Multi-Biometrics 2D and 3D Ear Recognition

Ping Yan and Kevin W. Bowyer[1]

Department of Computer Science and Engineering,
University of Notre Dame, Notre Dame IN 46545, USA,
pyan, kwb@cse.nd.edu

Abstract. Multi-biometric 2D and 3D ear recognition are explored. The data set used represents over 300 persons, each with images acquired on at least two different dates. Among them, 169 persons have images taken on at least four different dates. Based on the results of three algorithms applied on 2D and 3D ear data, various multi-biometric combinations were considered, and all result in improvement over a single biometric. A new fusion rule using the interval distribution between rank 1 and rank 2 outperforms other simple fusion rules. In general, all the approaches perform better with multiple representations of a person.

1 Introduction

Fingerprints, face and iris have received wide attention both in academic research and in the biometrics industry. Fingerprint and iris are considered as generally more accurate than face, but face is more flexible for use in surveillance scenarios. However, face by itself is not yet as accurate and flexible as desired. Ear images can be acquired in a similar manner to face images, and at least one previous study suggests they are comparable in recognition power [1], so additional work on ear biometrics has promise to lead to increased recognition flexibility and power.

Three algorithms have been explored on 2D and 3D ear images, and based on that, three kinds of multi-biometrics are considered: multi-modal, multi-algorithm and multi-instance. Various multi-biometric combinations all result in improvement over a single biometric. Multi-modal 2D PCA together with 3D ICP gives the highest performance. To combine 2D PCA-based and 3D ICP-based ear recognition, a new fusion rule using the interval distribution between rank 1 and rank 2 outperforms other simple combinations. The rank one recognition rate achieves 91.7% with 302 subjects in the gallery. In general, all the approaches perform much better with multiple images used to represent one subject. In our dataset, 169 subjects had 2D and 3D images of the ear acquired on at least four different dates, which allows us to perform multi-instance experiments. The highest rank one recognition rate reaches 97% with the ICP approach used to match a two-image-per-person probe against a two-image-per-person gallery. In addition, we found that different fusion rules perform differently on different combinations. The min rule works well when combing the multiple presentations of one subject, while the sum rule works well when combining multiple modalities.

2 Data Acquisition

All the images used in this paper were acquired at the University of Notre Dame in 2003-2004. In each acquisition session, the subject sat approximately 1.5 meters away from the sensor, with the left side of the face facing the camera. Data was acquired with a Minolta Vivid 910 range scanner. One 640x480 3D scan and one 640 x 480 color image are obtained near simultaneously.

The earliest good image for each of 302 persons was enrolled in the gallery. The gallery is the set of images that a "probe" image is matched against for identification. The latest good image of each person was used as the probe for that person. This results in an average of 4.3 weeks time lapse between the gallery and probe. Including the images for multi-instance experiments, there are a total of 942 pairs of 3D and 2D images used in this work (302+302+169+169). A subset of 202 persons of data was used in initial experiments to explore algorithm options.

3 Algorithms

Three different algorithms have been examined. The PCA (Principle Component Analysis) based approach has been widely used in face recognition [2–4]. In our experiments, a standard PCA based algorithm [5] is used on both 2D and 3D ear data. Based on the observation that edge images of the range image are much cleaner than for the 2D edge images, we develop an edge-based Hausdorff distance method for 3D ear recognition using the range image. Also, Besl and McKay's classic ICP algorithm [6] has been applied on 3D ear data. Approaches considered include a PCA ("eigen-ear") approach with 2D intensity images, achieving 63.8% rank-one recognition; a PCA approach with range images, achieving 55.3%; Hausdorff matching of edge images from range images, achieving 67.5%, and ICP matching of the 3D data, achieving 84.1%. Results of these four single-biometric experiments are represented as CMC curves in Figure 1 [7].

