3D Signatures for Fast 3D Face Recognition

Chris Boehnen, Tanya Peters, and Patrick J. Flynn
Department of Computer Science and Engineering
University of Notre Dame, USA {cboehnen, tpeters, flynn}@nd.edu

Abstract. We propose a vector representation (called a 3D signature) for 3D face shape in biometrics applications. Elements of the vector correspond to fixed surface points in a face-centered coordinate system. Since the elements are registered to the face, comparisons of vectors to produce match scores can be performed without a probe to gallery alignment step such as an invocation of the iterated closest point (ICP) algorithm in the calculation of each match score. The proposed 3D face recognition method employing the 3D signature ran more than three orders of magnitude faster than a traditional ICP based distance implementation, without sacrificing accuracy. As a result, it is feasible to apply distance based 3D face biometrics to recognition scenarios that, because of computational constraints, may have previously been limited to verification. Our use of more complex shape regions, which is a trivial task with the use of 3D signatures, improves biometric performance over simple spherical cut regions used previously [1]. Experimental results with a large database of 3D images demonstrate the technique and its advantages.

Keywords: Biometrics, 3D Face, Fast, Surface Distance, 3D Signature

1 Introduction

Techniques for biometric recognition of people have applications in electronic commerce, homeland security, access control, and forensics. Much recent research work in biometrics has addressed face recognition, where the reflectance and/or shape of the face is measured and distinguishing features derived for the purpose of identification. 2D face images can potentially be captured quickly, inexpensively, and surreptitiously. 3D face shape has been investigated less deeply than face appearance, but offers potential advantages relating to decreased sensitivity to lighting and intrinsic reflectance variations in faces. Retention of the third dimension may allow for easier partial matching techniques than those needed for 2D face images captured under nonfrontal pose, where the effects of the projection transform must be accommodated.

Many traditional 3D face recognition techniques use Besl and McKay's iterative closest point (ICP) alignment approach [5] to get a distance between 3D face models as the basis for computation of the matching score; the alignment error measured can serve as a measure of match quality [1][2][3] representing the distance between the two faces. However, ICP is iterative and involves a nearest-neighbor computation at each iteration, which can require a long running time, especially when the application is identification and involves the processing of a large subject gallery.

This paper proposes a new face representation (called a 3D signature) which is a vector representation of the surface with vector entries tied to a canonical face coordinate system obtained by a one-time per region alignment operation. The 3D signature representation is thus 'pre-registered' to the surface, obviating the need for an iterative alignment procedure such as ICP to be executed at matching time. This yields a much shorter match computation time than an ICP-based matcher in recognition scenarios. For verification, the running time would remain the same, as the 3D signature generation process would dominate the running time. The 3D signatures are similar to a vector representation used previously with principal component analysis for whole 3D faces that has been previously studied [4], but their construction employs a different search criterion and PCA is not required for our proposed approach. In addition, the proposed representation uses small face regions extracted using arbitrary geometric cuts, made simple with our 3D signature approach, that are more complex than the spherical cuts utilized in prior work (e.g., [1]).

2. Previous Work

The invocation of ICP [5] in each match score computation has been a part of many of the best-performing 3D face matching techniques proposed in the literature. The calculation of N match scores by such a technique requires that ICP be executed N times for a typical running time of $O(Nn_d \log n_m)$. Here, we present a review of the most successful ICP-based 3D face matching techniques. Finally, we review 3D face match score calculation techniques that are not based on ICP, although these techniques may utilize some form of alignment in preprocessing.

2.1 ICP Match-Score 3D Face Biometrics

Several ICP based 3D face matchers have been proposed in the literature. Lu et al. [3] aligned 3D faces utilizing automatically detected feature points. During the computation of every match score, they refined the alignment utilizing ICP. The match score was the root mean squared distance between the two surfaces. Chang et al. [2] experimented with whole face 3D matching by shrinking spherical regions around the nose for improved biometric performance. They first aligned two 3D face models with ICP and provided a Nearest Neighbor matching score.

