Actively Exploring Creation of Face Space(s) for *Improved* Face Recognition

Nitesh V. Chawla and Kevin W. Bowyer
{nchawla, kwb}@cse.nd.edu
Department of Computer Science and Engineering
University of Notre Dame, USA

Abstract

We propose a learning framework that actively explores creation of face space(s) by selecting images that are complementary to the images already represented in the face space. We also construct ensembles of classifiers learned from such actively sampled image sets, which further provides improvement in the recognition rates. We not only significantly reduce the number of images required in the training set but also improve the accuracy over learning from all the images. We also show that the single face space or ensemble of face spaces, thus constructed, has a higher generalization performance across different illumination and expression conditions.

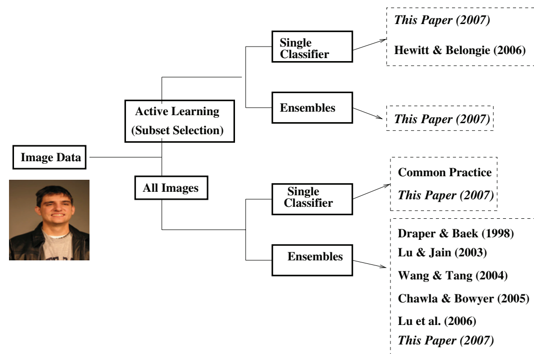

Figure 1: Related Work

Introduction

Quoting from the Government Security Newsletter: "It turns out that uncooperative subjects, poor illumination and the difficulty of capturing comparable images often make it difficult for face recognition systems to achieve the accuracy that government officials might seek in large-scale anti-terrorism applications." We want to be able to construct a face space that has an effective generalization capacity across subjects, illumination and expression variations. In addition, the substantially increasing number of images available for training require scalable methods for learning and evaluation.

It is very likely that subjects can have multiple images enrolled in a database, albeit with an expression, illumination or temporal variance. Hence, it becomes pertinent to be able to appropriately select the subset of the available images that leads to a performance improvement. This brings us to the important question: *How do we select the most useful subset of a set of images for training a face recognition system?* The question motivates the requirement of a methodology that can filter the "redundant" images of a subject.

Our Contributions Face recognition classifiers typically use all the images for either constructing a single classifier or an ensemble of classifiers (Draper & Baek 1998; Lu & Jain 2003; Wang & Tang 2004; Lu *et al.* 2006; Chawla & Bowyer 2005). Most of the work in ensembles has looked at re-sampling and random subspaces. (Lu *et al.* 2006) apply boosting with LDA for face recognition, wherein they develop a weakness analysis theory to work with both weak and strong learners. On the contrary, in this work we use active learning to learn a single classifier and ensemble to establish improvements in face recognition with PCA and the nearest neighbor classifier. (Hewitt & Belongie 2006) also apply active learning for face recognition, albeit from a video stream and in a limited setting; we will elucidate the key differences between our and their work in the next section. Figure 1 summarizes the main related works. We would also like to point out that our testbed is more comprehensive and diverse than other related work.

The fundamental research questions that we address in our paper are: a) Can we select an appropriate and relevant subset for training classifier(s) to achieve a comparable, if not better, rank-one recognition rate than the classifier trained with all the available images? b) Can these classifier(s) give consistent performances across different expressions and/or illumination conditions? c) How does the ensemble of such classifiers perform as compared to an ensemble constructed from randomly sampled images?

Experimental Scenarios

Temporal aging in the images: There is a time lapse between the images used in the training, gallery, and probe sets.
Varying expression and illumination conditions: We consider different combinations of expressions and illumina-

tions for training and evaluation (gallery and probe).
Unique subjects in gallery and probe: The subjects and images in the training set are mutually exclusive of the gallery and probe sets.
Streaming data: The images arrive in a conceptually streaming fashion. We start from the first data acquisition session, build a classifier, and then actively mine from subsequent sessions.

Typically in biometrics, the *probes* are the set of testing images, and the *gallery* is the set of images of the subjects that the probe images are matched against.

