Inspired by Bertillon – Recognition Based on Anatomical Features from 3D Face Scans

Štěpán Mráček*, Christoph Busch†, Radim Dvořák* and Martin Drahanský*
*Brno University of Technology
Faculty of Information Technology
Brno, Czech Republic
{imracek,idvorak,drahan}@fit.vutbr.cz
†Hochschule Darmstadt, CASED and
Fraunhofer Institute for Computer Graphics (IGD)
Darmstadt, Germany
christoph.busch@igd.fraunhofer.de

***Abstract*—We present an automatic 3D face recognition algorithm that is inspired by Alphonse Bertillon's anthropometry. Our recognition pipeline consists of several steps. First, the facial landmarks such as the tip of the nose or the inner eye corners are detected. Subsequently the head rotation is compensated during the orientation normalization process. The facial features are extracted by performing 61 different measures. We also present a feature evaluation function that rates individual components of the feature vector. Finally, our results are compared with two other 3D face recognition methods. We show that the multi-algorithmic system consisting of the anatomical-based recognition together with the eigenfaces method and the recognition using histogram-based features reaches significantly better results than any of the employed methods individually.**

I. INTRODUCTION

Face recognition is one of the most used biometric techniques. Humans are able to localize a face in a very large and complicated scene. Also the detection of anatomical features, such as the position of the nose, the eyes, and the mouth within the face, doesn't pose us difficulties. Furthermore, we can recognize faces from various angles, even if we are confronted with a face expression or if a part of a face is covered. Many activities that we are doing completely automatically with no effort, become quite difficult if we try to describe this process mathematically.

Nevertheless, a lot of research has been done in recent years in the area of the biometric face recognition, especially addressing three-dimensional recognition methods. With an additional three-dimensional measurement, a face recognition system becomes more robust with respect to faked biometric characteristics (e.g., a 2D-facial print out). The facial biometric reference has become a mandatory part of the biometric passports in all member states of the ICAO (International Civil Aviation Organization).

The face recognition method presented in this paper is based on the facial anatomical features that describe the morphology of the facial geometric landscape. Contrary to the other well-known methods like the eigenface approach [TP91] or histogram-based recognition [ZSBF08], the specific location of facial landmarks, their mutual positions and another topological-related measures are used to form a feature vector that is consequently used to the direct comparison between a gallery 3D reference record and a probe record.

In contrast to some other purely 3D face recognition methods, for example those introduced by Cadoni et al. [CBG09] or Lu et al. [LCJ04], we extract only significant data into feature vectors that are subsequently compared. There is no need to store the entire model of the face in the biometric database in order to compare two samples. This approach has a positive impact on the data storage size as well as on comparison through-put (speed), especially when an identification process (one-to-many comparison) is in operation.

II. RECOGNITION PIPELINE

Our approach covers all parts of recognition, including training on significant representatives of the target population, preprocessing, calculating the comparison score, and the final decision. The recognition pipeline is modular, therefore each module deals with the specific problem and can be replaced with another one which does the same task.

The overview of the proposed recognition pipeline is in Fig. 1. The landmark detection module is responsible for locating distinguished points in the facial scan. The input of this module is the facial surface representation usually given as point-cloud, processed mesh or range image. The output is the set of located landmarks, e.g., the nose tip, nasal bridge, and inner eye corners.

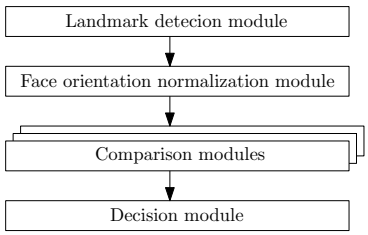

Fig. 1. Recognition pipeline

The control is subsequently passed to the normalization module. This module aligns the face to the predefined position within the space. Utilization of the planar symmetry of the face as well as positions of located landmarks are used in this step.

In the comparison modules, the features are extracted and the resulting feature vector is compared with a template stored in the database. Multiple comparison modules are implemented such that besides the recognition based on anatomical features, the eigenface method utilizing PCA and face recognition algorithm using histogram-based features are applied.

Finally, the comparison scores are normalized and fused together. We are using the weighted sum of gained comparison scores. The weight coefficients are derived from the equal error rates provided by the utilized recognition methods.

