3D Face Recognition

I. Kakadiaris[†] G. Passalis[†,‡] G. Toderici[†] N. Murtuza[†] T. Theoharis[†,‡]
[†] CBL, Dept. of Computer Science, University of Houston, TX, USA
[‡] CGL, Dept. of Informatics, University of Athens, Greece
http://www.cbl.uh.edu/

Abstract

In this paper, we present a new 3D face recognition approach. Full automation is provided through the use of advanced multi-stage alignment algorithms, resilience to facial expressions by employing a deformable model framework, and invariance to 3D capture devices through suitable preprocessing steps. In addition, scalability in both time and space is achieved by converting 3D facial scans into compact wavelet metadata. We present results on the largest known, and now publicly-available, Face Recognition Grand Challenge 3D facial database consisting of several thousand scans. To the best of our knowledge, our approach has achieved the highest accuracy on this dataset.

1 Introduction

Among the many biometric identification modalities proposed for verification and identification purposes, face recognition is high in the list of subject preference, mainly because of its non-intrusive nature. However, from the operator's point of view, face recognition faces some significant challenges that hampers its widespread adoption. Accuracy is the most important of these challenges. Current 2D face recognition systems can be confounded by differences in pose, lighting, expressions and other characteristics that can vary between captures of a human face. This issue becomes more significant when the subject has incentives *not* to be recognized (i.e., non-cooperative subjects).

It is now widely accepted that in order to address the challenge of accuracy, different capture modalities (such as 3D or infrared) and/or multiple instances of subjects (in the form of multiple still captures or video) must be employed [5]. However, the introduction of new capture modalities brings new challenges for a field-deployable system. The challenges of 3D face recognition, which concern the current paper, are as follows: *1) Accuracy Gain:* A significant accuracy gain compared to 2D face recognition systems must result to justify the introduction of a 3D system, either for sole use or in fusion with other modalities. *2) Efficiency:* 3D capture creates larger data files per subject which implies significant storage requirements and slower processing. The conversion of raw 3D data to efficient meta-data must thus be addressed. *3) Automation:* A field-deployable system must be able to function fully automatically. It is therefore not acceptable to assume user intervention for locating key landmarks in a 3D facial scan. *4) Capture Devices:* 3D capture devices were primarily developed for medical and other low-volume applications and suffer from a number of drawbacks when applied to face recognition. These include artifacts, small depth of field, long acquisition time, multiple types of output, and high

BMVC 2006 doi:10.5244/C.20.89

price. *5) Testing Databases:* There is a lack of large and widely accepted databases for objectively testing the performance of 3D face recognition systems.

In this paper, we address the major challenges of a 3D field-deployable face recognition system. We have developed a *fully automatic* system which uses a composite alignment algorithm to register 3D facial scans with a 3D facial model, thus achieving complete *pose-invariance*. Our system employs a deformable model framework to fit the 3D facial model to the aligned 3D facial scans, and in so doing measures the difference between the facial scan and the model in a way that achieves a high degree of *expression invariance* and thus *high accuracy*. The 3D differences (the deformed facial model) are converted to a 2D geometry image and then transformed to the wavelet domain; it has been observed that a small portion of the wavelet data is sufficient to accurately describe a 3D facial scan, thus achieving the *efficiency* goal. Some of the issues of 3D *capture devices* were addressed; specifically, artifacts were handled by median cut and smoothing filters.

Our contributions are the following: 1) We develop a multistage hybrid mesh alignment algorithm to initialize the model fitting process; 2) develop a parameterized, annotated, anthropometrically-based, subdivision deformable face model which is used for identifying the facial features from the raw 3D data; 3) introduce a novel distance metric based on the wavelet representation of the geometry image and normal map corresponding to the subdivided and deformed face model. The result of these contributions is a system for 3D face recognition that achieves the highest accuracy on the FRGC v2 database [19].

This paper is organized as follows: Section 2 reviews related work, while Section 3 describes the methods utilized by our approach. Section 4 presents a performance evaluation using extensive and publicly available databases, while Section 5 summarizes our approach.

2 Related Work

Despite the introduction of commercial grade 2D face recognition systems, their performance remains unreliable. Extensive experiments conducted using the FERET dataset [20] and during the Face Recognition Vendor Test (FRVT) 2002 study indicate that the success rate is not sufficient for critical applications. It appears that 2D face recognition techniques have exhausted their potential as they stumble on inherent problems of their modality (mainly pose and illumination differences).