Actively Creating Face Spaces

Active learning trains a classifier by starting with an intial set of labeled training examples, and selecting additional examples based on their "effectiveness" score (Seung, Opper, & Sompolinsky 1992; Lewis & Gale 1994; Abe & Mamitsuka 1998; Roy & McCallum 2001; Iyengar, Apte, & Zhang 2000; Saar-Tsechansky & Provost 2004; Cohn, Atlas, & Ladner 1994). (Hewitt & Belongie 2006) use active learning for face recognition in video. We differ from their work in multiple ways. Firstly, we work in a general setting of 2-D face recognition, wherein the images have an element of temporal, illumination, and expression variation. Secondly, they build a face model by using annotated facial features, and use the model's error to select frames. We use PCA and nearest neighbor approach with the notion of "dissimilarity" to select images. Our algorithmic approach to active learning is very different. Thirdly, they have a small dataset of only six subjects, thus mitigating significant conclusions. We perform a much larger and comprehensive evaluation in a practical setting. Lastly, we also demonstrate performance improvements by learning ensembles after active learning.

The main goal of our work is to select a subset of relevant images that are representative and informative of the available images/subjects in the domain, without making any assumptions of the environment in which the images were acquired. We adopt active learning to enable an efficient search of the image space; we use the broad definition of active learning that gives the learning program control on the inputs on which it trains, contrasting with random selection of inputs (Cohn, Atlas, & Ladner 1994). The selected images should be "sufficient" for the recognition task and be generally applicable to different subjects or images captured under different environments. An accompanying goal is to reduce the computational complexity of recognition by significantly reducing the number of images required for recognition.

Learning Face Spaces

Two-dimensional face images are usually represented as high-dimensional matrices, where each matrix cell holds a gray-level intensity value. These raw feature vectors can be very large and highly correlated. For example, a 150 $\times$ 130 image (as typically used) can be unwrapped and represented by a vector of size 19,500. To counter this curse of dimensionality and highly correlated features, Principal Component Analysis (PCA) (Turk & Pentland 1991) is applied after performing geometric normalization and histogram equalization on the face images. This PCA transformed space is called the *face-space*. For classification purposes, we generate the projections of the face-space to the Mahalanobis space by whitening the PCA dimensions (unit variance transformation). Then, we use the Mahalanobis Cosine-based (*MahaCosine*) distance metric for the nearest neighbor classifier to perform recognition. Formally, the *MahaCosine* measure between the images *i* and *j* with projections *a* and *b* in the Mahalanobis space is computed as: $MahaCosine(i, j) = cos(\theta_{ij}) = \frac{|a||b|cos(\theta_{ij})}{|a||b|}$ (Beveridge & Draper).

The general procedure for classifying images is as follows. We assume that there are n_g gallery images in G, n_p probe images in P, and each probe image is indicated by p. A probe image is matched against each gallery image by computing the distance δ using the *MahaCosine* measure. For each probe image, we sort the gallery images by decreasing δ. The closest gallery image (or the 1-nearest neighbor) will be at the rank 0, and the furthest gallery image will be at rank $n_g - 1$. The classifier correctly recognizes a person if the probe image and the top ranked image in the gallery are of the same person. We define a rank-indicator function $r(p, G)$ that returns the first successful match (k) of a probe image in the gallery. For instance, $r(p, G) = 3$ implies that the match for a probe image is at the fourth-nearest neighbor. We are primarily interested in the number of correct matches at 1 nearest neighbor: $r(p, G) = 0$.

