Multi-Modal 2D and 3D Biometrics for Face Recognition

Kyong I. Chang Kevin W. Bowyer Patrick J. Flynn

Computer Science & Engineering Department
University of Notre Dame
Notre Dame, IN 46556
{kchang,kwb,flynn}@cse.nd.edu

Abstract

Results are presented for the largest experimental study to date that investigates the comparison and combination of 2D and 3D face data for biometric recognition. To our knowledge, this is also the only such study to incorporate significant time lapse between gallery and probe image acquisition. Recognition results are presented for gallery and probe datasets of 166 subjects imaged in both 2D and 3D, with six to thirteen weeks time lapse between gallery and probe images of a given subject. Using a PCA-based approach tuned separately for 2D and for 3D, we find no statistically significant difference between the rank-one recognition rates of 83.1% for 2D and 83.7% for 3D. Using a certainty-weighted sum-of-distance approach to combining 2D and 3D, we find a multi-modal rank-one recognition rate of 92.8%, which is statistically significantly greater than either 2D or 3D alone.

1. Introduction

The identification of the human face in 2D has been investigated by many researchers, but relatively few 3D face identification studies have been reported. Each imaging modality has its own benefits and problems in the task of human face recognition. 2D images are generally easier and less expensive to acquire. The perceived benefits from using 3D relative to 2D data include less variation observed due to factors such as makeup and less sensitivity to illumination changes. According to a recent literature survey [1], there are two main strategies in 2D face recognition, statistical approach and neural network approach based on facial features [2, 3]. One of the main motivations of 3D face recognition is to overcome the problems in general 2D recognition methods resulting from the illumination changes, expression or pose variations. There are 3D face recognition methods proposed by several studies [4, 5, 6, 7, 8].

This study deals with face recognition using multiple sensors (CCD and range finder). Each sensor captures different aspects of facial features, 2D intensity representing surface reflectance and 3D depth values representing face shape data. Even though each imaging modality has its own advantages and disadvantages depending on certain circumstances, there is often some expectation that 3D data should yield better performance. However, no rigorous experimental study has been reported to validate this expectation. The experiments reported in this study are aimed at (1) testing the hypothesis that 3D face data provides better biometric performance than 2D face data, using the PCA-based method, and (2) exploring whether a combination of 2D and 3D face data may provide better performance than either one individually.

One aspect of combining different biometrics is how to combine results provided by individual sources effectively during the decision process. Many different approaches could be envisioned for combining multiple types of biometric information. In general they can be thought of as occurring at the image level, the metric level, or the rank level. In this study, the combination of face data at the metric level has been considered.

2. Previous Work

In this section, methods that use multiple types of facial data for identification purposes, multi-modal biometrics, are reviewed. The term "multi-modal biometrics" is used here to refer to the use of different sensor types without necessarily indicating that different parts of the body are used. The important aspects of these multi-modal studies are summarized in Table 1. Wang *et.al.* computed Gabor filter response in 2D images and point signatures in the 3D range data to obtain features for face recognition. A similarity function and SVM were tested for the classification. They concluded that SVM classifies better than the similarity function, and that integrated features (Gabor coefficients and point signature) perform better than a single feature alone [9]. Face profile data obtained with 2D and 3D facial images for automatic face verification is proposed by Beumier *et.al.* [10]. The full facial surface is constructed based

Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG'03)

Table 1: Previous studies that integrated multiple types of facial data for recognition.

Source	Facial Data *(Subjects)*	Methods	Fusion
[9]	2D frontal & 3D range image *(50)*	Gabor filter response & point signature	support vector machine (SVM)
[10]	2D frontal & 3D profiles *(120)*	profile matching	weighted sum
[11]	2D frontal & 2D profile *(30)*	HMM & eigenfaces & profile matching	logarithmic score transformation

on geometric features of the external contour along with the profile-based approach. A weighted sum of the 2D and 3D scores is used to deliver the fusion process. Combination of both frontal and profile view of 2D face data for identification through the combination of face classifiers is reported in [11]. While the profile view alone provides lower performance than frontal view, the profile classifier combined with HMM or eigenfaces using the frontal view performs better than recognition with single frontal view alone.