4 Multi-Biometrics

Recently, multi-biometrics have been investigated by several researchers [8–11]. Multi-biometrics can be divided into three simple classes, according to the method of combination. These are multi-modal, multi-algorithm and multi-instance. In general, multi-modal uses different modalities of biometrics, like face, voice, fingerprint, iris and ear of a same subject. Also we consider that for a given biometric, the data from different sensors are one kind of multi-modal, like 2D intensity data and 3D range data. Multi-algorithm uses different algorithms on the same data. For example, we can use both PCA and ICP on 3D ear data. Multi-instance has more than one representation for a given subject. For example, if we took three 2D ear images of the same person on different dates, then the three images together can be treated as a representation of this person.

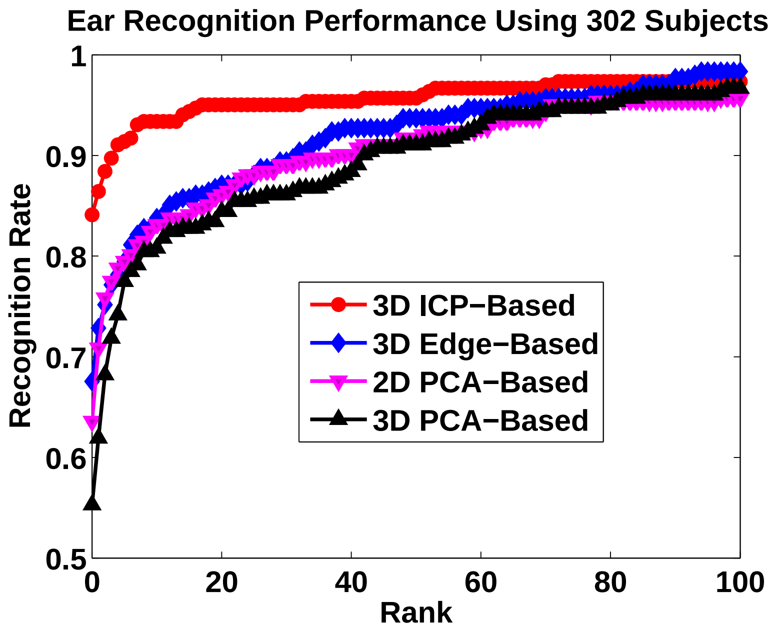


Fig. 1. Performance of Different Approaches

A complex combination can involve more than one kind of multi-biometrics. For example, we can combine 2D PCA and 3D ICP, which includes both multi-modal and multi-algorithm biometrics.

4.1 Fusion Levels and Score Normalization

Each simple biometric has four steps: (1) obtain the data from the sensor, (2) extract the interesting area or features from the raw data, (3) compare the data to a group of enrolled data to obtain the matching score and (4) determine the correct or incorrect matching based on the matching score [12]. Based on these different steps, there are several possible fusion levels. Sensor level fusion combines the raw sensor outputs. Feature extraction level fusion combines multiple extracted features from each biometric. Matching score level fusion combines the matching scores from each biometric. Decision level fusion uses the results from different biometrics and makes the final decision based on all of them.

In our study, the fusion rules work at the matching score level. Since each simple biometric has different meaning, range and distribution of matching scores, score normalization is required in order to combine them. In our experiments, min-max score normalization has been applied on all the results before we do the fusion: $s' = (s - min)/(max - min)$.

4.2 Multi-modal Biometrics

Multi-modal biometrics in this paper refers to the combination of 2D intensity data and the 3D range data. There are three algorithms based on 3D range data, and one on 2D intensity data. Therefore, the combinations include 2D PCA with 3D ICP, 2D PCA with 3D PCA, and 2D PCA with 3D edge-based approach.

First, two simple fusion rules are tried on all three combinations. As shown in Table 1, the sum rule performs much better than the min rule. This result is similar to the conclusion in [11, 4]. Also an advanced sum rule is tested. The rank one matching in each modality is given an additional weight, which measures the distance between itself and the rank two match. The advanced sum rule yields better results than the simple sum rule.