Having defined the procedure for training the face space and matching, we now describe the active sampling process. Before we proceed, we would again like to point out that we assume the distribution is known and the labels of all images are readily available at zero cost. Thus, we relax the definition of active learning given known distribution and concepts (Cohn, Atlas, & Ladner 1994). Our main goal is to find a smaller, but consistent, training set by allowing the classifier to selectively sample the most relevant and complementing set of images. We generate an initial training (or gallery) set L by randomly selecting n images from a pool of available data of size T. The remainder, $T - L$, forms the "unlabeled" (or probe) set UL. Note that UL is not really unlabeled, but we call it unlabeled for testing purposes and to distinguish it from the training set. We construct a face-space from L. The nearest neighbor classifier provides the distance (δ) between the images in UL and the images in L. The distances are indicative of the quality of match between an image in UL and an image in L, and result in a rank-ordering of images in UL. A sampling distribution D_s is derived from this rank-ordering as $\frac{r(p,L)}{\sum_{i \in UL} r(i,L)}$. The higher the rank of an image in UL, the higher the probability of selection. The intuition behind it is that we want to target the "difficult" images, where difficulty is defined by the lack of a representative in the training set. We then remove the selected images from UL without replacement, attach the true labels, add them into the training set L, and re-create the face space. Thus, L keeps increasing and UL keeps shrinking over the iterations. Note that once an image is removed from UL and added to L, it is not evaluated on again. The goal here is to actively maximize the coverage with most representative images for a person or subject. We only need to capture those images of a person that maximally differ from each other. This distinguishes us from the boosting algorithm as well.

The procedure continues for a prescribed number of iterations or until a stopping criterion is met. We introduce a heuristic for deciding the stopping criterion: the sum of ranks in UL should be within a threshold m. Recall, that the rank of an image goes from 0 to $n_g - 1$, where n_g is $|L|$. A rank 0 for an image in UL will imply that person is already in the training database. We did not notice any significant difference as we varied m from 30 to 100. Thus, we set $m = 50$. We will call this approach $Active_n$, where n indicates the starting sample size of the training set (L). It is summarized as follows.

- **Input:** pool of images T; size of sample n and n_s.
- Randomly select n images from T to create the initial training set L. Set $UL = T - L$.
- For (i=1; until stopping criterion; i++)
  - Construct a face space from L (PCA transformation).
  - Compute the distances for each (transformed) image in UL to each image in L.
  - Compute the rank-ordering of the images in UL. Convert the ranks to a distribution D_s.
  - Based on D_s sample without replacement n_s images from UL and add to L. Update $UL = UL - n_s$; $L = L + n_s$.
- **Output:** The final set L.

Data Collection

We acquired the collection B database from the University of Notre Dame (Flynn, Bowyer, & Phillips 2003). The subjects participate in the acquisition repeatedly (at most once in a week) over a period of time. Images of the subjects are captured with two side lights on (**LF**) and two side lights and a center light on (**LM**). In addition the subjects are imaged with two expressions: neutral (**FA**) and smile (**FB**). There is a total of 82 subjects in Spring'02, 333 subjects in Fall'02, and 334 subjects in Spring'03. There is a total of 487 unique subjects across the three semesters. Note that we guaranteed that all the images are different among the training, gallery, and probe sets.For the purposes of our experiments, we constructed three variants of training, gallery, and probe sets with different number of subjects/images and temporal/expression/illumination variations.

Experiment-1 Data: We randomly selected 121 subjects with four images each for the training set. We chose not to select all the subjects to maintain the difficulty of recognition: having a set of unique subjects in the gallery and probe sets. Then, we considered all the subjects that had at least three acquisition sessions over the course of the semester. The second session of acquisition for those subjects became the gallery set and the last session became the probe set. We used the longest possible time-gap between enrolling a gallery image and the subsequent probe image, thus permitting the greatest possible natural variations in the subjects' images. This gave us 381 subjects (381 $\times$ 4 images) in each gallery and probe set. We tried to mimic a setting that may be used in a face recognition system — a) the subjects enrolled in the gallery may not always be in the training data; b) the images of the subjects in the gallery or probe set will have an element of time-lapse from the training set images; and c) the probe set images will have a time-lapse from the gallery images.

Experiment-2 Data: There are a total of 487 unique subjects in the collection. For this experiment, we retained all the subjects along with one image from each of FA-LF, FA-LM, FB-LF, FB-LM, thus giving us a total of 1948 images in the training set. A similar approach to constructing the gallery and probe sets was followed for Experiment-2, resulting in 381 subjects in the probe and gallery sets.