In addition to recognition methods based solely on the human face, there are other recognition methods using multiple biometric sources. In conjunction with face data, gait [12], ear [13], voice [14, 15, 16, 17], voice and lips [18], fingerprints [19], hand geometry and fingerprints [20, 21], and profile [11] have been used to improve overall recognition reliability. One commonality of the studies described above is that the identification rate based on multiple sensors / biometrics sources provides overall performance improvement.

3. Methods and Materials

3.1. 2D and 3D Face Recognition Using PCA

Extensive work has been done on face recognition algorithms based on PCA, popularly known as "eigenfaces" [22]. A standard implementation of the PCA-based algorithm [23] is used in the experiments reported here.

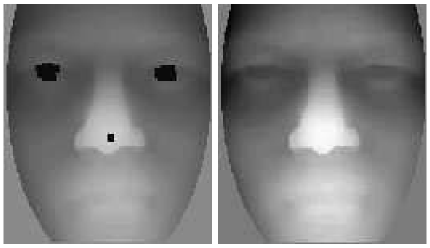
3.2. Normalization

The main objective of the normalization process is to minimize the uncontrolled variations that occur during the acquisition process and to maintain the variations observed in facial feature differences between individuals. The normalized images are masked to "gray out" the background and leave only the face region (See Figure 3). This is done to projected 2D data (Figure 3 - (b)) and 3D data (Figure 3 - (c)) only. In our data acquired by the Minolta Vivid-900 range scanner, every data point has a depth value as well as intensity value. While each subject is asked to gaze at the camera during the acquisition, it is inevitable to obtain data with some level of pose variations between acquisition sessions.

The 2D image data is typically treated as having pose variation only around the Z axis, the optical axis. The PCA software uses two landmark points (the eye centers) for geometric normalization to correct for rotation, scale, and position of the face for 2D matching. However, the face is a 3D object, and if 3D data is acquired there is the opportunity to correct for pose variation around the X, Y, and Z axes. Because the range sensor acquires a color texture map registered with the 3D data, it is in principle possible not only to correct the 3D data to a standard pose, but to then also create a projected 2D image from that same standard pose.

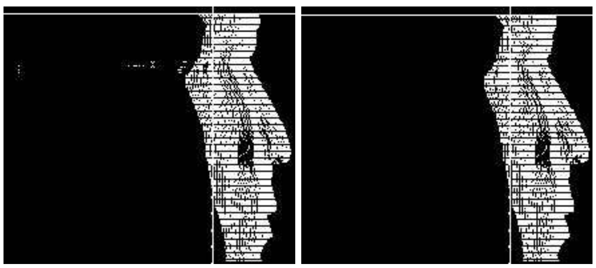
A transformation matrix is first computed based on the surface normal angle difference in X (roll) and Y (pitch) between manually selected landmark points (two eye tips and center of lower lip) and predefined reference points of a standard face pose and location. Pose variation around the Z axis (yaw) is corrected by measuring the angle difference between the line across the two eye points and a horizontal line. At the end of the pose normalization, the nose tip of every subject is transformed to the same point in 3D relative to the sensor (See Figure 2). After the 3D data points are transformed, a projected 2D intensity image is created from the color texture map that is associated with the 3D data.

Creating a projected 2D image from the texture map associated with the 3D data after the 3D pose correction might, at first, seem to have only advantages. After all, it corrects for more real pose variation than correcting the 2D image only for pose variation around the Z axis, as is done for applying PCA to standard 2D images. However, there are complications that can occur in creating a projected 2D image from sensed 3D data. As the 3D data is originally acquired, there is a "complete" 2D color texture map; that is, there is a color texture value even for some points that fail to produce a valid 3D value. This failure to produce a valid 3D value at some potential sample points is due to the particular method of sensing 3D, in this case, structured light using a projected stripe. There may be missing or invalid 3D data in regions of the face such as eyeballs or eyebrows, even though there is a 2D color texture sample at these points. When the original pose of the 3D data is changed, the projected 2D image that is created for the new 3D pose will have "holes" where there is missing 3-D data. Thus the projected 2D image is more fully pose corrected than the original 2D image can be, but it will also sometimes have some



