i = \exp\{\frac{S(x_i, g_1) - \mu_i}{\sigma_i}\}$$

Perfect Recognition

- Perfect Recognition: duplicate gallery images as query images.

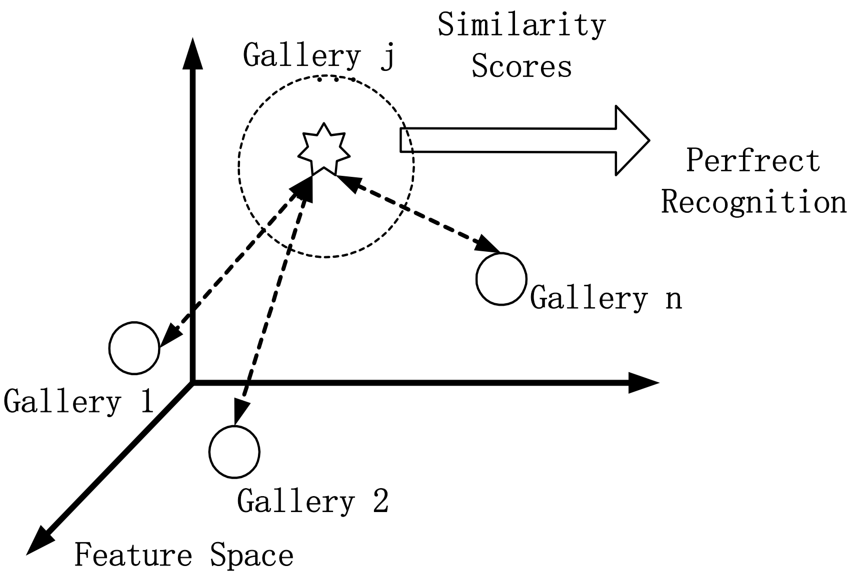
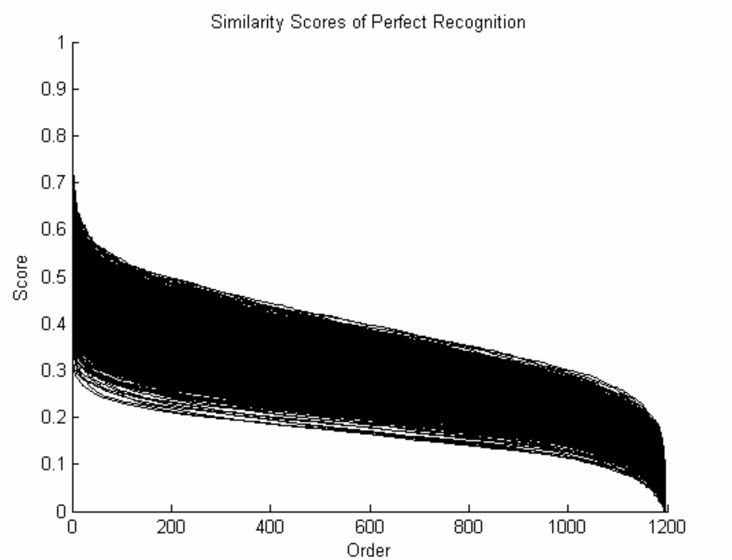


Illustration of perfect recognition

- Perfect Recognition Similarity Scores (PRSS): similarity scores between gallery images.

Modeling Intrinsic Performance

- Perfect recognition similarity scores (PRSS) are calculated for all the gallery images using only gallery data
- For each gallery data, calculate its performance metric f_i



An example of perfect recognition similarity scores

- Using average performance metric to represent perfect recognition

$$f = \frac{1}{N}\sum_i f_i$$

System Parameters Tuning

- f measures the performance of a FR system as a function of the intrinsic components
- Intrinsic parameters for a PCA based face recognition include
  - number of PCA coefficients
  - types of similarity measurements: L1, L2, Cosine
  - Measurement space: Euclidean or Mahalanobis

System Parameters Tuning (cont'd)