Given the shortcomings of the pure 2D approaches, a number of 3D and 3D+2D multimodal approaches have recently been proposed. Bronstein *et al.* [2] use a canonical-forms approach to represent 3D faces. They present promising results on a proprietary database. Blanz and Vetter [4] describe an approach for 2D face recognition using a 3D morphable model. The results presented are also promising, but they also use a proprietary database. An excellent recent survey of the field is given by Bowyer *et al.* [5]. Due to the lack of available 3D databases, the majority of these approaches have not been extensively tested. To address this issue, NIST introduced the Face Recognition Grand Challenge (FRGC) and FRVT 2006 [1], and made two multimodal databases publicly available. FRGC v2 [19] includes over 4000 scans with facial expressions. In this section, we present here a small sample of relevant work that is not meant to be exhaustive, and is targeted on the approaches that utilize this database. The performance metrics used for evaluation are described in Section 4.

On the extensive FRGC v2 database, Chang *et al.* [7, 6] examined the effects of facial expressions using two different 3D recognition algorithms. They reported a 92% rank-one recognition rate. Husken *et al.* [11] presented a multimodal approach that uses hierarchical graph matching (HGM). They extended their HGM approach from 2D to 3D but the reported 3D performance is lower than the 2D equivalent. Their fusion, however, offers competitive results, 96.8% verification rate at 0.001 false acceptance rate (FAR), compared to 86.9% for the 3D only. Maurer *et al.* [18] also presented a multimodal approach tested on the FRGC v2 database and reported a 87% verification rate at 0.01 FAR.

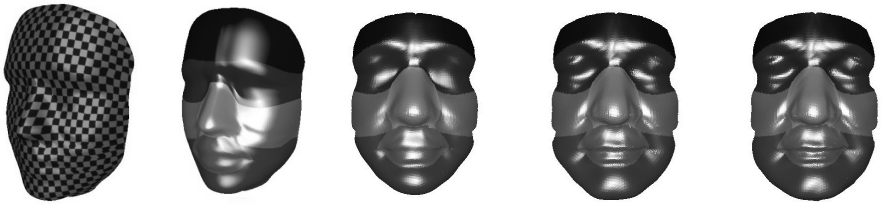
3 Methods

The main idea behind our approach is to annotate facial data using a deformed facial model. The deformed model captures the details of an individual's face and represents this 3D geometry information in an efficient 2D structure by utilizing the model's parameterization. This structure is analyzed in the wavelet domain and the spectral coefficients define the metadata that are used for comparison among different subjects. The physical (geometrical) modeling of the human face allows greater flexibility, better understanding of the face recognition issues, and requires minimum training compared to a statistical modeling approach.

Our face recognition procedure can be divided in two phases, enrollment and authentication. **Enrollment:** Raw data are converted to metadata and stored in the database as follows: *Acquisition:* Raw data are acquired from the sensor and converted to a 3D polygonal representation using sensor-dependent preprocessing. *Alignment:* The data are aligned into a unified coordinate system using a scheme that combines three different alignment algorithms. *Deformable Model Fitting:* An Annotated Face Model (AFM) is fitted to the data. *Geometry Image Analysis:* Geometry and normal map images are derived from the fitted model and wavelet analysis is applied to extract a reduced coefficient set as metadata. **Authentication:** Metadata retrieved from the database are directly compared using a distance metric.

Data Preprocessing: The purpose of preprocessing is to minimize the impact of the quality of input data. The preprocessing consists of the following filters, executed in the given order: *1) Median Cut:* This filter is applied to remove spikes from the data. Spikes are more common in laser range scanners, therefore stronger filtering is needed in this case. *2) Hole Filling:* Laser scanners usually produce holes in certain areas (e.g., eyes, eyebrows) thus a hole filling procedure is applied. *3) Subsampling:* The deformable model fitting effectively resamples the data, making the method insensitive to data resolution. Therefore, the resolution is decreased to gain efficiency to a level that does not sacrifice accuracy.