Experiment-3 Data: We split the Spring'02 and Fall'02 data into m sessions, where m is the maximum number of times one of the subjects appeared. More subjects come for fewer sessions. There are 333 subjects in Fall'02 that participate in at least one session, and subsequently the number of subjects who come over all the multiple sessions in Fall'02 declines. We then provided the images in a (conceptually) streaming fashion, ordered by time of acquisition, to actively sample from. We split up the sessions for each of the expression and illumination combinations. This allowed us to evaluate the temporal aging of images without any illumination and expression variations. We constructed the gallery and probe sets from the Spring'03, resulting in 333 subjects in both gallery and probe sets. This permitted an "out-of-time" testing as no acquisitions of Spring'03 were included in the training set. Table 1 summarizes the datasets' sizes.

Table 1: The size reflects the aggregate number of images across the expression and illumination variations.

Data	Training Set Size	Gallery Set Size	Probe Set Size
Experiment-1	484	1,524	1,524
Experiment-2	1,948	1,524	1,524
Experiment-3	10,564	1,332	1,332

Experiments

We set up our experiments to answer the questions set out in the Introduction. We will report on each of the pre-defined experiments individually, given the different objectives and properties. We established the following common classification schemes as benchmarks for each of the experiments, in addition to their respective $Active_n$ runs. We report average rank-one recognition rates and standard deviations, including statistical test of significance, where applicable.

1. Single *specialized* face space: This is the face space trained on a particular expression and illumination combination, and only tested on the corresponding combination.
2. Complete face space (*All*): This is constructed on the complete training set consisting of all the images with illumination and expression variations. The only exception

Table 2: Rank-one recognition rate of various approaches for Experiment-1. $Specialized$ is denoted by $Sp.$

	$Active_{n=50}$	$Random_{k=250}$	All	$Sp.$
FA-LF	67.7 ± 0.8	65.8 ± 0.9	67.9	66.1
FA-LM	72.6 ± 1.2	70.3 ± 1.3	72.7	69.5
FB-LF	73.0 ± 0.9	70.9 ± 1.2	72.4	70.3
FB-LM	75.5 ± 1.0	72.2 ± 1.7	74.0	72.2

to this is Experiment-3, where we considered each illumination and expression combination separately (the *specialized* case). The goal of Experiment-3 is to evaluate the temporal aging scenario.

3. $Random_k$: As a primary benchmark, we randomly sampled approximately the same number of images (k) as selected (on an average) by the $Active_n$ methodology. This was also performed 10 different times. The purpose was to see if the randomly selected set of images can achieve the same rank-one recognition rate as compared to the "selected" set.

4. *Ensemble approaches*: We also voted the scores (using sum rule) from each of the 10 random runs for both $Random_k$ and $Active_n$. Our thesis is that the actively learned classifiers will potentially be more diverse than the randomly generated classifiers. Hence, we expect the ensemble generated from actively learned classifiers to outperform the ensemble generated from classifiers learned from randomly selected images. We will call the ensemble approaches $Ensemble_A$ and $Ensemble_R$ for active and random, respectively. The $Active_n$ based ensemble, $Ensemble_A$, is in a similar theme as boosting. However, unlike boosting, it does not include all the possible images (training set examples) in subsequent iterations. Secondly, each subsequent iteration of $Active_n$ is sampling without replacement: when images are selected they are removed from the UL set and are not evaluated on again for "correctness" or "incorrectness".

Experiment-1

We considered the data outlined for Experiment-1. Using $n = 50$ for $Active_n$, we ran the experiment 10 different times. The choice of the initial 50 images can trigger a different exploration space. Thus, each face space can have a potentially different representational bias. On an average the $Active_n$ procedure resulted in a selection of less than or equal to 250 images. *This is less than 50% of the original number of images.*

Table 2 compares the performance of the various outlined approaches. The individual face spaces constructed from $Active_n$ outperform $Random_k$, where $k = 250$ (resultant number of images from $Active_n$); the *Specialized* face space; and *All*. Note that none of the face spaces in Table 2 have been tuned by dropping the eigen vectors from front or behind. $Active_{n=50}$ outperforms $Random_{k=250}$ for all the illumination and expressions combinations. This difference is **statistically significant**, using the paired t-test for