(a) (b)

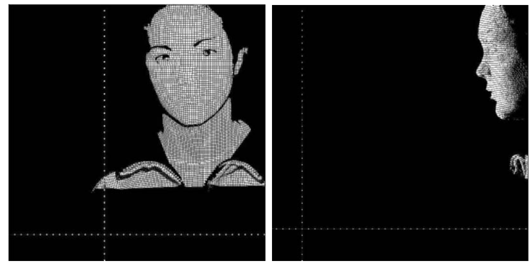
Processing missing data points in range data



(c) (d)

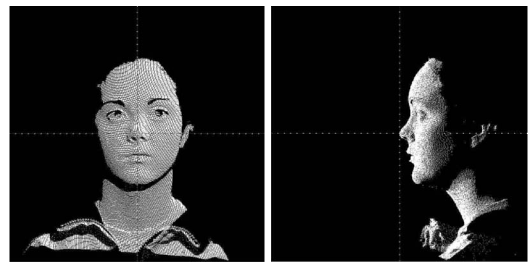
Processing spike noise in range data

Figure 1: Processing the missing holes in 3D



(a) X-Y plane (b) Y-Z plane

Initial pose of a subject in 3D space



(a) X-Y plane (b) Y-Z plane

Corrected pose of a subject in 3D space

Figure 2: Pose normalization

missing data around the regions of the eyes and eyebrows.

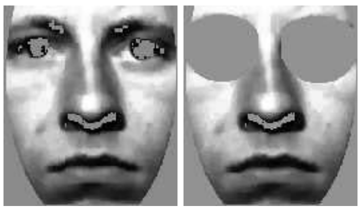
This problem with the 3D is alleviated to some degree by preprocessing the 3D data to fill in holes and remove spikes (See Figure 1). This is done by median filtering followed by linear interpolation using valid data points around a hole. However, even though we attempt to fill in the missing holes in 3D, there are regions where filling holes is not sufficient, such as the nostrils area after the pose correction. Because recovering missing holes in 3D shape data is in principle related to interpolating missing 2D intensity data, each imaging modality uses its own mask. Also, this indicates that the normalization process needs to be applied according to the level of data quality acquired by each sensor. The histogram equalization is performed to normalize only the intensity 2D face images. For the missing data points in intensity images, a mask is used that ignores the eye regions, where data is severely corrupted due to the specular surface (See Figure 3-(b)). As illustrated in the Figure 3-(c), however, missing holes on eye area are reliably filled compared to holes in nostrils area in 3D. This encourages us to block nostrils area in 3D.

3.3. Data Collection

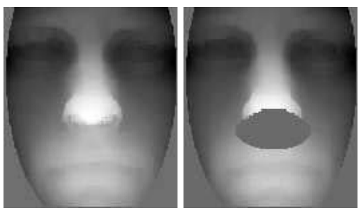
A gallery image is an image that is enrolled into the system to be identified. A probe image is a test image to be matched against the gallery images. Images were acquired at the University of Notre Dame between January and May 2003. Two four-week sessions were conducted for data collection, approximately six weeks apart. The first session is to collect gallery images and the second session is to collect probe images. Thus, for a given subject in our study, there is at least six and as many as thirteen weeks time lapse between the acquisition of their gallery image and their probe image. All subjects completed an IRB-approved consent form prior to participating in each data acquisition session. A total of 278 different subjects participated in one or more data acquisition sessions. Of these 278 subjects, 166 participated in both a gallery acquisition and a probe acquisition. Thus, for the experiments in our study, there are 166 individuals in the probe set, the same 166 individuals in the gallery, and 278 individuals in the training set. The 278 in the training set are the 166 in the gallery plus the 112 for whom good data was not acquired in both the gallery and probe sessions.