Table 1. Fusion on Multiple Modalities (302 subjects)

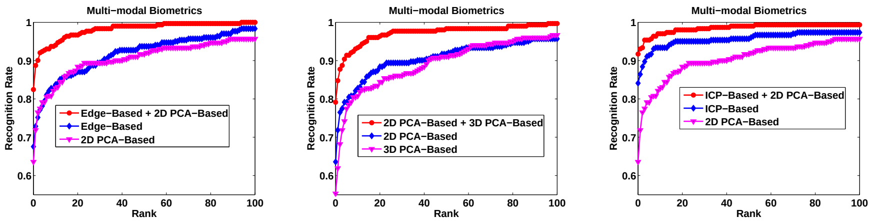
Multi-modals	**MIN**	**Simple SUM**	**Advanced Sum**
2D PCA + 3D ICP	76.4%	81.1%	82.5%
2D PCA + 3D PCA	72.2%	78.8%	79.1%
2D PCA + 3D Edge	73.5%	80.5%	82.5%

The sum rule adds individual matching scores from different matches. Equal weights are assigned to each modality without any bias. However, in general, some modalities have better performance than others. In order to show the bias of several modalities, different weights are assigned to individual modalities. We test the weight assignment by using 202 subjects on 2D PCA combining with 3D ICP. As shown in Table 2, the highest performance is 93.1%, obtained when the weight of ICP is 0.8, and the weight of PCA is 0.2.

Table 2. Different Weights for Fusing the ICP and PCA results

Weight 2D PCA	**Weight 3D ICP**	**Performance (202 Subjects)**	**Performance (302 subjects)**
1	0	71.4%	63.6%
0	1	85.1%	84.1%
0.9	0.1	73.3%	66.9%
0.8	0.2	76.7%	68.9%
0.7	0.3	78.2%	73.8%
0.6	0.4	81.7%	78.5%
0.5	0.5	84.2%	82.5%
0.4	0.6	86.6%	88.7%
0.3	0.7	89.1%	90.7%
0.2	0.8	93.1%	90.4%
0.1	0.9	91.6%	86.4%

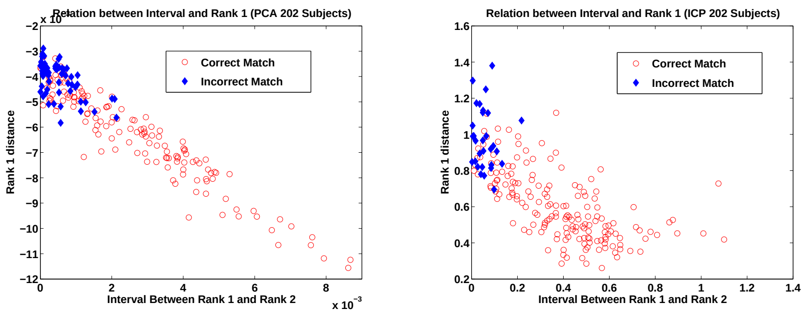
Applying the same weighted sum rule to the other two combinations, the best performance is obtained when there is equal weight for each modality. This is because 2D PCA, 3D PCA and edge-based approaches have similar performance. The rank one recognition is 79.1% when combining 2D PCA and 3D PCA, and it is 82.5% when combining 2D PCA and 3D edge-based algorithm. CMC curve are shown in Figure 2.



(a) (b) (c)

Fig. 2. Multi-modal Biometrics Performance

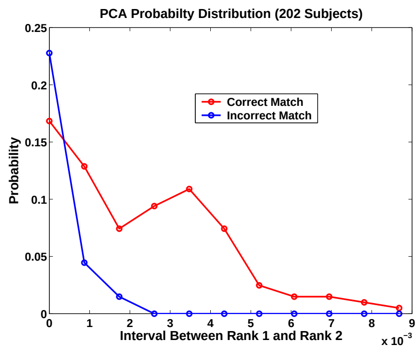
Our third combination rule is based on the analysis of the interval between rank 1 and rank 2 in both PCA and ICP results. Figure 3 shows that the overlap area between the correct matches and incorrect matches is much less in ICP than in PCA, which means that it is easier to use a threshold to separate the correct and incorrect matches in the ICP than in the PCA results.