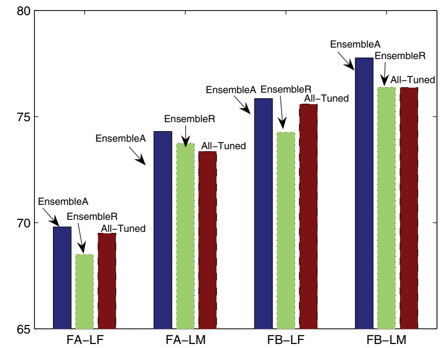

Figure 2: Rank-one recognition rate of Ensembles for Experiment-1. The ensembles are compared with the well-tuned *All* classifier.

means over the 10 random runs, for the FA-LM and FB-LM conditions.

The face space trained on all the images, All, is comparable to $Active_n$. Interestingly, the *specialized* classifier does not lead to any performance improvement. That is, there is really no gain in evaluating the probe point using the classifier specifically trained for that probe space. This adds evidence to our conjecture that there is natural variation in expression and illumination, and we want diversity in our face space to increase the recognition rate.

Figure 2 shows the performance of the ensembles of actively learned classifiers, $Ensemble_A$, and the ensemble of classifiers from random selection of images, $Ensemble_R$. The results agree with our conjecture that the diversity in the choice of the training images through active learning provides a significant improvement over random ensembles. We also compare to the performance achieved by the well-tuned All, where tuning was performed via dropping of eigen vectors from front and behind and evaluating directly on the testing set. Thus, the tuned face space is essentially an empirically observed upper limit on the recognition rate achieved by All. Ideally, any tuning should be performed on a validation set. $Ensemble_A$ just constructed from 10 classifiers, without any individual tuning of a face space, easily achieves the performance attained by the well-tuned single face space.

Experiment-2

Experiment-2 considered the second larger dataset of 1948 images. For consistency, we again used the starting point of $n = 50$ images for $Active_n$. We again ran 10 random runs for both $Active_n$ and $Random_k$. Based on the number of images discovered by active learning — approximately 550 out of 1948 images on an average over 10 runs — we randomly selected the same number of images.

Table 3 includes the performance of various approaches on the Experiment-2 data. $Active_{n=50}$ **statistically significantly** outperforms $Random_k$, using paired t-test for means at 95% confidence, for all the four expression and illumination variations. *It is noteworthy that the actively generated*

Table 3: Rank-one recognition rate of various approaches for Experiment-2. $All - 1160$ is the classifier with the default tuned face space (60% of the vectors are dropped from behind).

	$Active_{n=50}$	$Random_{k=550}$	All	$All - 1160$	$All - 389$	$Specialized$
FA-LF	77.47 ± 0.96	75.28 ± 0.93	70.6	72.7	76.90	77.69
FA-LM	77.69 ± 0.84	75.67 ± 0.93	68.24	71.39	76.37	77.69
FB-LF	79.58 ± 0.32	77.09 ± 1.09	69.29	75.06	78.21	79.00
FB-LM	80.29 ± 0.73	78.29 ± 1.27	71.35	75.59	79.79	80.57

face space from 28% of the total number of images outperforms the face space constructed from all the images. The classifier learned on the entire training set, All, significantly under-performs. Potentially, the multiple images of subjects are contributing to the low energy and low variance eigen vectors. To evaluate this, we dropped 60% of the vectors from behind, which are typically associated with the low variance (Chawla & Bowyer 2005). This dropped the number of eigen vectors to 1160, and resulted in a significant improvement in performance. However, this "default-tuned" face space still was not competitive to active learning approach.