A. Face Surface Processing

Most of the employed algorithms in the recognition pipeline operate on a range image representation of the face. Since our testing dataset is using the point-cloud representation of the three-dimensional face scans, conversion to the range image has to be done first. We are using parallel projection where each point in the three-dimensional space is linearly projected to the xy-plane that is divided into a regular grid. One cell in the grid corresponds to one pixel in the resulting range image and pixel values indicate the distance of the surface from the plane. If more points are projected into the same resulting pixel, then the point with the maximal z-coordinate is used as the pixel value to solve the potential problem of concealed surfaces.

The acquisition of the three-dimensional data in inappropriate lighting conditions may cause presence of a noise or spikes (unusually protruding points) on the face surface. In order to reduce the impact of this deficiency on the recognition performance, a spike removal operator followed by Laplacian smoothing is performed.

During spike removing, a simple method is used to identify outliers. The value of each pixel is compared with the average value of the surrounding pixels. If the difference between the pixel and its surroundings exceeds a specific threshold, then the pixel is considered as a spike and its depth value is replaced by the average of its neighbors.

Subsequently, the iterative Laplacian smoothing [Fie88] is performed. The depth value of each pixel x_i is modified in every iteration according to its N neighbors:

$$x_i = x_i + \alpha \frac{\sum_{x_j \in neighbours(x_i)} x_j}{N} \quad (1)$$

The coefficient α and the number of iterations affects the degree of the smoothing.

In the further analysis of the surface, distinguishing landmarks are detected. The landmark detection in our recognition pipeline relies on the curvature analysis of the face surface. The principal idea of this curvature estimation method is illustrated in Fig. 2. The curvature at some given surface point C is estimated from the difference d between the elevated point C and the connecting line between adjacent points A and B. This method provides larger values for those points located on a bended part of surface and smaller values for points on a smooth part of the surface. The curvature c at the given point C is then estimated as $c(C) \approx d$.

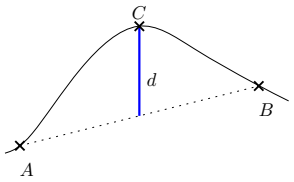

Fig. 2. Curvature at the point C is estimated as distance d between line AB and point C

The curvature at every point is calculated in four directions (horizontal, vertical, and two diagonal). From those four curvatures the two principal curvatures are selected as the pair with maximized difference, e.g., if the difference between the horizontal and the vertical curvatures is greater than between the diagonal pair, then the horizontal and the vertical curvature are selected as principal curvatures k_1 and k_2.

For further classification of the surface the following curvature descriptions are used:

- mean curvature: $H = \frac{1}{2}(k_1 + k_2)$
- Gaussian curvature: $K = k_1 \cdot k_2$

Various types of a surface morphology, such are pits, peaks, saddles, and valleys can be detected by comparison of the signs of the Gaussian and mean curvature, see Table I in Segundo et al. [SQBS07].

TABLE I
SURFACE CLASSIFICATION

$K \backslash H$	< 0	$= 0$	> 0
< 0	saddle ridge	minimal	sadle valley
$= 0$	ridge	flat	valley
> 0	peak	(none)	pit

B. Landmark Detection

The landmark detection algorithm employed in our method is based on the classification of the face surface morphology. We are using a slightly modified approach of the method presented by Segundo et al [SQBS07]. During this process, the nose tip, the inner eyes corners, the nasal bridge, and the lower as well as outer nose corners are detected (see Fig. 3).

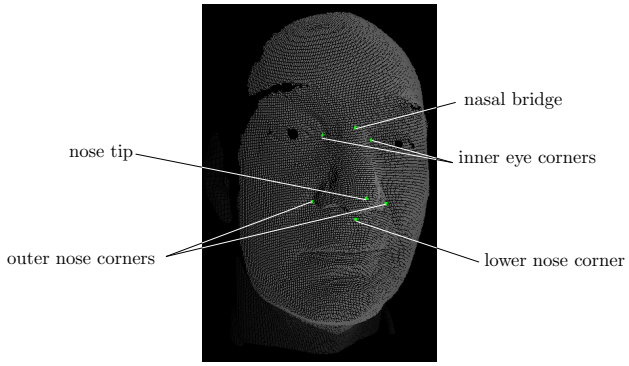

Fig. 3. Result of the landmark detection

Initially, the area of interest of the face has to be separated such that surface data representing the clothes, neck, and hair

is masked out. This is achieved in two steps. First, the K-means clustering [JD88] separates the data into two clusters. This will isolate parts like ears, shoulders, and other parts that are not protruding and therefore they are not considered as a surface which may contain landmarks of interest. However, some protruding parts, like shirt collar or haircut, may be still present after the K-means clustering. Surface areas stemming from these regions are compared and the largest one is finally considered as the region that contains landmarks of interest.