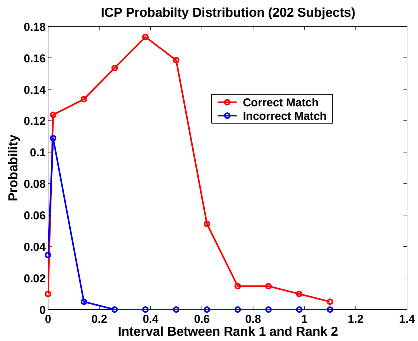
(a) (b)

Fig. 3. Relationship Between Correct Matches and Incorrect Matches

Figure 4 shows the probability distribution of the different intervals between the correct matches and incorrect matches. In general, the greater the gap between the rank 1 and rank 2, the higher the possibility that it is a correct match. When the interval in ICP is greater than 0.2, they are all correct matches. The corresponding value in PCA is 0.002. For both ICP and PCA, we split the interval range into 10 steps. All the interval values are placed into these 10 steps. The percentage of the correct over incorrect matches in each interval step is shown in Table 3.



(a)



(b)

Fig. 4. Interval Distribution Between Correct Matches and Incorrect Matches

Table 3. Fraction of the Correct Match in the Different Interval Level

	1	2	3	4	5	6	7	8	9	10
PCA	0.4250	0.7429	0.8333	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ICP	0.2222	0.5319	0.9999	1.000	1.000	1.000	1.000	1.000	1.000	1.000

When an interval falls into a certain range, we can determine the possibility that it is a correct or incorrect match from Table 3. Using this information we can combine the PCA and ICP in a smarter way. Before the combination, the interval between the rank 1 and rank 2 is computed first for each comparison in the ICP and PCA. Then the corresponding percentage of the correct match and incorrect match is obtained according to Table 3. Using this strategy to combine the PCA and ICP results on 202 subjects, the rank one recognition rate is 93.1%, which is the same as the best results obtained from the simple weight scheme shown in table 2.

Till now, all the results are calculated from 202 subjects. Since the small dataset has a distribution similar to the larger dataset (302 subjects), we predict the distribution of the larger dataset by using the value in Table 3. The rank one recognition rate is 91.7%, which is even better than the results (90.1%) using simple weighted sum scheme. Thus it seems that performance may be increased by using a smart fusion step. However the increase is not statistically significant and this issue deserves further exploration.

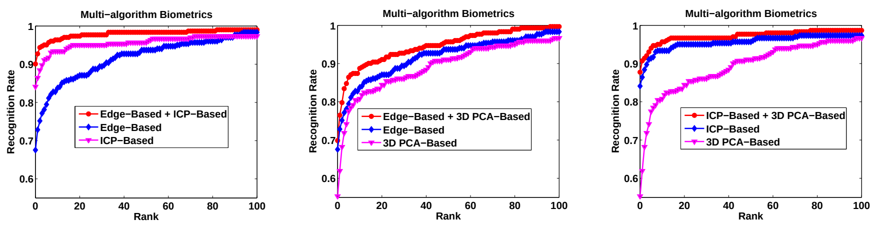
4.3 Multi-algorithm Biometrics

Three different algorithms have been developed to use on the 3D data. These are the ICP-based algorithm, PCA-based algorithm and edge-based algorithm. After score normalization, the weighted sum rule is used for combinations. Rank one recognition rates are demonstrated in Table 4. The best performance is achieved when combining ICP and edge-based algorithm on the 3D data.

Table 4. Multi-algorithm Biometrics Using Weighted Sum Rule

	3DICP	**3DPCA**	**3DEdge**	**Performance**
ICP + PCA	0.90	0.10		87.70%
ICP + Edge	0.80		0.20	90.2%
PCA + Edge		0.40	0.60	69.9%

From Table 1, if we only consider with those not so good performance, like 2D PCA, 3D PCA and 3D edge-based approach, the multi-modal biometrics has better performance than the multi-algorithm.



(a) (b) (c)

Fig. 5. Multi-algorithm Biometrics Performance

4.4 Using Multiple Images to Represent a Person

In general, approaches perform better with a multiple-sample representation of a person, and scale better to larger datasets. We have 169 subjects that have at least 4 good images in both 2D and 3D data. Each pair of 2D and 3D images were taken on a different date. In this section, we will concentrate on the multi-galleries or multi-probes using 2D PCA and 3D ICP algorithms.