The best (as measured by rank-one recognition rate) ICP-based 3D face recognition performance at the time this paper was written was reported by Faltemier et al. [1]. Their system's performance matches that of other non-ICP–based techniques [10]. Utilizing multiple regions on the face (obtained using spherical cuts around distinguished surface points), they found that 28 regions combine for the best performance and achieved a 97.2% rank-one rate on the University of Notre Dame Fall 2003 and Spring 2004 3D face datasets (Phillips et al. [7] provide additional information on these datasets). Faltemier et al. [1] showed that the use of a larger number of regions improves performance; they also analyzed the best performance for different number and choices of regions. Choosing the best combination of 8 regions, they were able to achieve a 92.9% rank-one performance. They promote their system for verification but not for recognition because of the running time that their ICP-

based system entails (calculating the match scores for the total of 28 regions that they utilize requires 2.38 seconds, not including the time for preprocessing, result fusion, or file I/O).

2.3 Non-ICP-Based 3D Face Biometrics

The methods discussed in this section may utilize ICP in the overall process as we do; however, the key is that they do not use ICP as part of each match-score computation. So regardless of gallery size, ICP is utilized as part of probe preprocessing and is not utilized for each match-score computation.

Colbry et al. [9] recognized the issue of a need for a faster ICP technique that does not utilize ICP for each match score computation and presented an approach for faster matching. An input face is aligned to a stored model face and a depth map is rendered. After rendering the depth maps various sparse feature representations for analysis such as Gabor wavelets were utilized. These alternate match score computations were accomplished without using ICP during each match score calculation. Although eliminating the use of ICP as a part of each match score did dramatically decrease the running time, they found that it had a negative impact on biometric performance.

Kakadiaris et al. [10] produced depth maps utilizing a spin image. Next, the spin image was converted into a wavelet representation utilizing Haar and pyramid coefficients for comparison. Their approach performed well at 97.3% rank one, with performance comparable to that of Faltemier et al. [1]. Their approach did not require an alignment step, such as ICP, for each match score computation but did utilize ICP in the initial alignment wavelet-generating process. As such, their running times were much faster than those of Faltemier et al. [1] but utilized a distinctly different process. This process is distinct from much of the previous 3D face recognition approaches in that it is not based upon a 3D distance metric.

Russ et al. [4] proposed using a 3D PCA-based approach to calculate a unique 3D face vector that was used with PCA to perform recognition. The vectors were produced utilizing a combination of nearest neighbor and normal searches in order to ensure a complete vector. Next, the vectors were fed into principal component analysis to produce PCA coefficients utilized for recognition.

3. 3D SIGNATURE GENERATION FOR FACE RECOGNITION

In this section, we will describe the process of generating 3D signatures. We employ a gallery $G = \{ X_1, X_2, \ldots X_{nG}\}$ of 3D face images captured for the purpose of enrollment, and a training set T of 3D images from fifty subjects not in T is used to bootstrap the generation of two generations of reference faces. Each 3D image is assumed to be roughly aligned to the same location and orientation.

3.1 Generation of Initial Reference Face

An initial reference face F_1 was generated by aligning, smoothing, and merging 3D images of ten distinct faces in T. We employed the Raindrop Geomagic [12]

commercial software for this step. The reference face defines a canonical coordinate system. The reference face generated from this process can be seen in Figure 1.

A 3D signature is a vector representation $S = [\ p_1, \ldots p_n]$ (with p_n representing 3D points) of the 3D surface constructed from a structured traversal of face positions defined on a reference region.

Figure 1: Initial Reference Face Generated

Figure 2: Complete Reference Face

After computation of the initial reference faces, a 3D signature is computed for each of the fifty face images in T using the following process.