Annotated Face Model: Our approach utilizes an AFM, which needs to be constructed only once using conventional graphics modeling techniques. This model is an integral part of our approach since it is used in alignment, is deformed in the fitting stage, and its properties allow the creation of the metadata. Based on Farkas' work [9] we ensured that the model is anthropometrically correct. Using information from facial physiology, we have annotated the AFM into different areas. Figure 1(a) depicts our annotation with each area denoted by a different shade. Finally, we applied a continuous global two dimensional parameterization on the model, as depicted in Figure 1(b). The injective



(a) (b) (c) (d) (e)

Figure 1: (a) AFM Parameterization. (b) Annotated facial areas and (c) Fitting progress: AFM after 8, (d) 32, and (e) 64 iterations.

property of the specific parameterization allows us to map all vertices of the model's surface from $\mathbb{R}^3$ to $\mathbb{R}^2$ and vice versa. Therefore we can define the model both in $\mathbb{R}^3$ as polygonal data and in $\mathbb{R}^2$ as a geometry image [10, 13, 14]. A geometry image is a regularly sampled 2D image that has three channels, each one encoding geometric information (X, Y and Z components of a vertex in $\mathbb{R}^3$). Since neighboring information is retained in the geometry image, each element (geometry pixel) can form a triangle in $\mathbb{R}^3$ with two of its neighboring elements, thus allowing an easy transition back to the polygonal representation.

Alignment: Alignment is a key part of any geometric approach. In fact, a misalignment error cannot be rectified in the later steps of this or other similar approaches. To this end, we present a novel multistage alignment method that offers robust and accurate alignment even in the presence of facial expressions.

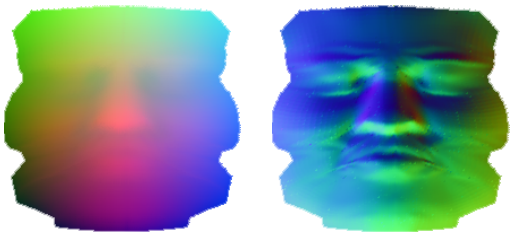
Specifically, we align each new raw dataset with the AFM before the fitting process starts. The alignment computes a rigid transformation that includes rotation and translation. The multistage alignment method consists of three algorithmic steps. Each step uses as input the output of the previous one; early steps offer greater resilience to local minimums while later steps offer greater alignment accuracy:

Spin Images: The purpose of the first step is to establish a plausible initial correspondence between the model and the data. If we do not expect arbitrary rotations and translations in the database this step can be omitted. We utilize the spin image algorithm presented by Johnson [12].

Iterative Closest Point (ICP): The main step of our alignment pipeline uses the ICP algorithm [3], extended in a number of ways. By exploiting the model's annotation we assign per-area weights, and compute a weighted least squares solution for the rigid transformation. Additionally, pairs containing points on surface boundaries are rejected [24]. This ensures that no residual error is introduced into ICP's metric from the non overlapping parts of two surfaces. Finally, if the obtained transformation is not satisfactory we have an option of running the trimmed ICP algorithm [8].

Simulated Annealing on Z-Buffers: This is a refinement step that ensures that the model and data are correctly aligned. We apply the global optimization technique known as Enhanced Simulated Annealing [22] on the difference of the Z-buffers of the model and the data.

Deformable Model Fitting: The AFM is fitted to the data in order to capture the



(a) (b)

Figure 2: (a) Geometry and (b) normal map images of a subject's upper face area. The X, Y and Z vertex components mapped to blue, green and red color components for visualization.

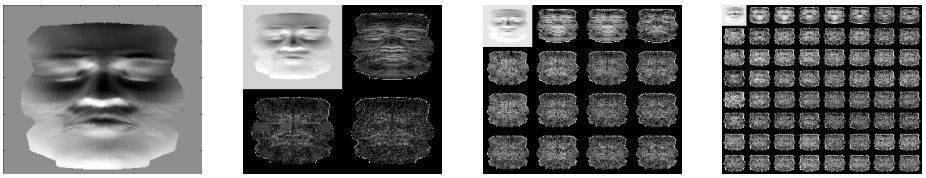
geometric information of an individual's face. We utilize the deformable model fitting framework. The analytical formulation is given by: $M_q \frac{d^2q}{dt^2} + D_q \frac{dq}{dt} + K_q q = f_q$, where M_q is the mass matrix, D_q is the damping matrix, K_q is the stiffness matrix, and f_q are the external forces. The external forces drive the deformation, the stiffness matrix controls the resistance to the deformation and the mass and damping matrix control the velocity and acceleration of the vertices respectively. This equation is solved iteratively using a finite element method (FEM) approximation. An example of the fitting progress for subsequent iterations is presented in Figure 1(c).