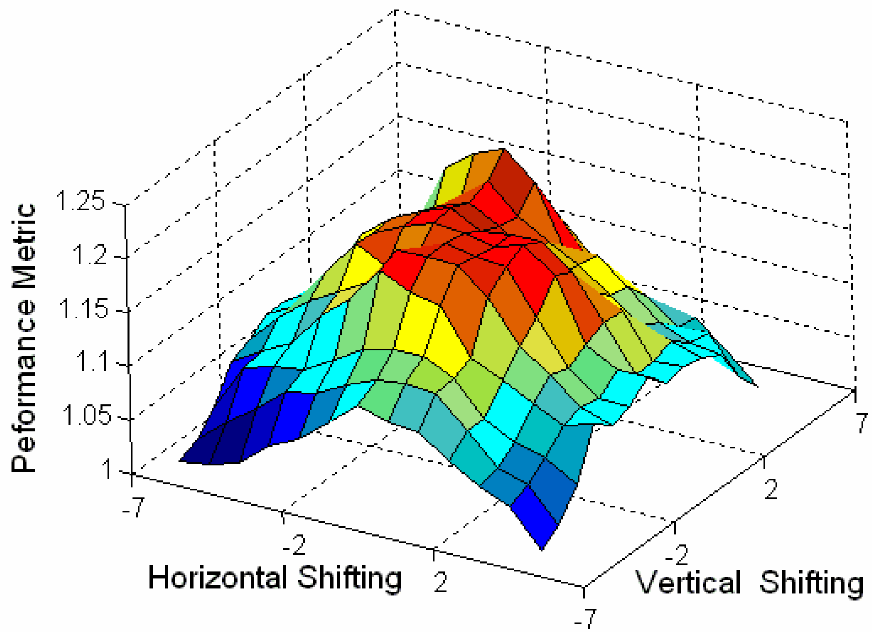
- Offline System Tuning: adjust algorithm parameters to achieve the largest f

Query Set	Accuracy of selected parameter	Offline selected parameters	Accuracy range
FERET FB	80.0%	[200, Cos., Maha.]	[70.2% , 82.0%]
FERET FC	49.4%	[200, Cos., Maha.]	[5.2% , 50.7%]
FERET Dup1	34.7%	[200, Cos., Maha.]	[22.6% , 38.8%]
FRGC Exp. 1	75.1%	[120, Cos., Maha.]	[32.7% , 75.5%]
FRGC Exp. 4	23.4%	[120, Cos., Maha.]	[4.9% , 27.0%]

Summary of offline parameter selection and actual recognition accuracy

Online Eye Adjustment

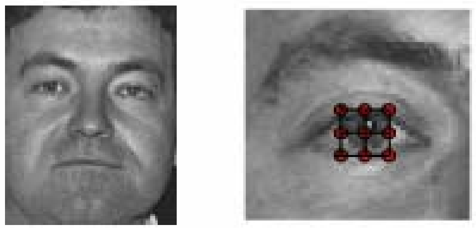
- Different eye locations provide different face alignment, therefore providing different values of performance metric f_i.



performance metric f_i varies with eye locations

- Search eye positions around initial eyes, and select the alignment candidate corresponding to the largest f_i

Online Eye Adjustment (cont'd)



(a) (b)

(c)

Adjust eye localization online. (a) face aligned on initial eyes (b) eye candidates to be adjusted, (c) eye locations before and after adjustments

Results show that the automatically adjusted eyes can provide better accuracy than manual eye localizations!

Data Set	Manual eyes	Adjusted on manual eyes	Automatic eyes	Adjusted on automatic eyes
FERET FB	79.8%	85.1%	74.8%	84.8%
FERET FC	49.3%	59.8%	43.3%	57.2%
FERET Dup1	34.8%	44.6%	30.6%	42.9%

Recognition accuracy with adjusted eyes

Online Eye Adjustment (cont'd)

Table 1. Summary of rank 1 recognition rate with adjusted eyes

Data	PCA (MahaCosine)		PCA (Eucldean)		LDA		EBGM	
	original	adjusted	original	adjusted	original	adjusted	original	adjusted
FERET FB	85.6%	86.6%	74.2%	75.2 %	71.9%	76.7 %	90.1%	91.4%
FERET FC	64.9%	65.5%	4.6%	5.7%	46.9%	55.7 %	40.7%	41.8%
FERET Dup1	44.2%	48.4%	33.7%	35.6%	22.9%	29.8 %	46.0%	47.2%

Online Performance Prediction

- Objective: predict if a probe image is correctly recognized

- Definition:
  - "success" recognition: the probe image is correctly recognized
  - "failed" recognition: the probe image is incorrectly recognized
  -

- Our method:
  - Extract features from the differences between actual recognition similarity scores and the corresponding perfect recognition similarity scores
  - Train a performance predictor using extracted features
  - Predict recognition results online for query images

