3D FACE RECOGNITION USING TOPOGRAPHIC HIGH-ORDER DERIVATIVES

Ali Cheraghian[1], *Farshid Hajati*[1,2], *Ajmal S. Mian*[3], *Yongsheng Gao*[2]

[1]Electrical Engineering Department, Amirkabir University of Technology, Iran
[2]School of Engineering, Griffith University, QLD 4111, Australia
[3]Computer Science and Software Engineering, The University of Western Australia, WA 6009, Australia
cheraghian@aut.ac.ir, f.hajati@griffith.edu.au, ajmal@csse.uwa.edu.au, yongsheng.gao@griffith.edu.au

ABSTRACT

This paper presents a novel feature, Topographic High-order Derivatives (THD) for 3D face recognition. THD is based on the high-order micro-pattern information extracted from face topography maps. Face topography maps are partitioned into polar sectors, and THDs are computed using directional high-order derivatives within the sectors. Local features are extracted by encoding directional high-order derivatives within polar neighborhoods. To evaluate the proposed method, we use Bosphorus and FRGC 3D face databases which include pose and expression changes. The performance of the proposed method is higher compared to the state-of-the-art benchmark approaches in 3D face recognition.

Index Terms— Topography, high-order derivatives, 3D face, face recognition

1. INTRODUCTION

Local features have been used in 2D face recognition algorithms due to their robustness to variations such as pose and illumination changes. Bruneli and Poggio [1] proposed one of the first methods in local approaches. They used a set of automatically selected facial features and computed the normalized cross-correlation between the region of the facial features in the probe and the same region in all gallery images. Nanni and Maio [2] proposed a Weighted Sub-Gabor method in which the Gabor wavelet was applied at each sub-pattern and the extracted Gabor features were projected to a low dimensional space. Ahonen et al. [3] applied Local Binary Patterns (LBP) to the face recognition task. In LBP, face images are considered as a composition of micro-patterns and a binary code is computed for pixels regarding their neighbors. Kanan et al. [4] proposed a local approach for face representation and recognition based on Adaptively Weighted Patch Pseudo Zernike Moment Array (AWPPZMA) using only one exemplar image per person. Zhang et al. [5] proposed a local pattern descriptor called Local Derivative Pattern (LDP). The LDP templates extract high-order local information from 2D images by encoding various distinctive partial relationships contained in the local regions.

Recent research shows that 3D images have more advantages in uncontrolled conditions compared to the traditional 2D images [6]. Cook et al. [7] partitioned 3D face images into a predefined number of overlapping spatial sub-regions and applied Log-Gabor transform to each sub-region at three different scales. Chang et al. [8] segmented overlapped sub-regions around the nose using ICP and combined the sub-regions at the score level. Alyuz et al. [9] used a fast and accurate region-based registration scheme to partition the 3D face images into 15 patches and construct seven meaningful regions. They applied Linear Discriminant Analysis (LDA) [10] to the individual patches and the results were fused at the decision level. Hajati et al. [11] proposed Patch Geodesic Distance (PGD) for expression and pose invariant 3D face recognition. In PGD, 3D face images are partitioned into distinct equal-sized square patches. Local geodesic paths within the patches and global geodesic paths for the patches are then combined to encode the shape adjusted textures into patterns.

We present a novel 3D face descriptor called Topographic High-order Derivatives (THD). The THD extracts directional high-order derivatives from the face topographic data. The proposed method describes the 3D face images by spatial micro-patterns extracted from both depth and texture data which are more invariant under expression and pose variations compared to 3D benchmark approaches.

2. TOPOGRAPHIC HIGH-ORDER DERIVATIVES

One of the most common shape representation methods is the topography map [12]. The distinctive characteristic of a topographic map is that the shape of the surface is shown by level curves (see Figure 1(a)). Level curves are lines that join points of equal height on a surface above or below a reference point (usually the highest point in the topography map) [13]. Given a surface $z = f(x, y)$, the cross-section between the

surface and a horizontal plane, the plane which is parallel to x-y plane, is called a level curve and the surface between two level curves is called a level strip. The level strip with the height index h, $f^h(x,y)$, is defined as

$$f^h(x,y) = \{(x,y,z)|z_{\max} - h.w_z < f(x,y) < z_{\max} - (h-1).w_z\} \quad (1)$$

where z_{max} is the maximum height in the surface, w_z is the step size of the level strip and h is the height index of the level strip defined as

$$h = [z/w_z] + 1 \ ; \ 0 \le z < z_{\max} \quad (2)$$

where w_z varies from 1 to z_{max}.