Mandal *et al.* [16, 17] combined the deformable model idea with subdivision surfaces. Compared to parametric surfaces, subdivision surfaces offer increased flexibility and computational efficiency. We constructed a subdivision surface based on our AFM using Loop's scheme [15]. The analytical formulation remains unchanged whereas the FEM implementation is integrated with the subdivision surface computation. This allows us to solve the equation at a specified resolution (limit surface) and simultaneously apply the computed forces back in the low resolution model (control mesh). For an AFM with n vertices and m triangles we have $3n$ degrees of freedom in the control mesh and $4^l m$ finite elements (where l is the level of the limit surface).

During fitting, the polygonal data of an individual's face act as attractors for the vertices of the subdivision surface, thus driving the deformation. This involves several nearest neighbor searches at each iteration. For efficiency reasons, we employ an octree, thus lowering the average cost of a search from $O(k)$ to $O(\log k)$, where k is the number of triangles of the facial data.

Geometry Image Analysis: The deformed model that is the output of the fitting process is converted to a geometry image, as depicted in Figure 2(a). The geometry image regularly samples the deformed model's surface and encodes this information on a 2D grid. The grid resolution is correlated with the resolution of the AFM's subdivision surface. From the geometry image a normal map image is constructed, as depicted in Figure 2(b).

We divide the geometry and normal map images into their three components (X,Y and Z) and treat them as separate images. Each component is analyzed using a transform and the coefficients are stored. We use two different transforms thus obtaining two sets of wavelet coefficients: the Haar and the pyramid, with the latter being significantly



(a) (b) (c) (d)

Figure 3: Haar wavelet analysis for the normal map image from Figure 2 (displayed with contrast stretch for better visibility on paper): (a) zero level, (b) first level, (c) second level, (d) third level.

more computationally expensive. We apply the Haar transform on the combined normal/geometry images and the pyramid transform on the geometry images.

Haar Transform: The first transform is a decimated wavelet decomposition using tensor products of the full Walsh wavelet packet system [23] (Figure 3). The 1D Walsh wavelet packet system is constructed by repeated application of the Haar filter bank. Both channels output the result of convolving a 1D discrete signal with a Haar filter and then subsampling by a factor of two. The low–pass and high–pass Haar filters are g and h, respectively: $g = \frac{1}{\sqrt{2}}[1 \quad 1]$ and $h = \frac{1}{\sqrt{2}}[1 \quad -1]$. For images, we use tensor products of these 1D filters, resulting in a four channel filter bank with channels LL, LH, HL, and HH (corresponding to the filters $g^t * g, g^t * h, h^t * g$ and $h^t * h$ respectively). Channel LL (low–pass) captures the local averages of the image, and channels LH, HL and HH (high–pass) capture horizontal, vertical and diagonal edges, respectively. We recursively apply this decomposition to each of the four output channels to construct the full Walsh wavelet packet tree decomposition. We store a subset of the coefficients produced by the decomposition for efficiency reasons. The same coefficients are stored for every image which allows two wavelet decompositions to be compared directly without reconstructing the images.

Pyramid Transform: The second transform decomposes the images using the complex version of the steerable pyramid transform [21], a linear multi-scaled, multi-orientation image decomposition algorithm. The resultant wavelet representation is translation invariant and rotation-invariant. This feature is desirable to address possible positional and rotational displacements caused by facial expressions. Our algorithm applies a 3-scale, 10-orientation complex steerable pyramid transform to decompose each channel of the deformation image. Only the low-pass orientation subbands at the furthest scale are stored. This enables us to compare two images directly and robustly using multiple orientations.