To further explore the face space learned from all the images, without any active learning, we implemented a wrapper-based dimension selection approach. We carefully tuned the face space by dropping eigen vectors and evaluating directly on the testing set. We retained only 389 dimensions from the original 1947 dimensions in the face space. Interestingly, once the number of eigenvectors was reduced down to 389, the performance significantly improved, almost to the level of $Active_n$. We hasten to add that without any tuning whatsoever, $Active_n$ is able to achieve and even overcome the empirical upper bound of the recognition rate obtained by the single face space trained on all the images. It resonates with the premise for active learning: ignoring the redundant information.

Figure 3 includes the performance of the ensemble methods and the well-tuned single classifier. For clarity in comparisons, we only include the performance of the ensemble methods and the single face space (All) that is tuned directly on the testing set. The comparisons to other individual face spaces can be done by juxtaposing with the results in Table 3. $Ensemble_A$ significantly improves performance over all the techniques, including $Active_k$. $Ensemble_R$, while improving over $Random_k$, does not provide any significant improvements over $Active_k$. That is, the single face space constructed from actively exploring the image space is comparable to the ensembles generated from randomly selected images. It is not very surprising that $Ensemble_A$ outperforms the $Ensemble_R$ as each individual member of former has its own selection bias towards sampling various images in the training set.

Experiment-3

We divided the subjects in Spring'02 and Fall'02 into as many sessions, in chronological order, as they appeared over the course of two semesters. This resulted in 23 sessions over two semesters. This division was performed for each individual pair of illumination and expression conditions. Given the much larger number of images, we did not construct a single face space from all of the approximately 10,564 images as that became computationally prohibitive with its memory requirements. Note that we did not run any ensemble experiments for this data, as we were more interested in constructing a single classifier from a streaming set of images.

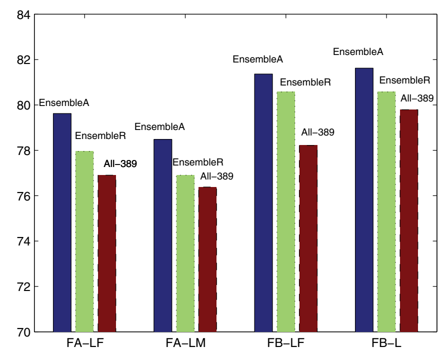

Figure 3: Rank-one recognition rate of Ensembles for Experiment-2. $All - 389$ is the single face space tuned directly on the testing set, resulting in only 389 dimensions.

Experiment-3 evaluated the effect of temporality in selection, keeping expression and illumination constant. Since, the first session was comprised of 82 images, we set $n = 82$ for $Active_n$. We compare $Active_n$ with $Random_k$ and *Specialized*. Also, we only selected the same number of random images as approximately discovered by $Active_n$. Our results demonstrate that the *Specialized* classifier constructed from all the corresponding images performs poorly as compared to both active and random selections. Default tuning improved the performance for the $specialized$ classifiers, but it is still lower than $Active_n$. We only used the default tuning, as the computational cost of searching the optimal dimensions quickly became prohibitive with the much larger number of images in Experiment-3.

Figure 4 shows the results. Again, using the very different data setting for active learning, there is a definite advantage of using active learning to select the reduced number of images and improve the performance. While the probe and gallery sets had the same illumination and expressions as the trained face space, learning the specialized classifier still did not improve the performance. Once again, active learning permits a **significant reduction** in the number of required images. We are able to select less than 20% of the images on average, and still perform better than the specialized classifier (19.35% for FA-LF; 23.13% for FA-LM; 18.63% for FB-LF; and 15.11% for FB-LM).

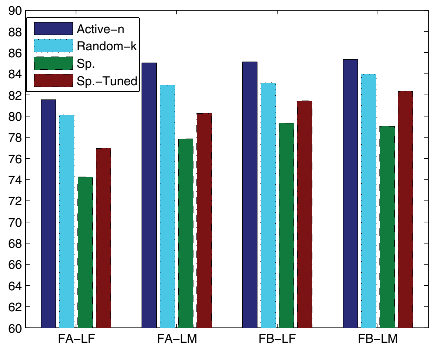

Figure 4: Rank-one recognition rate for Experiment-3. $Specialized$ is denoted by $Sp.$ and default tuned $Specialized$ is denoted by $Sp - Tuned$.