Since the level strips can be represented more accurately in the polar coordinates, we first transform the level strips into polar coordinates. Given a level strip $f^h(x,y)$, the transformed level strip in the polar coordinates, $F^h(r,\theta)$, is defined as

$$F^h(r,\theta) = f^h(x,y) \quad (3)$$

$$r = \sqrt{(x-x_p)^2 + (y-y_p)^2} \quad (4)$$

$$\theta = \begin{cases} 0, & ; x = x_p, \ y = y_p \\ \tan^{-1}(y - y_p/x - x_p), & ; x \ge x_p \\ \tan^{-1}(y - y_p/x - x_p) + \pi & ; x < x_p \end{cases} \quad (5)$$

Here, x_p and y_p are the coordinates of the pole (i.e. nose tip).

In order to extract local features, the level strips are partitioned into the predefined patches and the derivatives are computed within patches. We partition the level strips into non-overlapping polar patches with w_r and w_θ as the patch size in the radial and the angular directions, respectively. After partitioning, the point (r,θ,z) in the level strip $F^h(r,\theta)$ will be located in the (n,m)th patch where n and m indices are determined as

$$n = [r/w_r] + 1 \ ; \ 0 \le r < r_{\max} \quad (6)$$

$$m = [\theta/w_\theta] + 1 \ ; \ 0 \le \theta < 2\pi \quad (7)$$

where w_r and w_θ are the patch sizes in the radial and the angular directions, respectively. We define the patch sizes in the radial and the angular directions (w_r and w_θ) as

$$w_r = r_{\max}/N_r \quad (8)$$

$$w_\theta = 2\pi/3(2n-1) \quad (9)$$

where r_{max} and N_r are the maximum radius and the number of sections in the radial direction, respectively. n is the patch index number computed using Equation (6).

The total number of patches, N_p, in the partitioned topography map can be defined using Equation (10).

$$N_p = \sum_{k=1}^{[r_{\max}/w_r]+1} 3(2k-1) \quad (10)$$

Fig. 1. Example of face topography map and its partitioning into polar patches.

Figure 1(b) illustrates an example of partitioning a face topography map into polar patches. In the partitioned topography map, the (n,m)th patch of the level strip with the height index h, $F^h_{nm}(r^{nm},\theta^{nm})$, is represented as

$$F^h_{nm}(r^{nm},\theta^{nm}) = F^h(w_r(n-1) + r^{nm}, w_\theta(m-1) + \theta^{nm}) \quad (11)$$

where $F^h(r,\theta)$ is the level strip with the height index h. r^{nm} and θ^{nm} are the coordinates of the (n,m)th patch of the partitioned level strip.

Figure 2 shows an 8-neighborhood in the (n,m)th patch of the partitioned level strip. Let (r^{nm},θ^{nm}) be a point of the (n,m)th patch of the level strip with the height index h, $F^h_{nm}(r^{nm},{}^{nm})$, the eight neighbors of the point are defined as

$$F^h_{nm}(r^{nm}_1,\theta^{nm}_1) = F^h_{nm}(r^{nm} + \Delta r, \theta^{nm} + \Delta\theta) \quad (12)$$

$$F^h_{nm}(r^{nm}_2,\theta^{nm}_2) = F^h_{nm}(r^{nm} + \Delta r, \theta^{nm}) \quad (13)$$

$$F^h_{nm}(r^{nm}_3,\theta^{nm}_3) = F^h_{nm}(r^{nm} + \Delta r, \theta^{nm} - \Delta\theta) \quad (14)$$

$$F^h_{nm}(r^{nm}_4,\theta^{nm}_4) = F^h_{nm}(r^{nm}, \theta^{nm} - \Delta\theta) \quad (15)$$

$$F^h_{nm}(r^{nm}_5,\theta^{nm}_5) = F^h_{nm}(r^{nm} - \Delta r, \theta^{nm} - \Delta\theta) \quad (16)$$

$$F^h_{nm}(r^{nm}_6,\theta^{nm}_6) = F^h_{nm}(r^{nm} - \Delta r, \theta^{nm}) \quad (17)$$

$$F^h_{nm}(r^{nm}_7,\theta^{nm}_7) = F^h_{nm}(r^{nm} - \Delta r, \theta^{nm} + \Delta\theta) \quad (18)$$

$$F^h_{nm}(r^{nm}_8,\theta^{nm}_8) = F^h_{nm}(r^{nm}, \theta^{nm} + \Delta\theta) \quad (19)$$

where Δr and $\Delta\theta$ are the radial and angular resolutions of the topography map.