In each acquisition session, subjects were imaged using a Minolta Vivid 900 range scanner. Subjects stood approximately 1.5 meter from the camera, against a plain gray background, with one front-above-center spotlight lighting their face, and were asked to have a normal facial expression ("F_A" in FERET terminology [24]) and to look directly at the camera. The height of the Minolta Vivid scanner was adjusted to the approximate height of the subject's face, if needed. The Minolta Vivid 900 uses a projected light stripe to acquire triangulation-based range data. It also captures a color image near-simultaneously with the range data cap-

(a) Original 2D mask image (no blocking)



(b) Projected 2D mask image (before / after blocking)



(c) 3D mask image (before / after blocking)

Figure 3: Examples of mask images in 2D and 3D

ture. The result is a 640 by 480 sampling of range data and a registered 640 by 480 color image. (Note that there can be some regions of missing values in the 3D data.) These are the 3D and 2D images used in the experiments reported here.

3.4. Distance Metrics

2D data represents a face by intensity variation whereas 3D data represents a face by shape variation. It is obvious that the "face space" could be very different between modalities. Thus, during the decision process, certain metrics might perform better in one space than in the other. In this experiment, Euclidean distance and Mahalanobis distance metrics were explored for possible use during the decision process for the gallery matching [25]. Mahalanobis performed best in both cases. Eigenvector selection for the "face space" was done separately for each modality.

3.5. Data Fusion

The pixel level provides perhaps the simplest approach to combining the information from multiple image-based biometrics. The images can simply be concatenated together to form one larger aggregate 2D-plus-3D face image. The metric level focuses on combining the match distances that are found in the individual spaces. Having distance metrics from two or more different spaces, a rule of how to combine the distances across the different biometrics for each person in the gallery can be applied. The ranks can then be determined based on the combined distances.

One of the early tasks in data fusion is to normalize the scores, which are the results of a metric function. Scores from each space need to be normalized to be comparable each other. There are several ways of transforming the scores including linear, logarithm, exponential, logistic, etc. [11]. The scores are normalized so that the distribution and the range of score values are mapped to the same domain between for both modalities.

There are many ways of combining different metrics to achieve the best decision process, including majority vote, sum rule, multiplication rule, median rule, min rule, average rule and so on. Depending on the task, a certain combination rule might be better than others. It is known that sum rule and multiplication rule provide generally plausible results [26, 11, 12].

In our study, a weight is estimated based on the distribution of the top three ranks in each space. The motivation is that a larger distance between first- and second-ranked matches implies greater certainty that the first-ranked match is correct. The level of the certainty can be considered as a weight representing the certainty. The weight can be applied to each metric as the combination rules are applied. The multi-modal decision is made as follows. First the 2D probe is matched against the 2D gallery, and the 3D probe against the 3D gallery. This gives a set of N distances in the 2D face space and another set of N distances in the 3D face space, where N is the size of the gallery. A plain sum-of-distances rule would sum the 2D and 3D distances for each gallery subject and select the gallery subject with the smallest sum. We use a confidence-weighted variation of the sum-of-distances rule. For each of 2D and 3D, a "confidence" is computed using the three distances in top ranks as (second distance - first distance) / (third distance - first distance). If the difference between the first and second match is large compared to the typical distance, then this confidence value will be large. The confidence values are used as weights in the sum of distances. A simple product-of-distances rule produced similar combination results, and a min-distance rule produced only slightly worse combination results. Thus it appears that any of a variety of combination rules can give good results.

4. Experiments

There are three parts to this study. The first part is to examine the performance of original 2D and projected 2D. The second part is to evaluate the performance of 2D and 3D independently. Data fusion is considered, in the third part,

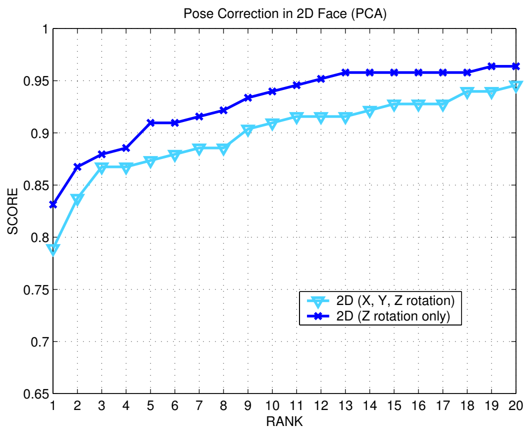


Figure 4: 2D recognition performance in different pose corrections.