- The reference face and the training face image T are aligned using ICP, with an initial coarse correspondence obtained from nose alignment.
- An indexed sequence of locations $S = \{p_1, \ldots p_{nS}\}$, $p_i = (x_i, y_i, z_i)$ consisting of all points on the initial reference face was defined using its canonical coordinate system. The initial reference face model consists of 15,000 points. We show this initial face in Figure 1. The initial distribution of points was made approximately uniform utilizing a uniform point sampling function in Raindrop Geomagic [12].
- Images of fifty images X_j in T were aligned with F_1 using ICP.
- *Calculation of 3D Signature*: Each of the aligned faces was sampled by searching along the normal vector on F_1 at each point p_i in S, retrieving the closest intersecting point on X_j. This yields an ordered set of 3D points on the aligned version of X_i, and a set of fifty correspondents $\{p_l, \ldots p_{50}\}$ for each p_i.
- The correspondents $\{p_1, \ldots p_{50}\}$ for each point p_i are averaged, yielding an average 3D signature which also serves as a second generation reference face. This face model is shown in Figure 2. We cropped the eye and nose region from this representation because the data from this area are typically noisy and contain holes due to sensor limitations.

To determine individual reference regions, the averaged reference face was manually cropped into different regions using Raindrop Geomagic [11]. This is done once during the initial reference region signature generation. Using this method, we produced eight random regions, as are shown in Figure 4. Regions 5 and 6 are subsets of region 2. However, the resulting 3D signatures are not simply subsets. The reason for this is that each reference region is aligned separately. The different alignments result in different 3D signatures. Each region utilized here yields an increase in the performance of the fusion result.

Selection of the reference region shape and point density plays a critical role in 3D signature generation. However, determining the ideal region shape and density to utilize is a problem that has not been addressed and most likely deserves more attention. Our method of reference-region generation has three main advantages over traditional hard-coded region-generation approaches.

- Ease of producing odd-shaped regions
- Ability to increase or decrease the number of points for comparison, regardless of the initial point density of the model
- Ability to control the point density for different parts of the same region

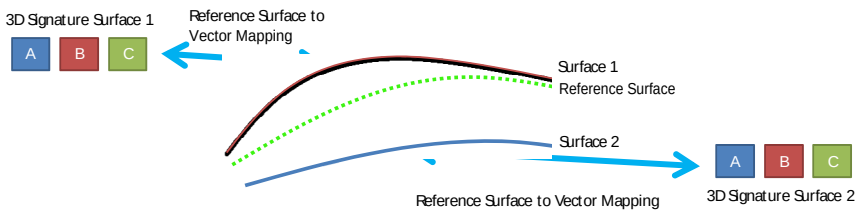

Figure 3: 3D signature correspondence for two surfaces: Green dotted surface is the reference surface, and the red and blue surfaces are input surfaces; the double arrows represent correspondence

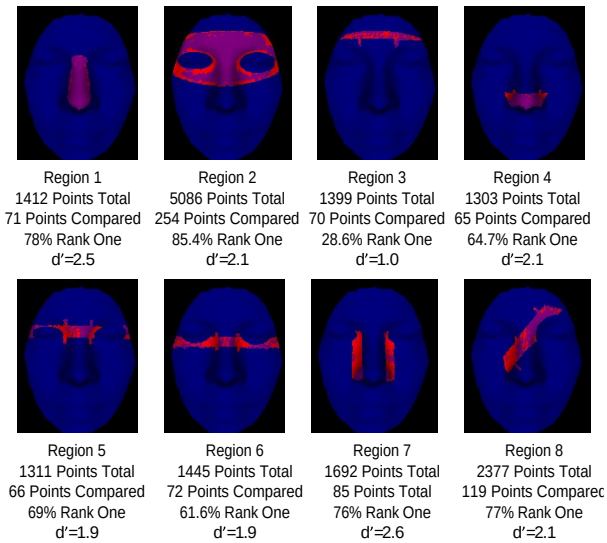

Figure 4: 8 Reference Regions

The ability to customize region shape definition and point density has traditionally required more complex programming as the regions are defined via hard coding in methods such as Faltemier et al. [1]. Faltemier et al. [1] employed hard coded

spherical regions (since they are relatively simple to define) and utilized the point densities of the existing gallery and probe face models as a basis for comparison. Our approach offers the ability to tune region shapes and point density to highly discriminatory face features by modifying the reference face regions.