Distance Metrics: In the authentication phase, the comparison between two subjects (gallery and probe), is performed in the wavelet domain using the metadata. We utilize different distance metrics for the two wavelet types:

Haar Metric: For the Haar wavelet coefficients we employ a simple L^1 metric on each component independently. For example, the X component is computed as follows: $\text{Score}_x(P,G) = \sum_{i,j} |\text{P}_x[i,j] - \text{G}_x[i,j]|$ where P and G are the probe and gallery images, respectively. The total distance score is the sum of the scores of all components: $d(P,G) = \text{Score}_x(P,G) + \text{Score}_y(P,G) + \text{Score}_z(P,G)$

Pyramid Metric: For the pyramid coefficients we employ the complex wavelet structural similarity index algorithm [25]. For the X component, a window of size 3 traverses

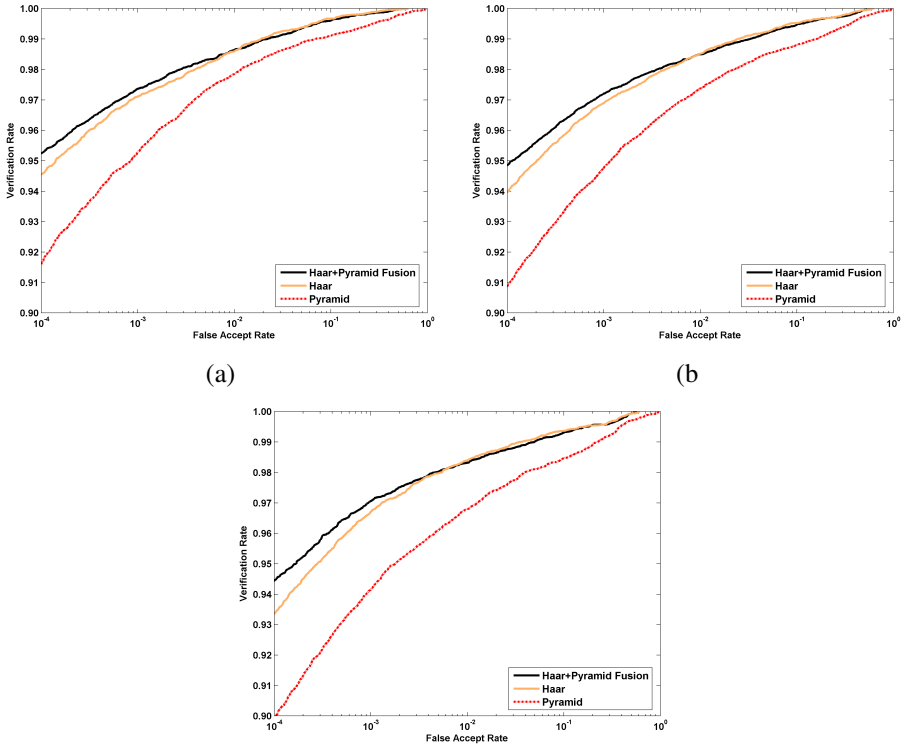


Figure 4: System performance using various wavelet transforms (Haar and Pyramid) as well as their fusion on the FRGC v2 database. (a) ROC I, (b) ROC II and (c) ROC III.

the image one step at a time. In each step, we extract all wavelet coefficients, resulting in two sets of coefficients $p_{w,i} = \{p_{w,i} | i = 1,...,N\}$ and $g_{w,i} = \{g_{w,i} | i = 1,...,N\}$, drawn from the probe image and the gallery image, respectively. The distance metric is computed as follows:

$$\tilde{S}(p_w, g_w) = 1 - \left(\frac{2\sum_{i=1}^{N} |p_{w,i}||g_{w,i}| + K}{\sum_{i=1}^{N} |p_{w,i}|^2 + \sum_{i=1}^{N} |g_{w,i}|^2 + K} \right) \cdot \left(\frac{2|\sum_{i=1}^{N} p_{w,i} g^*_{w,i}| + K}{2\sum_{i=1}^{N} |p_{w,i} g^*_{w,i}| + K} \right)^r$$

where w is the current step of the window, N is the number of coefficients in the window and r is an experimentally determined exponent. K is a small positive value which is used to make the result numerically stable. The first component measures the equivalence of the two coefficient sets while the second reflects the consistency of phase changes. If $p_{w,i} = g_{w,i}$ for all i's, the distance is 0. The weighted sum of the local scores from all windows provides the distance score in the X component: $Score_x(P,G) = \sum_{w=1}^{N} (b_w \cdot \tilde{S}(p_w, g_w))$ where b_w is a predefined weight depending on which subband the local window lies on. The weight assigned to a particular subband is determined by its descriptive power and performance. Similarly, we compute the scores for the Y and Z components and, to compute the total distance, we sum the scores as with Haar.