After the face region is selected, the nose tip is located as a point with the highest peak density inside the region. An example of the range image and corresponding peak density map is given in Fig. 4. The inner eye corners are located in an analogical manner as those points with the highest pit density. An additional heuristic is used during the eye corners detection. Since the location of the nose tip is known and with the assumption that the face is not rotated upside down, a highest pit density point can be searched only within the reduced area above the nose tip.

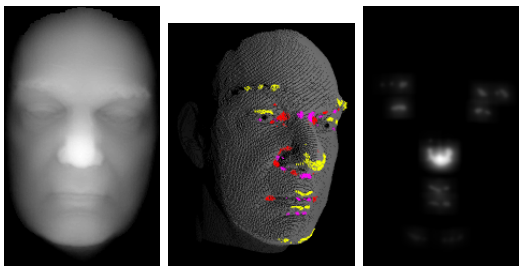

Fig. 4. Left: input range image, Center: highlighted peaks (yellow) and pits (red) on the face surface, Right: corresponding peak density map

The point on the connecting line between the inner eye corners with the highest curvature is marked as the nasal bridge. However, the localization of the nasal bridge and nose tip is not completely accurate. Therefore, additional processing has to be done. The nasal bridge is usually situated a bit higher on the face and the nose tip location may be shifted due to nose asymmetry or false peak detection. The area between the temporary located nasal bridge and the nose tip is divided into horizontal curves. In each curve, that consists of pixels from one range image row, a distinct point located that represents the highest horizontal curvature in that row and added to a set P. A line that fits all points in set P in a least square mean approach is calculated from the set P. This line is then projected to the face surface and a reliable profile curve is defined in this way.

The final position of the nose tip as well as the nasal bridge and the lower nose corner are located on the local extremes of the second derivation from the profile curve.

C. Orientation Normalization

For the purpose of the orientation normalization, the nose tip, nasal bridge, and lower nose corner are used. The whole model is moved first, such that the nose tip is situated in the origin of coordinates. After that, the rotation around all three main axes is performed. Although for example both inner eyes corners together with the nose tip can be used to align the face to some predefined position, we observed that the more robust orientation normalization is achieved when additional information is gained from the planar symmetry of the face. The iterative orientation normalization algorithm is as follows:

1) The vector between the nose tip and the nasal bridge is rotated in such a manner that x-coordinates of its ending points are equal to zero and the angle between the xz-plane and the nasal bridge vector is 1.1 radians. This situation is illustrated in Fig. 5. We empirically observed that the angle 1.1 radians is the most suitable value.
2) The rotation along the y axis is estimated with the aid of the planar symmetry of the face. From both left and right part of the face the centroid point is calculated, which represents the cumulative z coordinates of the points on left or right side. For a correctly-oriented face both centroid values are equal. The rotation angle is calculated from the angle between the centroid connecting line and the x axis.
3) Points 1 and 2 are repeated until both rotation angles are sufficiently small. It has been observed that 5 iterations are usually enough for a robust face orientation.

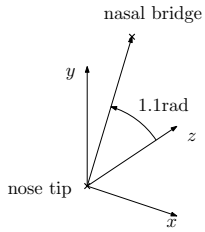

Fig. 5. Nasal vector normalization – the nose tip and the nasal bridge is rotated to a predefined position

III. FEATURE EXTRACTION

As a 3D-scan of a human face is a metrically correct presentation we can directly extract anatomical facial features from the properly oriented face scan with located landmarks. We are using 61 different measures that can be divided into three main categories. In the first category, basic Bertillon-like measures are present, which describe facial surface morphology. This category includes measures described in Table II.