3.2 Correspondence Search

Given a pair of 3D surfaces, one a reference region and the other an image of an unknown face, it is possible to determine 3D signature correspondence via a nearest neighbor (NN) search, a normal vector search, a hybrid of these two approaches, or a different correspondence method. The notion critical to a 3D signature is that a method of determining correspondence between two points is used.

The nearest-neighbor search assigns each point in the unknown face signature to the closest correspondent on the unknown face to that of the point in the reference face. Previous work [4] has documented that the NN approach does not pick the most representative points when areas of curvature differ between the reference and input surface.

A normal search searches along the normal vector to the reference region surface at each reference region point to find the intersection between the normal and the unknown face surface which is then the corresponded point used for the 3D signature. As such, it is possible that the reference region normal does not intersect the unknown face surface due to surface shape or holes in the 3D surface.

A hybrid technique proposed by Russ et al. [4] repairs missed correspondences by using a NN search to produce correspondences for probe points where the normal search fails. This is necessary for use with PCA as missing vector entries cannot be tolerated.

3.3 Individual Region Match Score Calculation

Calculation of match scores is simple when two 3D signatures are compared. No ICP step is needed during the match score calculation (as in [1][2][3]) since the signatures are pre-aligned. In our method, we calculate the average squared Euclidean distance between the vectors by accumulating the squared distance between each corresponding point pair. When utilizing a normal only search criteria we can ignore any points that lack a correspondent due to a failed normal search. No iterative techniques are used.

3.4 Score Fusion

After obtaining match scores from multiple regions, it is necessary to utilize all of the results to determine identity in a process commonly referred to as fusion. The most commonly used fusion methods include sum, product, minimum, and consensus. The concept of combining independent match scores into one result via fusion was originally proposed by Kittler et al [12].

Faltemier et al. [1] found optimal performance with a modified Borda count technique operating at the rank level. Each rank value was compared to a threshold. Ranks higher than the threshold were replaced by the threshold value and then all ranks were summed. This modified Borda approach performed 4.5% better than sum.

All rank-based fusion techniques, lose the proximity of match scores between in rank computation. We propose a new method named the Match Sum Thresholding (MST) approach. MST operates on the set of scores resulting from multiple 3D signature comparisons. As with the modified Borda count technique, MST replaces any region scores that exceed a predetermined matching score threshold with the value of that threshold.

The threshold used in the MST method is tuned for each 3D signature. Determining an individual MST score threshold for each region is necessary, because match score ranges vary from region to region. We precomputed the MST thresholds using the following approach. We limited the possible values of the thresholds to .01 increments searching for a biometric performance local maxima.

```
1.  Set initial score thresholds to include all match scores
2.  Continue until score thresholds do not change or change falls
    below a set threshold
    a.  For each region
      i.   For a range of threshold values (such as   i=0; i<2;
           i+=.01 )
         1.  Find best threshold for current region based upon
             other region thresholds and MST calculation
         2.  Set threshold for current region to best
```

4. RESULTS

Our experiments used a subset of the FRGC 2.0 3D face data corpus [8]. Images were acquired at the University of Notre Dame using a Minolta Vivid 900 3D scanner. We used all subjects with two or more 3D face images, yielding a dataset comprising 410 subjects and 3939 total images, and conducted closed-set recognition trials. The data contains a variety of facial expressions. All experiments were performed with the earliest image as the gallery image and subsequent images as probes. This standard experiment has been commonly utilized in previous work and is referred to as Experiment 2 in the FRGC protocol. For more information on this dataset or experiment, please refer to Philips et al. [8].