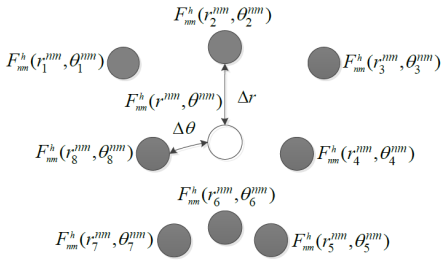


Fig. 2. The 8-Neighborhood of a point in polar coordinates.

After determining the 8-neighborhood of the points, the derivative along radial and angular directions with different

orders can be computed. The first-order polar derivatives along the radial and the angular directions are defined as

$$\partial F^h_{nm}(r^{nm},\theta^{nm})/\partial r = F^h_{nm}(r^{nm},\theta^{nm}) - F^h_{nm}(r^{nm}+\Delta r,\theta^{mn}) \quad (20)$$

$$\partial F^h_{nm}(r^{nm},\theta^{nm})/\partial \theta = F^h_{nm}(r^{nm},\theta^{nm}) - F^h_{nm}(r^{nm},\theta^{mn}+\Delta\theta) \quad (21)$$

Similarly, the second-order polar directional derivatives can be defined as

$$\partial^2 F^h_{nm}(r^{nm},\theta^{nm})/\partial r^2 = \partial F^h_{nm}(r^{nm},\theta^{nm})/\partial r - \partial F^h_{nm}(r^{nm}+\Delta r,\theta^{mn})/\partial r \quad (22)$$

$$\partial^2 F^h_{nm}(r^{nm},\theta^{nm})/\partial \theta^2 = \partial F^h_{nm}(r^{nm},\theta^{nm})/\partial \theta - \partial F^h_{nm}(r^{nm},\theta^{mn}+\Delta\theta)/\partial \theta \quad (23)$$

Generally, the l^{th}-order polar directional derivatives are defined using the $(l-1)^{th}$-order polar directional derivatives as

$$\partial^l F^h_{nm}(r^{nm},\theta^{nm})/\partial r^l = \partial^{(l-1)} F^h_{nm}(r^{nm},\theta^{nm})/\partial r^{(l-1)} - \partial^{(l-1)} F^h_{nm}(r^{nm}+\Delta r,\theta^{mn})/\partial r^{(l-1)} \quad (24)$$

$$\partial^l F^h_{nm}(r^{nm},\theta^{nm})/\partial \theta^l = \partial^{(l-1)} F^h_{nm}(r^{nm},\theta^{nm})/\partial \theta^{(l-1)} - \partial^{(l-1)} F^h_{nm}(r^{nm},\theta^{mn}+\Delta\theta)/\partial \theta^{(l-1)} \quad (25)$$

We encode the computed derivatives into binary bits using the unit step function, thereby forming a binary pattern. The l^{th}-order directional Topographic High-order Derivatives Patterns (THDP) in the radial and the angular directions in the (n,m)th patch of the level strip with the height index h, $F^h_{nm}(r^{nm},\theta^{nm})$, are defined as

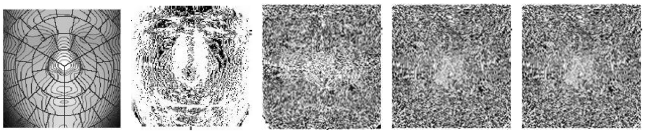
$$THDP^l_r(F^h_{nm}(r^{nm},\theta^{nm})) = \sum_{i=1}^{8} 2^i \cdot u(\partial^{(l)} F^h_{nm}(r^{nm},\theta^{nm})/\partial r^{(l)} \cdot \partial^{(l)} F^h_{nm}(r^{nm}_i,\theta^{nm}_i)/\partial r^{(l)}) \quad (26)$$