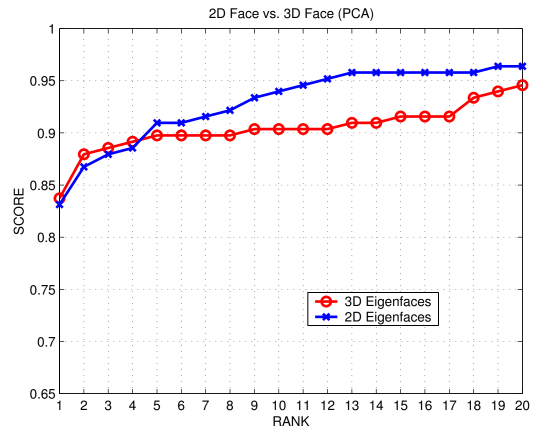


Figure 5: 2D face versus 3D face using eigenfaces.

to combine results at the metric level with different fusion strategies.

The eigenvectors for each face space are tuned by dropping the first M and last N eigenvectors to obtain an optimum set of eigenvectors. Thus, we expected to have a different set of eigenvectors representing 2D face space versus representing 3D face space. The cumulative match characteristic (CMC) curve is generated to present the results. The McNemar statistical significance test is considered based on rank-one recognition rates.

4.1. Experimental Results: Original 2D face versus projected 2D face

Considering the problems encountered in using the projected 2D images from the pose-corrected 3D data, we also examined using the original 2D images. 2D recognition is examined with a set of original intensity images that have pose correction only around Z axis, against projected 2D intensity data that has pose correction in X, Y and Z axes. Figure 4 shows the CMC curves for the two types of 2D images. It turns out that the performance of original 2D images with only the correction for 2D rotation around the Z axis is greater than that of the projected 2D images created using the pose-corrected 3D data. The performance using the original 2D images is 83.1% (M = 5 and N= 9) versus 78.9% by the projected 2D images (M = 5 and N= 5). This can be interpreted as an indication that the uncorrected pose variation in the original 2D images is not as damaging to recognition as is the loss of data due to the mask needed when using the projected 2D image. Thus, we decided to use the original 2D images rather than the projected 2D images in later experiments.

4.2. Experimental Results: 2D face versus 3D face in biometrics

This experiment is to investigate the performance of individual 2D eigenface and 3D eigenface methods. The null hypothesis is that there is no significant difference in the recognition rate between 2D or 3D, given (1) the use of the same PCA-based algorithm implementation, (2) the same subject pool represented in training, gallery and probe sets, and (3) the controlled variation in one parameter, time, of image acquisition between the gallery and probe images. After the eigenvectors are tuned, M = 6, and N = 7 vectors are dropped in 3D to create the face spaces. With the given optimal set of eigenvectors in 2D or 3D, the results show that rank-one recognition rate is 83.1%, and 83.7% for 3D (See Figure 5). This difference in rank-one recognition rates is clearly not statistically significant. Thus the results of our experiment do not provide evidence for rejecting the null hypothesis; we do not find a statistically significant difference in accuracy between PCA-based recognition using 2D and 3D face data.