Along with these basic measures, in a second category another 48 measures connected with facial curves on the face were extracted. These curves are shown in Fig. 6. The similarity between the selected subject specific curve and the corresponding curve extracted from an average face forms a measure. We applied four different coefficients to describe the similarity of a subject specific curve X and an average face curve Y:

Pearson product-moment correlation coefficient:

$$r = \frac{\sum_{x \in X}(x - \mu_X)\sum_{y \in Y}(y - \mu_Y)}{N\sigma_X\sigma_Y} \qquad (2)$$

TABLE II
BERTILLON-LIKE MEASURES

Measure name	Description
Nose width	distance between outer nose corners
Nose depth	distance between line connecting the outer nose corners and the nose tip
Nose height	distance between the lower nose corner and the nasal bridge
Nose sharpness	Angle between the nasal bridge, nose tip and the lower nose corner
$\frac{\text{Nosedepth}}{\text{Nosewidth}}$	Ratio between the nose depth and the nose width
Nose volume	Count of voxels that are located under nose area
Bridge of the nose depth	z-axis difference between the eye corners and the nasal bridge

where μ_X is the arithmetic mean of the set X, σ_X is its standard deviation and N is number of elements in each set.

Cross correlation:

$$r = \frac{\langle X, Y \rangle}{\|X\| \cdot \|Y\|} \tag{3}$$

where $\langle X, Y \rangle$ stands for the dot product and $\|X\|$ is norm of X.

Hausdorff distance:

$$d_H(X, Y) = \max\{\sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y)\} \tag{4}$$

where $d(x, y)$ can be arbitrary metric function in the given metric space. We are using Euclidean distance.

Sum of differences:

$$d = \sum_{i=1}^{N} (X_i - Y_i) \tag{5}$$

where N is size of sets X and Y.

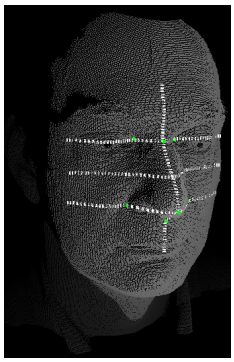

Fig. 6. Facial curves

Each curve was compared with the corresponding curve extracted from the average face using all similarity coefficients mentioned above. Moreover, the similarity of the first and the second derivation of all facial curves to their corresponding average counterparts were calculated.

Another 6 local measures were extracted from the curvature at the nose tip, nasal bridge, and the point halfway between the nose tip and the nasal bridge. Both vertical and horizontal curvatures were measured in all three points.

The suitability of anatomical features for the face recognition is illustrated in Fig. 7. Only three measures were selected to form the anatomical feature vector. As multiple scans were available for all data subjects the distribution of these scans according to the three primary anatomical features is of interest. Fig. 7 illustrates that scans from the same subject, which are visualized by symbol and color, are clearly creating clusters in the low dimensional feature space.

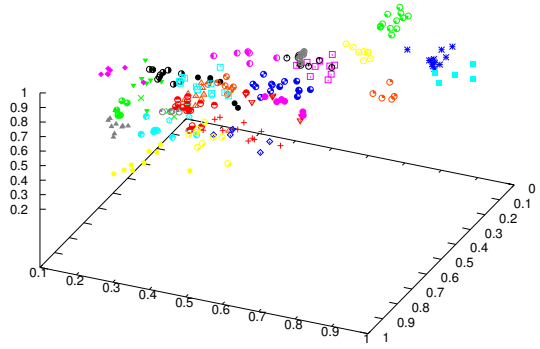

Fig. 7. Distribution of selected anatomical features in the face feature space

IV. PERFORMANCE EVALUATION

A. Reliability of Anatomical Features

Our approach has been tested on the Face Recognition Grand Challenge (FRGC) version 2 database.

A biometric performance testing was conducted on the "spring2004" part of the FRGC v2 database [PFS+05] in order to confirm the reliability of the selected anatomical features. We select only those scans that belong to subjects with more than four scans. This set contains 1837 scans from 226 subjects. Out of these scans, we choose 200 scans as the training set. We used both Euclidean (see Eq. 6) and city block distance (see Eq. 7) as comparison score. The overall results are given in Table III. Better performance was achieved utilizing the city block distance.

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{6}$$

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i| \tag{7}$$

However, not all anatomical features have the same discriminating capability. To select only those measures that have a positive contribution to the recognition performance, we are using a selection procedure based on a discriminative potential function. This function provides greater values for those measures where the intra-class variability is low and

TABLE III
INITIAL PERFORMANCE EVALUATION FOR THE RECOGNITION BASED ON THE ANATOMICAL FEATURES

Utilized distance function	Achieved EER
Euclidian	15%
City block	12%

the inter-class variability is high. After all measured values are normalized to the range between 0 and 1, the intra-class variability is expressed using these three values:

- mean deviation μ – mean value range for each subject.
- standard deviation s – standard deviation of value ranges for each subject
- maximum deviation m – maximal value range for each subject

In order to achieve good inter-class variability, measured values have to be uniformly distributed among the space. Therefore the inter-class variability is expressed as the correlation c of measured values to the uniform probability density function.