For each subject, there are four 2D and 3D images available. We consider three possible multiple-instance representations based on these images. These are (a) 1 in the gallery and the other 3 images in the probe, (b) 2 in the gallery and the other 2 in the probe, and (c) 3 in the gallery and the other 1 in the probe. Two fusion rules, min and sum, are attempted to combine the results, shown in Table 5. It is interesting here that we have multi-instance better than multi-modal. Using one gallery and one probe for these 169 subjects, the rank one rate is 73.4% for 2D PCA, and 81.7% for 3D ICP. Combining the results of 2D PCA and 3D ICP, the best performance obtained is 88.2%.

Table 5. Fusion on Multiple Galleries and Probes (169 Subjects)

	2D PCA		**3D ICP**	
1G1P	73.4%		81.7%	
	MIN	**SUM**	**MIN**	*SUM*
1G3P	82.2%	83.4%	95.3%	81.1%
2G2P	84.0%	87.5%	97.0%	81.7%
3G1P	81.7%	80.5%	91.1%	81.7%

In the multi-galleries and multi-probes experiments, the best performance is achieved when 2 images are put into the gallery and the other 2 put into the probe. This is true in both 2D PCA and 3D ICP algorithms. This combination gives us 4 matches, whereas the other combinations give 3 matches. Also we noticed that the min rule is much more powerful than the sum rule in the 3D ICP performance, while it has similar performance to the sum rule in the 2D PCA performance. We attribute the performance of the min rule to the possibility of minimizing "outliers" in the 3D matching.

Matching of 3D ear images has many sources of "outliers". There can be outlier noise in a given 3D image, such as a "spike" from 3D sensing. Also in matching one 3D image to another, incorrect point correspondences may arise, possibly due to points existing in one scan but not the other. Increasing the number of representations for a certain person in both the gallery and probe gives a better chance to find the correct correspondence between the points. Thus, the performance increases significantly in the ICP experiment.

5 Summary And Discussion

We find that multi-modal, multi-algorithm or multi-instance improve performance over a single biometric. The combination of the 2D PCA and 3D ICP gives the highest performance of any pairs of biometrics considered. Three different multi-biometric combinations were considered. All result in improvement over a single biometric. Among the four single modal ear biometrics, the ICP-based recognition outperforms the other three methods. And it is expected that the best combination includes the ICP as one of the components. Multi-modal with 2D PCA and 3D ICP gives the highest performance. As to the other three not as good methods, multi-modal biometrics turns out to have better performance than the multi-algorithm biometrics.

The fusion experiments on multi-modal, multi-algorithm and multi-instance biometrics yield different results. The sum rule outperforms the min rule on multi-modal and multi-algorithm biometrics, while the min rule performs well on the multi-instance biometrics, especially when using the ICP algorithm. Min rule has the power to reduce the noise from the original data, which is suitable for the application to multi-instance biometrics. The new fusion rule we introduced in combining 2D PCA and 3D ICP is based on analyzing the interval between rank one and rank two. And the performance result is the best of the fusion rules we used.

The multi-modal 3D ICP plus 2D PCA recognition was 87.7% on the 302 person dataset, as listed in Table 4. It is useful to ask how a multi-modal result compares to the multi-instance results for the individual modes. The multi-modal approach represents a person by two images, in both the gallery and as a probe. If we look at the two-image representation in each of the individual imaging modes, we get 87.5% for 2D PCA and 97% for 3D ICP on the subset of 169 of the 302 persons, Table 5. The multi-modal result for this same subset of 169 persons is 88.2%. Thus we find that the multi-modal result barely improves over the two-image 3D ICP result, and that the four-image 3D ICP result for multi-instance is substantially better than the multi-modal result. This is a different relative performance than found by Chang [13] in a study of multi-modal face recognition, where multi-modal 2D + 3D performance was greater than multi-image 2D or multi-image 3D. However, our work differs in several potentially important respects. One is of course that we study ear recognition rather than face recognition. But also, Chang used the same PCA-based approach for both the 2D face and the 3D face recognition, whereas we use an ICP approach for our 3D recognition. This is important because it appears in our results that the ICP-based approach is substantially more powerful than the PCA-based approach for 3D. Another potentially important difference is that in our multi-image results, the two images used to represent a person are taken at different times, at least a week apart. Chang used images from the same acquisition session in his multi-image results. It is quite possible that images taken on different days give a more independent sample, and so better performance.