4.3. Experimental Results: Single-modal biometrics versus multi-modal biometrics

This experiment is to investigate the value of a multi-modal biometric using 2D and 3D face images, compared against individual biometrics. The null hypothesis for this experiment is that there is no significant difference in the performance rate between uni-biometrics (2D or 3D alone) and multi-biometrics (both 2D and 3D together). According to Hall [27], a fusion can be usefully done if an individual probability of correct inference is between 50% and 95% with one to seven classifiers. From our second experiment, it is reasonable to fuse the two individual biometrics which meet this fusion criteria. Figure 6 shows the CMC with the

Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG'03)
0-7695-2010-3/03 $ 17.00 © 2003 IEEE

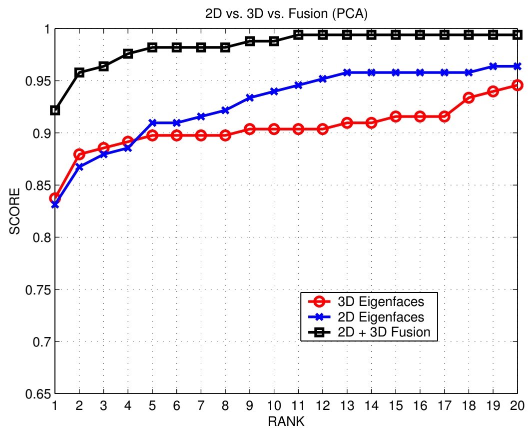


Figure 6: Single- versus multi-modal biometrics.

Table 2: Rank-one recognition rates achieved by different fusion methods

(Transformation)	Sum	Product	Minimum
Linear	92.2%	92.2%	90.1%
Logarithmic	92.2%	92.2%	90.1%
Exponential	91.0%	92.2%	90.1%
Weighted	92.8%	92.2%	90.4%

rank-one recognition rate of 92.8% for the multi-modal biometric, achieved by combining modalities at the distance metric level. Regardless of the particular fusion strategy, the combined 2D-plus-3D performs significantly better than either one alone (See Table 2). Figure 7 shows examples where the multi-biometric was correct when one of the individual biometric failed. A McNemar's test for significance of the difference in accuracy in the rank-one match between the multi-modal biometric and either the 2D face or the 3D face alone shows that multi-modal performance is significantly greater, at the 0.05 level.

5. Summary and Discussion

The value of multi-modal biometrics with 2D intensity and 3D shape of facial data in the context of face recognition is examined. This is the largest experimental study (in terms of number of subjects) that we know of to investigate the comparison and combination of 2D and 3D data for face recognition. In our results, each modality of facial data has roughly similar value as an appearance-based biometric. The combination of the face data from both modalities results in significant improvement over either individual biometric. In general, our results appear to support the conclusion that *the path to higher accuracy and robustness in biometrics involves use of multiple biometrics* rather than the best possible sensor and algorithm for a single biometric. The source of a biometric needs to be carefully examined to obtain complementary sources and the number of biometrics needs to be controlled in the context of data (sensor) fusion. Prior to adding a new modality to existing biometrics, an individual modality needs to be validated throughly so that it has a reasonable correct identification rate. One of the main purposes of sensor fusion is to reduce the ambiguity between domain experts. Thus, without clearly proven benefit, it cannot be expected to necessarily better performance by a newly added dimensionality to the decision domain.

The general quality level of the data in a 3D image collected by current range scanners is perhaps not as good as that of the 2D intensity image taken with current camera technology still. Range scanner technology has problems with missing and noisy data that do not occur with regular camera (CCD) technology. It is possible that the quality of 3D sensor data will improve more rapidly in the near future than will the relatively mature regular camera technology. If this happens, it could improve the usefulness of 3D face data relative to 2D face data.

There may still be some biometrics algorithm, other than PCA, for which one of the 2D face or the 3D face offers statistically significantly better recognition performance than the other. Also, there may be particular application scenarios in which it is not practical to acquire 2D and 3D face images that meet similar quality control conditions.

Even though data have been collected in a controlled external environment, such as lighting or facial expression, there is some degree of limitation that just cannot be controlled, such as slight movement around lips or eye area. This affects the performance rate since it actually changes the shape of face data occurring around the missing area. These problems more severely affect the performance in 3D than they do in 2D.