Conclusions

We summarize our conclusions, in the light of the questions posited in the Introduction.

- *Actively created face spaces allow a substantially smaller set of images to achieve the same performance as the larger unsampled set. The resultant face spaces also generalize effectively across a variety of expression and illumination conditions.* Using a variety of scenarios, we robustly demonstrated that it is indeed possible to actively sample a small subset images for training a classifier. we selected 50% of the available images for Experiment-1; 28% of available images for Experiment-2; and on an average 19% of images for Experiment-3. $Active_n$ also generalizes well across different sets of images, without any explicit assumption, thus mitigating the need of specialized and tuned classifiers. We were able to easily achieve or exceed the performance of a classifier (All) with face space tuned directly on the probe set. We showed that most of the improvements across different expression and illumination conditions were **statistically significant** at 95% confidence level using paired t-test, as compared to random selection. Our observations establish that by addressing the relevance of selected images with respect to the other images leads to an improvement in generalization performance of the classifier(s).
- *Ensembles from actively learned face spaces give a better performance than a single face space that is well-tuned directly on the testing set.* We also showed that the ensembles, $Ensemble_A$, constructed from the $Active_n$ framework outperform the other ensemble methods as well as the single well-tuned classifier (All) optimized directly on the probe set. Note that the well-tuned performance is actually an empirical upper bound as it is optimized directly on the testing set. $Ensemble_A$ constructed from just 10 classifiers, without any individual tuning of a face space, easily overcomes the performance attained by that well-tuned single face space. We believe that adding more diverse members in the ensemble should further improve the performance.

Acknowledgements

This work is supported by the NSF under grant CNS-0130839, by the CIA, and by the DOJ under grant 2004-DD-BX-1224.

References

Abe, N., and Mamitsuka, H. 1998. Query learning strategies using boosting and bagging. In *Fifteenth ICML*, 1–9.

Beveridge, D., and Draper, B. Evaluation of face recognition algorithms (release version 4.0).

Chawla, N. V., and Bowyer, K. W. 2005. Random subspaces and subsampling for 2-d face recognition. In *IEEE CVPR*, 582–589.

Cohn, D. A.; Atlas, L.; and Ladner, R. E. 1994. Improving generalization with active learning. *Machine Learning* 15(2):201–221.

Draper, B., and Baek, K. 1998. Bagging in computer vision. In *IEEE CVPR*, 144–149.

Flynn, P. J.; Bowyer, K. W.; and Phillips, P. J. 2003. Assessment of time dependency in face recognition: An initial study. In *International Conference on AVBPA*, 44–51.

Hewitt, R., and Belongie, S. 2006. Active learning in face recognition: Using tracking to build a face model. In *CVPR'06 Workshop*.

Iyengar, V. S.; Apte, C.; and Zhang, T. 2000. Active learning using adaptive resampling. In *ACM SIGKDD*, 92–98.

Lewis, D., and Gale, W. A. 1994. A sequential algorithm for training text classifiers. In *ACM SIGIR*, 3–12.

Lu, X., and Jain, A. K. 2003. Resampling for face recognition. In *International Conference on AVBPA*, 869 – 877.

Lu, J.; Plataniotis, K.; Venetsanopoulos, A.; and Li, S. 2006. Ensemble-based discriminant learning with boosting for face recognition. *IEEE Transactions on Neural Networks* 17(1):166 – 178.

Roy, N., and McCallum, A. 2001. Towards optimal active learning through sampling estimation of error reduction. In *ICML*, 441–448.

Saar-Tsechansky, M., and Provost, F. 2004. Active sampling for class probability estimation and ranking. *Machine Learning* 54(2):153 – 178.

Seung, H. S.; Opper, M.; and Sompolinsky, H. 1992. Query by committee. In *COLT*, 287–294.

Turk, M., and Pentland, A. 1991. Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3(1):71–86.

Wang, X., and Tang, X. 2004. Random sampling LDA for face recognition. In *IEEE CVPR*, 259–265.