6 Improved ICP Algorithm

The multi-biometric results presented in previous sections indicate that 3D shape matching with an ICP-based approach has strong potential for ear biometrics. Therefore, after the results in previous sections were completed, considered various refinements to this approach, several of which were incorporated into an improved algorithm. The amount of the ear shape used in the gallery and probe representations was adjusted to reduce interference from the background. An step to remove outlier point matches was added to reduce the effects of incorrect correspondences. Our improved algorithm produces substantially better results. Using the 302-person dataset, with a single 3D ear scan as the gallery enrollment for a person, and a single 3D ear scan as the probe for a person, the new algorithm achieves 98.7% rank-one recognition. This performance from a single modality and algorithm is high enough that a larger and more challenging data set is needed in order to experimentally evaluate its use in possible multi-biometric scenarios. We are currently developing such a dataset for future experiments.

Acknowledgements: This work was supported in part by National Science Foundation EIA 01-30839 and Department of Justice grant 2004-DD-BX-1224.

References

1. Chang, K., Bowyer, K., Barnabas, V.: Comparison and combination of ear and face images in appearance-based biometrics. In: IEEE Trans. Pattern Analysis and Machine Intelligence. Volume 25. (2003) 1160–1165
2. Turk, M., Pentland, A.: Eigenfaces for recognition. In: Journal of Cognitive Neuroscience. Volume 3(1). (1991) 71–86
3. Phillips, P., Moon, H., Rizvi, S., Rauss, P.: The FERET evaluation methodology for face recognition algorithms. In: IEEE Trans. Pattern Analysis and Machine Intelligence. Volume 22(10). (2000) 1090–1104
4. Chang, K., Bowyer, K., Flynn, P.: Face recognition using 2D and 3D facial data. In: Workshop on Multimodal User Authentication. (2003) 25–32
5. Beveridge, R., She, K., Draper, B., Givens, G.: Evaluation of face recognition algorithm (release version 4.0). (In: URL: www.cs.colostate.edu/evalfacerec/index.html)
6. Besl, P., McKay, N.: A method for registration of 3-D shapes. In: IEEE Trans. Pattern Analysis and Machine Intelligence. (1992) 239–256
7. Yan, P., Bowyer, K.W.: Ear biometrics using 2d and 3d images. In: Technical Report TR 2004-31, CSE Department, University of Notre Dame. (2004.)
8. Jain, A., Ross, A.: Multibiometric system. In: Communications of the ACM. Volume 47(1). (2004) 34–40
9. Bigűn, E., J.Bigűn, Fischer, S.S.: Expert conciliation for multi modal person authentication systems using Bayesian statistics. In: Proceedings of the International Conference on Audio and Video-Based Biometric Person Authentication. (Mar. 1997) 291–300
10. Brunelli, R., Falavigna, D.: Person identification using multiple cues. In: IEEE Trans. Pattern Analysis and Machine Intelligence. Volume 12(10). (1995) 955–966
11. Kittler, J., Hatef, M., Duin, R., Matas, J.: On combining classifiers. In: IEEE Trans. Pattern Analysis and Machine Intelligence. Volume 20(3). (1998) 226–239
12. Ross, A., Jain, A.: Information fusion in biometrics. In: Pattern Recognition Letters. Volume 24. (2003) 2115–2125
13. Chang, K., Bowyer, K., Flynn, P.: An evaluation of multi-modal 2d+3d face biometrics. Accepted to appear. In: IEEE Trans. Pattern Analysis and Machine Intelligence. (2005)