It is generally accepted that performance estimates for face recognition will be higher when the gallery and probe images are acquired in the same acquisition session, compared to performance when the probe image is acquired after some passage of time [28]. As little as a week's time is enough to cause a substantial degradation in performance [29]. While many performance results reported in the literature are obtained with datasets where the probe and gallery images are acquired in the same session, most envisioned applications for face recognition technology seem to occur in a scenario in which the probe image would be acquired some time after the gallery image. In this context, it is worth noting that the dataset used here incorporates a substantial time lapse between gallery and probe image acquisition.

The results presented in this study suggest that it is worthwhile to investigate biometrics combining multiple types of sources, such as combining 2D (appearance) and 3D (shape) with infrared imagery (thermal pattern). In future research, other fusion schemes will also be considered during the decision process.

The dataset used in the experiments reported here will eventually be made available to other research groups as a part of the Human ID databases. See http://www.nd.edu/~cvrl/ for more information about the dataset and the release agreement.

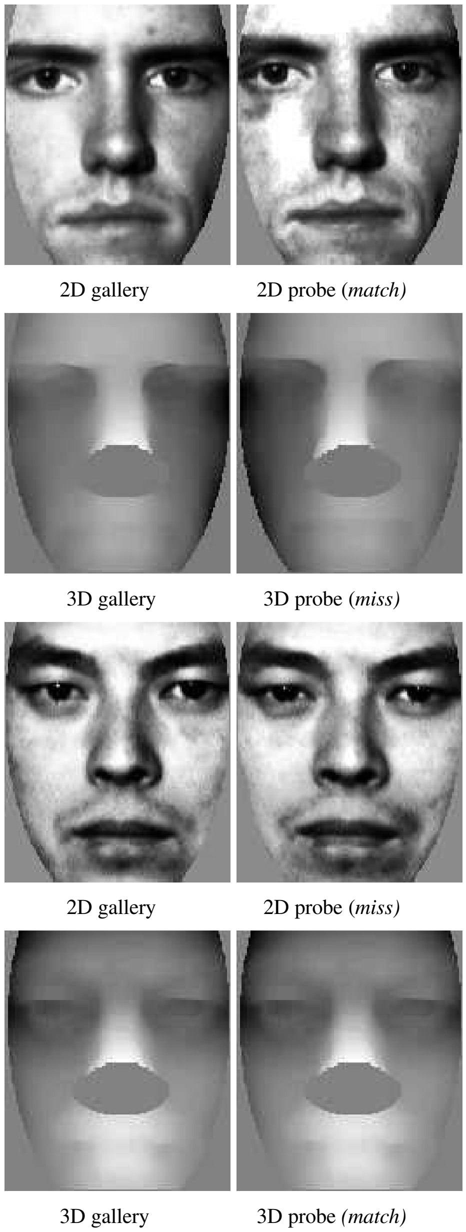


Figure 7: Two examples where multi-biometric corrects individual biometric.

6. Acknowledgments

This work was supported by DARPA through Office of Naval Research N000140210410 and by National Science Foundation EIA 01-30839.

References

[1] W. Zhao, R. Chellappa, A. Rosenfeld, and P. Phillips, "Face recognition: A literature survey," *To appear in ACM Computing Surveys*.

[2] M. Kamel, H. Shen, A. Wong, T. Hong, and R. Campeanu, "Face recognition using perspective invariant features," *Pattern Recognition Letters*, vol. 15, pp. 877–883, 1994.

[3] L. Wiskott, J-M. Fellous, N. Kruger, and C. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. and Mach. Intel.*, vol. 19, no. 7, pp. 775–779, July 1997.

[4] L. Lee, B. Moghaddam, H. Pfister, and R. Machiraju, "Silhouette-based 3D face shape recovery," *Graphics Interface*.

[5] A. Bronstein, M. Bronstein, and R. Kimmel, "Expression-invariant 3D face recognition," *Audio and Video based Biometric Person Authetication*, 2003.

[6] J. Huang, B. Heisele, and V. Blanz, "Component-based face recognition with 3D morphable models," *Audio and Video based Biometric Person Authetication*, 2003.

[7] C. Chua, F. Han, and Y. Ho, "3D human face recognition using point signature," *Int'l Conf. on Automatic Face and Gesture Recognition*, pp. 233–238, 2000.