The discriminative potential function of a measure x is defined as:

$$discriminative\ potential(x) = (1 - (\mu + s\ + m)) + c \quad (8)$$

In a subsequent performance test, we consecutively selected only those measures with a discriminative potential $p \geq t$, where $t \in \{0.0, 0.1, \ldots, 1.0\}$. Only the selected discriminative measures constitute the feature vector in this case. The test results are given in Fig. 8 indicating that the best performance (6.4% EER) was achieved by using the city block distance on those feature vector components with discrimination potential $p \geq 0.3$. Only 20 selected elements out of 61 possible features met this criteria.

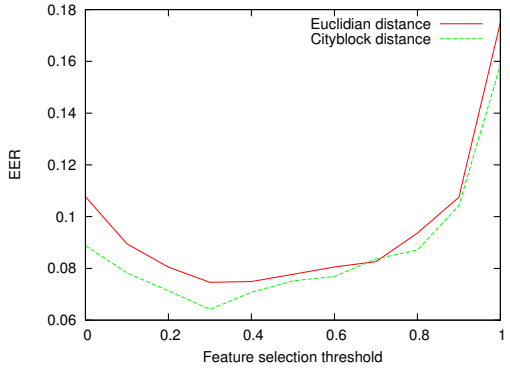

Fig. 8. Evaluation of recognition based on the anatomical features

B. Performance Improvement with Score-level fusion

Additionally to the recognition method based on the anatomical features, we implemented the eigenfaces method (PCA) and further a recognition method using histogram-based features.

As an input to the PCA we are using the shape index images of the face surface as introduced by Lu et al. [LCJ04]. The shape index value S of every pixel x is gained from its principal curvatures k_1 and k_2 using the following formula:

$$S(x) = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \frac{k_1 + k_2}{k1 - k2} \quad (9)$$

To avoid the impact of face parts that suffer from high intra-class variability, such as in the mouth region, only a selected region of interest (ROI) is used. An example of a PCA input image with highlighted region of interest is given in Fig. 9. The advantages of the PCA applied on processed range images has been investigated by Heseltine [Hes05]. A similar ROI selection was proposed by Colombo et al. [CCS06].

Fig. 9. Example of an input to the PCA with highlighted region of interest

For the histogram-based recognition method as presented by Zhou et al. [ZSBF08], the face range image is divided into N horizontal stripes. Each stripe is processed and points are assigned into K bins based on the z-coordinate values of the points. This process provides a feature vector of size $N \cdot K$ that can be directly used for calculating similarity score between the reference template and the probe scan captured from some data subject. We achieved the best results with the division to 10 rows and 6 bins within each row employing weighted city block distance function. The weight coefficients were derived from the discriminative potential function that was applied on all components of the feature vector.

TABLE IV
PERFORMANCE OVERVIEW OF APPLIED RECOGNITION METHODS

Method	EER
Anatomical features	6.4%
Eigenfaces (PCA)	8.8%
Histogram-based recognition	8.6%

The performance overview of applied recognition methods is given in Table IV. In order to fuse the methods on a score level, score normalization is required. We normalize the score to a range between 0 and 1.

We compute the normalized score s_{norm} based on the threshold s_{EER}, for which the EER is reached, and the comparison score s provided by one of the recognition methods. If s is lower or equal than s_{EER} then the normalized score $norm(s)$ is

$$norm(s) = \frac{s}{2 \cdot s_{EER}} \tag{10}$$

and when s is greater than s_{EER}, the normalized score s' is

$$norm(s) = \frac{1}{2} + \frac{s \ - s_{EER}}{2 \cdot (score_{max} - s_{EER})} \tag{11}$$

The weighted score fusion s for scores $s_1, s_2, \ldots, s_n$ is defined as:

$$s \ = w_1 \cdot norm(s_1) + w_2 \cdot norm(s_2) + \ldots + w_n \cdot norm(s_n) \tag{12}$$

where $w_1, w_2, \ldots, w_n$ are weights derived from equal error rates of each contributing method in the following way. The weight should reflect the reliability of the method, therefore resulting in a stronger weight if the EER_i of the corresponding method i is lower.

The fusion of the PCA and histogram-based recognition method together with the anatomical-based recognition provides an EER = 3.9%. Thus by this multi-algorithmic fusion approach a recognition pipeline is established that outperforms every contributing method. The DET curve of the fused system is displayed along with the DET curves of the contributing methods in Fig. 10.