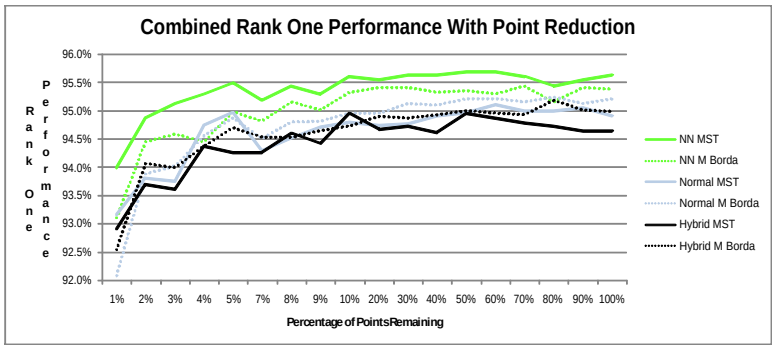

Figure 6: Performance of Hybrid, Nearest Neighbor, and Normal 3D Signatures

4.1 3D Signature Search Method

Previous work [4] utilized a hybrid normal/NN search method to produce vectors for PCA input. We compared the performance of a hybrid approach, NN approach, and normal approach for generating vectors, for a range of 3D signature sizes between 1% and 100% of the original size. The rank one recognition rate recognition results are shown in Figure 6 with different correspondence methods and the MST fusion method.

Contrary to Russ et al. [4], we found that the NN (nearest neighbor) search criteria for 3D signature generation performed best. The difference in performance between the NN approach and the previously utilized hybrid search is even greater (approximately 1%), depending upon the number of points utilized for match score computation. It is possible that the use of individual regions as opposed to entire faces (as used in [4]) explains this difference.

4.2 Fusion Method

The previous highest performing fusion method for multiple regions was proposed by Faltemier et al. [1]. They utilized a modified Borda count to combine the results from each region for optimal performance. Their search criteria was a nearest neighbor approach. As seen in Figure 6, we found that the best-performing 3D-signature search criterion was the NN search method and examine only that correspondence method in this section.

TABLE 1: THRESHOLD VALUES USED FOR MST

	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6	Region 7	Region 8
Region Thresh	0.55	1.2	0.4	0.6	0.85	0.8	0.85	0.8

Furthermore, our new MST fusion approach consistently outperformed the modified Borda approach by approximately 0.2% to 0.5%. The thresholds utilized for each region are shown in Table 1. We believe this demonstrates an advantage in this situation of using the matching score instead of the rank alone. The remainder of the experiments will be presented on the NN search criteria with MST fusion because those criteria obtained superior performance.

4.3 Running Time and Memory Requirements

In addition to a performance accuracy improvement, the 3D signature comparison technique yielded a decrease of nearly four orders of magnitude in matching execution time compared to the matcher proposed by Faltemier et al., due to the lack of an ICP step in each match pair comparison. The 3D-signature match-score computation running time and memory requirements are proportional to the total number of points in the 3D signature utilized for comparison.

Our test scenario utilized a single core of a dual-core laptop. The laptop utilized an Intel T7500 mobile processor operating at 2.2 GHz running a 32-bit version of Windows Vista. We found that on average we were able to compute 102,594 match scores per second for a single region. By comparison, Faltemier et al. [1] were able to

compute 12 match scores per second on average for a single region. This makes the proposed approach approximately 8700 times faster.

Non-ICP–based 3D matchers, such as that described by Kakadiaris et al. [10], have been able to achieve speed improvements similar to those seen here. However, their approach is distinctly different from our 3D-signature approach in that it [10] does not utilize a surface distance comparison (its match scores are based upon wavelet representations). Traditionally the distance based match score approach as used here has been more utilized in the literature. The advantages or disadvantages of this type of an approach or the wavelet approach by Kakadiaris et al. [10] is an open research area.

The proposed 3D signature approach requires the storage of 802 3D points when 95% of the original signatures' points are removed by subsampling. This requires 2406 vertices (x, y, z) to be stored, or approximately 32 kilobytes per gallery image. By comparison, a densely sampled 3D face model of 20,000 vertices would require in excess of 240 kilobytes for vertex storage alone.

5. Conclusion

We have proposed to represent a 3D surface as a 3D signature that can be directly compared to another 3D signature in a process similar to ICP match-based distance biometrics. Both approaches represent match score as an error distance between two surfaces. However, our approach eliminates the need for ICP as a part of the match-score computation.