Feature Extraction from Similarity Scores

- For a query image *x*, compute the actual recognition similarity scores (ARSS) and perfect recognition similarity scores (PRSS)
- Compute the differences between the kth-rank ARSS $S(x, j_k)$ and the corresponding m th-rank PRSS $S(j_m, j_k)$

$$d_k^m(x) = S(x, j_k) - S(j_m, j_k)$$

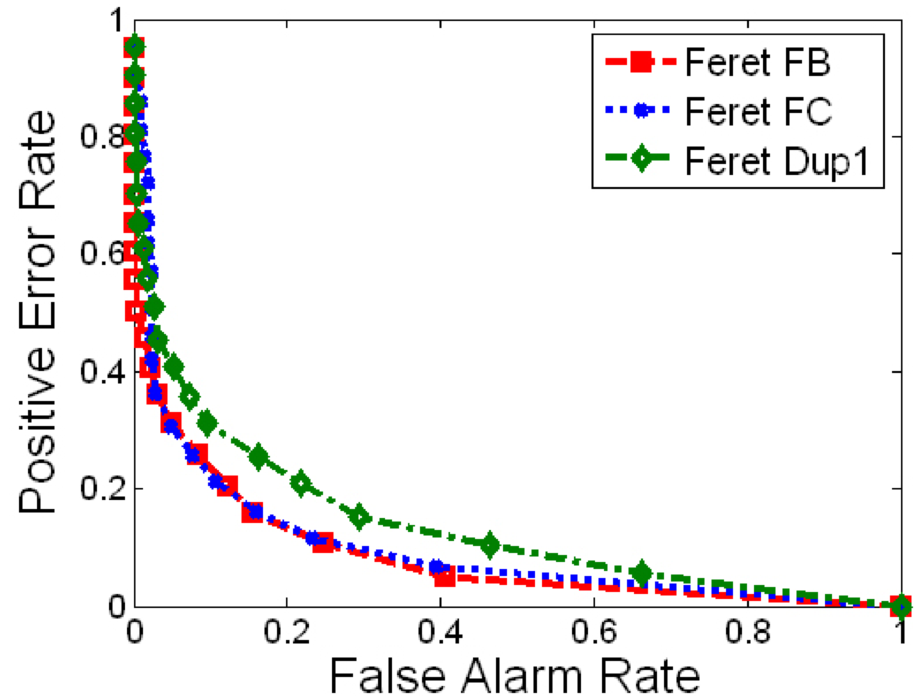
- Use multiple difference values between ARSS and PRSS at different ranks to form a feature vector (k=1,…,K, m= 1,…,M)

$$\begin{aligned} V = \{ \quad & d_1^1(x)w_1, \ldots, d_K^1(x)w_K, \\ & d_1^2(x)w_1, \ldots, d_K^2(x)w_K, \\ & \ldots, \\ & d_1^M(x)w_1, \ldots, d_K^M(x)w_K\} \end{aligned}$$

(w_K is a weighting factor)

Performance Predictor Training

- Train a performance predictor using extracted features with additional training data from the query images



ROC of trained performance predictor on FEREET probe sets

- A Support Vector Machine (SVM) is used as the predictor
- Supervised training with the FR as the supervisor
- Binary output: success or failed
  - False alarm: misclassify an failed recognition as success
  - Positive error: misclassify a corrected recognition as failed recognition