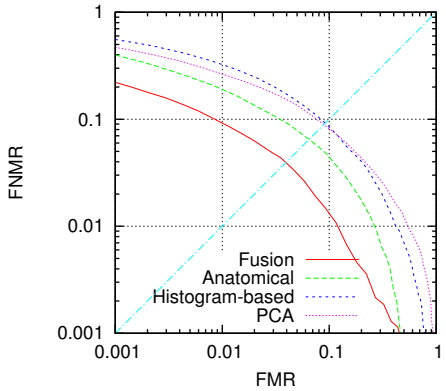

Fig. 10. DET curves of fused system and its components

V. CONCLUSION AND FURTHER WORK

A three-dimensional face recognition pipeline that includes landmark detection, face normalization feature extraction, model comparison and score fusion has been proposed in this paper. We have shown, that our recognition method based on the anatomical features of the face outperformed established approaches such as the histogram-based recognition and the PCA based recognition. In order to select only the relevant components of a feature vector, we employed a discriminative potential function. This function evaluates the inter-class and intra-class variability of individual feature vector components and thus provides a reasonable selection criteria.

The landmark detection is conducted merely on the 3D-face surface and remains a challenge. The proposed method operates very reliable thus that landmark detection, which is a prerequisite for the subsequent model normalization, failed in only 4 cases out of 1837 processed scans (0.2%). This failure rate was computed automatically during the batch landmark detection process. However in some cases the automatically place landmark will be dislocated from a position that would be determined by a human expert. Nevertheless the impact on the biometric performance is minimal.

Due to the robust orientation normalization process the system becomes independent on the head position and rotation. The observed recognition performance is still affected by facial expressions. In order to minimize the impact of non-neutral expression the recognition pipeline is focusing on the most rigid parts of the face, which is analyzed in the feature extraction step. However, by this region of interest approach may cause that a lot of possibly significant information is lost.

The implemented face recognition pipeline illustrates that just as Bertillon measured the anatomical features the same information can be extracted in a fully automated manner. Future work will focus on further improvements of the biometric performance by fusing 3D-information with features from the 2D-textured data.

VI. ACKNOWLEDGMENTS

This work was kindly supported by the Norwegian Information Security laboratory (NISlab), the Center for Advanced Security Research Darmstadt (CASED) and Fraunhofer IGD, which provided access to the FRGC v2 database.

REFERENCES

[CBG09] Marinella Cadoni, Manuele Bicego, and Enrico Grosso. 3d face recognition using joint differential invariants. In *ICB '09: Proceedings of the Third International Conference on Advances in Biometrics*, pages 279–288, Berlin, Heidelberg, 2009. Springer-Verlag.

[CCS06] A. Colombo, C. Cusano, and R. Schettini. 3d face detection using curvature analysis. *Pattern Recogn.*, 39(3):444–455, 2006.

[Fie88] D. A. Field. Laplacian smoothing and delaunay triangulations. *Communications in Applied Numerical Methods*, 4:709–712, 1988.

[Hes05] T. D. Heseltine. *Face Recognition: Two-Dimensional and Three-Dimensional Techniques*. PhD thesis, The University of York, 2005.

[JD88] Anil K. Jain and Richard C. Dubes. *Algorithms for clustering data*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1988.

[LCJ04] X. Lu, D. Colbry, and A. K. Jain. Three-dimensional model based face recognition. In *ICPR 04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR 04) Volume 1*, pages 362–366, Washington, DC, USA, 2004. IEEE Computer Society.

[PFS+05] P. Jonathon Phillips, Patrick J. Flynn, Todd Scruggs, Kevin W. Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, and William Worek. Overview of the face recognition grand challenge. pages 947–954, 2005.

[SQBS07] M. P. Segundo, C. Queirolo, O. R. P. Bellon, and L. Silva. Automatic 3d facial segmentation and landmark detection. In *ICIAP '07: Proceedings of the 14th International Conference on Image Analysis and Processing*, pages 431–436, Washington, DC, USA, 2007. IEEE Computer Society.

[TP91] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 586–591, 1991.

[ZSBF08] X. Zhou, H. Seibert, C. Busch, and W. Funk. A 3d face recognition algorithm using histogram-based features. *Eurographics Workshop on 3D Object Retrieval*, pages 65–71, 2008.