[8] B. Achermann, X. Jiang, and H. Bunke, "Face recognition using range images," in *Proceedings Int'l Conf. on Virtual Systems and MultiMedia '97, Geneva, Switzerland*, Sept. 1997, pp. 129–136.

[9] Y. Wang, C. Chua, and Y. Ho, "Facial feature detection and face recognition from 2D and 3D images," *Pattern Recognition Letters*, vol. 23, pp. 1191–1202, 2002.

[10] C. Beumier and M. Acheroy, "Automatic face verification from 3D and grey level clues," *11th Portuguese Conf. on Pattern Recognition*, Sept. 2000.

[11] B. Achermann and H. Bunke, "Combination of face classifiers for person identification," *Int'l Conf. on Pattern Recognition*, pp. 416–420, 1996.

[12] G. Shakhnarovich and T. Darrell, "On Probabilistic Combination of Face and Gait Cues for Identification," *Int'l Conf. on Automatic Face and Gesture Recognition*, pp. 169–174, 2002.

[13] K. Chang, K. Bowyer, V. Barnabas, and S. Sarkar, "Comparison and combination of ear and face images in appearance-based biometrics," *To appear in IEEE Trans. Pattern Anal. and Mach. Intel.*, vol. 25, 2003.

[14] N. Poh and J. Korczak, "Hybrid biometric person authentication using face and voice features," *Audio and Video based Person Authentication*, pp. 348–353, June 2001.

[15] S. Ben-Yacoub, "Multi-modal data fusion for person authentication using SVM," *Audio and Video based Person Authentication*, pp. 25–30, 1999.

[16] E. Bigun, J. Bigun, B. Duc, and S. Fischer, "Expert conciliation for multi modal person authentication systems by bayesian statistics," *Audio and Video based Person Authentication*, pp. 311–318, 1997.

[17] R. Brunelli and D. Falavigna, "Person identification using multiple cues," *IEEE Trans. Pattern Anal. and Mach. Intel.*, vol. 17, no. 10, pp. 955–966, 1995.

[18] R. Frischholz and U. Dieckmann, "BioID: A multimodal biometric identification system," *IEEE Computer*, vol. 33, no. 2, pp. 64–68, Feb. 2000.

[19] L. Hong and A. Jain, "Integrating faces and fingerprints for personal identification," *IEEE Trans. Pattern Anal. and Mach. Intel.*, vol. 20, no. 12, pp. 1295–1307, 1998.

[20] A. Ross and A. Jain, "Information fusion in biometrics," *Audio and Video based Person Authentication*, pp. 354–359, 2001.

[21] L. Hong, A. Jain, and S. Pankanti, "Can multibiometrics improve performance?," *Proceedings AutoID'99, Summit, NJ*, pp. 59–64, Oct. 1999.

[22] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[23] R. Beveridge and B. Draper, "Evaluation of face recognition algorithms (release version 4.0)," URL : http://www.cs.colostate.edu/evalfacerec/index.html.

[24] J. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Trans. Pattern Anal. and Mach. Intel.*, vol. 22, no. 10, pp. 1090–1104, October 2000.

[25] W. Yambor, B. Draper, and J. Beveridge, "Analyzing PCA-based face recognition algorithms: Eigenvector selection and distance measures," *2nd Workshop on Empirical Evaluation in Computer Vision, Dublin, Ireland*, July 1 2000.

[26] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. and Mach. Intel.*, vol. 20, no. 3, pp. 226–239, 1998.

[27] D. Hall, *Mathematical Techniques in Multisensor Data Fusion*, Artech House, Norwood, MA., 1992.

[28] J. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and M. Bone, "Facial recognition vendor test 2002: Evaluation report," *NISTIR 6965*, URL: http://www.frvt2002.org/FRVT2002/documents.htm.

[29] P. Flynn, K. Bowyer, and J. Phillips, "Assessment of time dependence in face recognition: an initial study," *Audio and Video based Biometric Person Authetication*, 2003.