This process has advantages over other consistent representations, such as depth maps, because a 3D signature is not a 2D projection of the 3D surface; it is, instead, dependent upon a 3D reference region retaining the resulting 3D shape in the new representation. This allows for the data to be represented in a face based coordinate system that does not suffer any data loss as a projection to a depth map does.

The process of computing the signature and performing the comparison presented performance and accuracy advantages over an ICP based approach. The lack of an ICP invocation in each match yields a dramatic speed increase without encountering an accuracy penalty. We were able to calculate match scores 8,700 times faster than previous state of the art ICP based methods [1]. Each match score is computed by calculating the distance between two 3D signatures that are stored in vector form. Previous exploration of vector representations [4] utilized a hybrid normal search/NN approach. Here we show that a purely NN approach achieves improved performance.

In addition, we present a new region fusion approach named MST, which outperforms the modified Borda count approach [1], which was previously the state-of-the-art fusion method for region based 3D face biometrics. MST does not require computing match rank, as does the modified Borda count approach. The superior performance of MST is a result of it retaining the distance between match scores. Utilizing more complex comparison regions and improved fusion method we are able to achieve a 95.5% rank one identification utilizing 8 regions as opposed to 92.9% using 8 regions from the previous state of the art method [1].

In closing, the proposed 3D-signature method outperforms previous state-of-the-art ICP match score-based biometrics [1][2][3] in running time and biometric

performance through the use of regions with more complex shapes and makes 3D face biometrics feasible for large applications with the increased speed and low memory requirements. In the future, we would like to further explore reference regions with a larger variety of shapes and densities. Although the regions here show the potential of complex regions, further exploration may yield further improvements in performance.

Acknowledgements

Biometrics research at the University of Notre Dame is supported by the Central Intelligence Agency, the National Science Foundation under grant CNS01-30839, by the US Department of Justice/National Institute for Justice under grant 2006-IJ-CX-K041, by iARPA and the Biometrics Task Force under US Army contract W91CRB-08-C-0093 and by UNISYS Corp.

References

1. Timothy Faltemier, Kevin W. Bowyer and Patrick J. Flynn,, "A Region Ensemble for 3D Face Recognition," *IEEE Transactions on Information Forensics and Security*, vol. 3, iss. 1, pp.62-73, 2008.
2. Kyong I. Chang, Kevin W. Bowyer, and Patrick J. Flynn, "Multiple Nose Region Matching for 3D Face Recognition under varying facial expression." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1695-1700, 2006.
3. Xiaoguang Lu, Dirk Colbry, and Anil K. Jain. "Three-Dimensional Model Based Face Recognition" *17th International Conference on Pattern Recognition*, 2004
4. T. Russ, C. Boehnen, and T.Peters. 2006. "3D Face Recognition Using 3D Alignment for PCA." *Computer Vision and Pattern Recognition*, New York, pp. 1391-1398, 2006
5. P. Besl and N. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 239– 256, 1992.
6. Bentley, J. L. 1990. K-d Trees for Semidynamic Point Sets. SCG '90: *Proc. 6th Annual Symposium on Computational Geometry* (1990), 187–197
7. P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 947–954, 2005.
8. D. Colbry, O. Folarin, and G. Stockman. "3D Face Identification—Experiments Towards a Large Gallery." *Society of Photographic Instrumentation Engineers Defense & Security*, 2008.
9. I.A. Kakadiaris, G. Passalis, G. Toderici, N. Murtuza, and T. Theoharis. "3D Face Recognition". *Proceedings of the British Machine Vision Conferenc*e, pp. 200-208, 2006
10. F.R. Al-Osaimi, M. Bennamoun, A. Mian, "Integration of local and global geometrical cues for 3D face recognition," *Pattern Recognition Letters*, pp. 1030-1040, 2008
11. http://www.raindropgeomagic.com/en/
12. Kittler, J.V., Matas, J., Jonsson, K., Sanchez, M.U.R., "Combining Evidence in Personal Identity Verification Systems," *Pattern Recognition Letters*, pp. 845-852, 1997