	ROC I	ROC II	ROC III
Fusion	97.3%	97.2%	97.0%
Haar	97.1%	96.8%	96.7%
Pyramid	95.2%	94.7%	94.1%

Table 1: Verification rates of our system at 0.001 FAR using various transforms on the FRGC v2 database.

4 Performance Evaluation

We perform our tests on the the FRGC v2 database that contains 4007 3D scans of 466 persons. The data were acquired using a Minolta 910 laser scanner that produces range images with a resolution of 640x480. The scans contain various facial expressions (e.g., happiness, surprise). Subjects are 57% male and 43% female, with the following age distribution: 65% 18-22 years old, 18% 23-27 and 17% 28 years or older.

The results are summarized using receiver operating characteristic (ROC) curves. For the FRGC v2 database, in order to produce comparable results, we utilize the three masks provided by FRGC v2. These masks, referred to as ROC I, ROC II, and ROC III, are defined over the square dissimilarity matrix (4007x4007), and are of increasing difficulty (the difficulty reflects the time elapsed between the probe and gallery acquisition sessions [19]).

Transforms: The first experiment is performed on the FRGC v2 database and its purpose is to evaluate the two wavelet transforms that we employ, as well as to provide a reference score for our system using publicly available databases and methods. In this experiment, our system using a fusion of the two transforms yielded a verification rate of 97.3% (for ROC I at 0.001 FAR), while separately for the Haar transform a rate of 97.1% and for the pyramid transform a rate of 95.2% were achieved (Figure 4). For the fusion, we experimentally found that the weighted sum is the most efficient yet simplest rule. Even though the pyramid transform is computationally more expensive it is outperformed by the simpler Haar wavelet transform (this can be attributed to the fact that in the current implementation, the pyramid transform utilizes only the geometry images and not the normal map images). The fusion of the two transforms offers more descriptive power, yielding higher scores, especially in the more difficult experiments of ROC II and ROC III (Table 1). To the best of our knowledge, this is the highest performance reported on the FRGC v2 database for the 3D modality. The computational time needed to generate a geometry image from the data is 15 seconds on average with unoptimized code. The time needed to compute the L^1 distance between a probe and a gallery metadata allows for more than 4000 datasets to be processed per second. The computational time was measured on an Opteron 2.4 GHz CPU.

Facial Expressions: The second experiment is focused on the effect of facial expressions on performance. The FRGC v2 database provides a categorization of the expressions that each individual assumes, allowing an easy division on two subsets: one containing only datasets where facial expressions are present, the other containing only datasets with neutral expressions.

	ROC I	ROC II	ROC III
Full Database	97.3%	97.2%	97.0%
Non-neutral Expressions	95.6%	95.6%	95.6%
Neutral Expressions	99.0%	98.7%	98.5%

Table 2: Performance of our system at 0.001 FAR on the full FRGC v2 database, and on subsets containing only non-neutral facial expressions and only neutral expressions.

The performance on the two subsets is measured and compared to the performance on the full database utilizing a verification scenario (Table 2). The verification rate is not decreased by a significant amount when expressions are present. The average decrease of 1.56% of the verification rate at 0.001 FAR between the full database and the facial expressions-only subset, is very modest (compared to other systems) given the fact that this subset contains the most challenging datasets from the whole database.

5 Conclusion

We presented algorithmic solutions to most of the challenges faced by 3D facial recognition systems. By utilizing a deformable model we map the 3D geometry information onto a 2D regular grid, thus combining the descriptiveness of 3D data with the computational efficiency of 2D data. A multistage fully automatic alignment algorithm and the advanced wavelet analysis resulted in robust state-of-the-art performance on the publicly available FRGC v2 database.

References

[1] Face Recognition Vendor Test 2006. `http://www.frvt.org/FRVT2006/`.

[2] R. Kimmel A. M. Bronstein, M. M. Bronstein. Three-dimensional face recognition. *Intl. Journal of Computer Vision*, 64(1):5–30, August 2005.