Improving Face Recognition Using Prediction

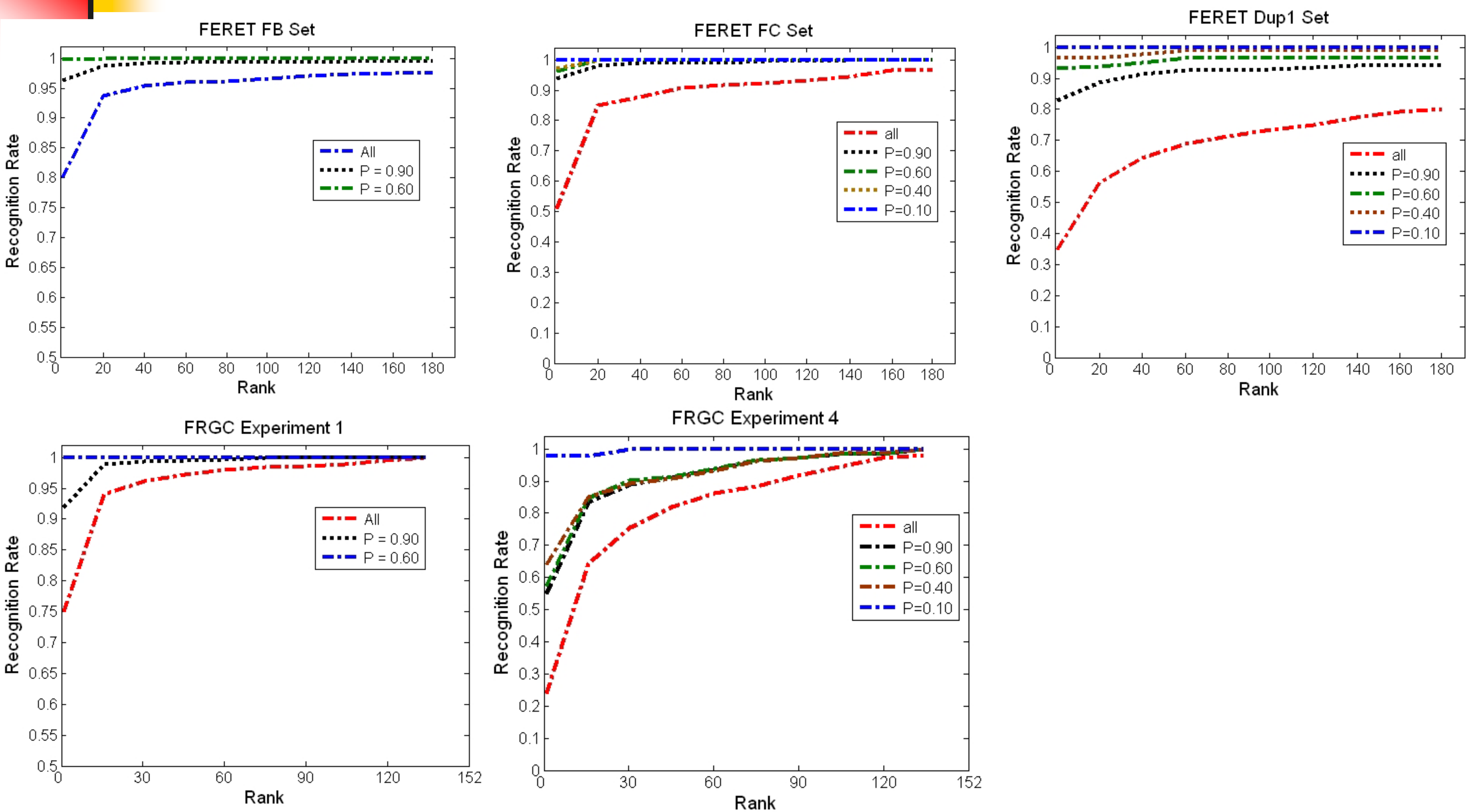
- **Cross-Validation on FRGC and FERET query sets**

  - Each set is divided into training and testing. The training set constitutes a small subset of the data set
  - Only the "success" data are retained while the "failed" data are removed from current recognition system
  - A threshold P is selected to preserve a certain percentage of "success" data

Data Set	All	P = 90%	P = 60%	P = 10%
FERET FB	80.0%	96.2%	99.7%	100.0%
FERET FC	49.3%	93.7%	96.4%	100.0%
FERET Dup1	34.7%	82.9%	93.1%	100.0%
FRGC Exp. 1	75.0%	91.8%	100.0%	100.0%
FRGC Exp. 4	23.9%	57.2%	64.3%	97.9%

Summary of recognition accuracy with/without using performance prediction

Improving Face Recognition Using Prediction (Cont'd)



CMC curves of face recognitions with/without using performance prediction

Amount of Training data v.s. Prediction Errors

Data Set	ratio = 0.2	ratio = 0.4	ratio = 0.6
FERET FB	0.1354	0.1370	0.1589
FERET FC	0.2296	0.2251	0.1801
FERET Dup1	0.0202	0.0496	0.0497

Summary

- We introduce performance metrics to model the intrinsic and extrinsic performance of a FR system, based on analysis of the similarity scores
- The proposed performance metrics can perform offline algorithm parameters tuning and online alignment adjustment using only gallery data
- The performance metrics can also be used for online performance prediction of query images
- The proposed methodologies may be extended for performance modeling and prediction of other similarity-based pattern recognition methods

More information about this work

- Peng Wang, Qiang Ji, and James Wayman, Modeling and Predicting Face Recognition System Performance Based on Analusis of similarity scores, IEEE Transactions on Pattern Analysis and Machine Intelligence, pages 665-670, Vol. 29, No. 4, April, 2006.

- Peng Wang, Lam Cam Tran, and Qiang Ji, Improving Face Recognition by Online Image Alignment, 18th International Conference on Pattern Recognition (ICPR'06) pp. 311-314, 2006.