[3] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[4] V. Blanz and T. Vetter. Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1063 – 1074, September 2003.

[5] K. W. Bowyer, K. I. Chang, and P.J. Flynn. A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. *Computer Vision and Image Understanding*, 101(1):1–15, January 2006.

[6] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Adaptive rigid multi-region selection for handling expression variation in 3D face recognition. In *Proc. IEEE Workshop on Face Recognition Grand Challenge Experiments*, page 157, Washington, DC, USA, 2005.

[7] K. I. Chang, K. W. Bowyer, and P.J. Flynn. Effects on facial expression in 3D face recognition. In *Proc. SPIE Biometric Technology for Human Identification II*, volume 5779, pages 132–143, Orlando, FL, March 28-29 2005.

[8] D. Chetverikov, D. Stepanov, and P. Krsek. Robust Euclidean alignment of 3D point sets: the trimmed iterative closest point algorithm. *Image Vision Comput.*, 23(3):299–309, 2005.

[9] L. Farkas. *Anthropometry of the Head and Face*. Raven Press, 1994.

[10] X. Gu, S. Gortler, and H. Hoppe. Geometry images. In *Proc. SIGGRAPH*, pages 355–361, San Antonio, TX, July 2002.

[11] M. Husken, M. Brauckmann, S. Gehlen, and C. von der Malsburg. Strategies and benefits of fusion of 2D and 3D face recognition. In *Proc. IEEE Workshop on Face Recognition Grand Challenge Experiments*, San Diego, CA, June 20-25 2005.

[12] A. Johnson. *Spin-Images: A Representation for 3-D Surface Matching*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August 1997.

[13] I.A. Kakadiaris, M. Papadakis, L. Shen, D. Kouri, and D.K. Hoffman. m-HDAF multiresolution deformable models. In *Proc. 14th International Conference on Digital Signal Processing*, pages 505–508, Santorini, Greece, July 1-3 2002.

[14] I.A. Kakadiaris, L. Shen, M. Papadakis, D. Kouri, and D.K. Hoffman. g-HDAF multiresolution deformable models for shape modeling and reconstruction. In *Proc. British Machine Vision Conference*, pages 303–312, Cardiff, United Kingdom, Sept. 2-5 2002.

[15] C. Loop. Smooth subdivision surfaces based on triangles. Master's thesis, Department of Mathematics, University of Utah, 1987.

[16] C. Mandal, H. Qin, and B. Vemuri. Dynamic smooth subdivision surfaces for data visualization. *IEEE Visualization*, pages 371–377, October 1997.

[17] C. Mandal, H. Qin, and B.C. Vemuri. A novel FEM-based dynamic framework for subdivision surfaces. *Computer-Aided Design*, 32(8-9):479–497, 2000.

[18] T. Maurer, D. Guigonis, I. Maslov, B. Pesenti, A. Tsaregorodtsev, D. West, and G. Medioni. Performance of Geometrix ActiveID^{TM}3D face recognition engine on the FRGC data. In *Proc. IEEE Workshop on Face Recognition Grand Challenge Experiments*, San Diego, CA, June 20-25 2005.

[19] 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. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 947–954, San Diego, CA, 2005.

[20] P.J. Phillips, A. Martin, C.L. Wilson, and M. Przybocki. An introduction to evaluating biometric systems. *IEEE Computer*, 33(2):56–63, February 2000.

[21] J. Portilla and E.P. Simoncelli. A parametric texture model based on joint statistic of complex wavelet coefficients. *Intl. J Computer Vision*, 40:49–71, 2000.

[22] P. Siarry, G. Berthiau, F. Durbin, and J. Haussy. Enhanced simulated annealing for globally minimizing functions of many-continuous variables. *ACM Transactions on Mathematical Software*, 23(2):209–228, 1997.

[23] E. Stollnitz, T. DeRose, and D. Salesin. *Wavelets for Computer Graphics: Theory and Applications*. Morgan Kaufmann Publishers, Inc, 1996.

[24] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *Proc. SIGGRAPH*, pages 311–318, Orlando, FL, 1994.

[25] Z. Wang and E.P. Simoncelli. Translation insensitive image similarity in complex wavelet domain. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, volume II, pages 573–576, Philadelphia, PA, March 2005.