$$THDP^l_\theta(F^h_{nm}(r^{nm},\theta^{nm})) = \sum_{i=1}^{8} 2^i \cdot u(\partial^{(l)} F^h_{nm}(r^{nm},\theta^{nm})/\partial \theta^{(l)} \cdot \partial^{(l)} F^h_{nm}(r^{nm}_i,\theta^{nm}_i)/\partial \theta^{(l)}) \quad (27)$$

where $THDP^l_r(F^h_{nm}(r^{nm},\theta^{nm}))$ and $THDP^l_\theta(F^h_{nm}(r^{nm},\theta^{nm}))$ are the l^{th}-order directional Topographic High-order Derivatives Patterns (THDP) at point (r^{nm},θ^{nm}) of the (n,m)th patch along the radial and the angular directions, respectively. $u(\cdot)$ is the unit step function determining the direction of local pattern transitions. The result of directional THDP computations for an example 3D face image along the radial direction is illustrated in Figure 3.

Finally, we model the distribution of the directional THDPs using the spatial histogram [14] and concatenate them to make a feature vector as

$$HTHDP^l(F^h_{nm}(r^{nm},\theta^{nm})) = \{HTHDP^l_r(F^h_{nm}(r^{nm},\theta^{nm})), HTHDP^l_\theta(F^h_{nm}(r^{nm},\theta^{nm}))\} \quad (28)$$



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

Fig. 3. THDP along the radial direction. (a) original partitioned topography map, (b) first-order THDP, (c) second-order THDP, (d) third-order THDP, (e) forth-order THDP.

where $HTHDP^l_r(F^h_{nm}(r^{nm},\theta^{nm}))$ and $HTHDP^l_\theta(F^h_{nm}(r^{nm},\theta^{nm}))$ are the THDPs histogram vectors extracted in the radial and the angular directions, respectively. In this paper, we also apply the THD operator to the texture data. The texture map is partitioned similary to the height map. The l^{th}-order directional Topographic High-order Derivatives Patterns along the radial and the angular directions of the texture of the level strip with the height index h, $T^h_{nm}(r^{nm},\theta^{nm})$, are defined as

$$THDP^l_r(T^h_{nm}(r^{nm},\theta^{nm})) = \sum_{i=1}^{8} 2^i \cdot u(\partial^{(l)} T^h_{nm}(r^{nm},\theta^{nm})/\partial r^{(l)} \cdot \partial^{(l)} T^h_{nm}(r^{nm}_i,\theta^{nm}_i)/\partial r^{(l)}) \quad (29)$$

$$THDP^l_\theta(T^h_{nm}(r^{nm},\theta^{nm})) = \sum_{i=1}^{8} 2^i \cdot u(\partial^{(l)} T^h_{nm}(r^{nm},\theta^{nm})/\partial \theta^{(l)} \cdot \partial^{(n)} T^h_{nm}(r^{nm}_i,\theta^{nm}_i)/\partial \theta^{(l)}) \quad (30)$$

We also model and concatenate the distribution of texture THDPs using the spatial histogram as

$$HTHDP^l(T^h_{nm}(r^{nm},\theta^{nm})) = \{HTHDP^l_r(T^h_{nm}(r^{nm},\theta^{nm})), HTHDP^l_\theta(T^h_{nm}(r^{nm},\theta^{nm}))\} \quad (31)$$

We used the histogram intersection method [15] to measure the similarity between the histograms of the height and the texture data in the query and model images as

$$S_{HI}(H,S) = \sum_{j=1}^{32} \min(H_j, S_j) \quad (32)$$

where $H = (H_1, H_2, , H_{32})$ and $S = (S_1, S_2, , S_{32})$ are 32-bin histograms of the query and the model, respectively.

Finally, we combine the height and the texture data at the decision level with a balancing weight as

$$S(H,S) = (S^h_{HI} + w_b . S^t_{HI})^{1/2} \quad (33)$$

where S^h_{HI} and S^t_{HI} are the histogram intersections of the height and the texture data, respectively, and w_b is the balancing weight.

3. EXPRIMENTAL RESULTS

Here, we investigate the robustness of the proposed method in the face recognition task under pose and expression variations. We used the Bosphorus [16] and the FRGC v2 [17]

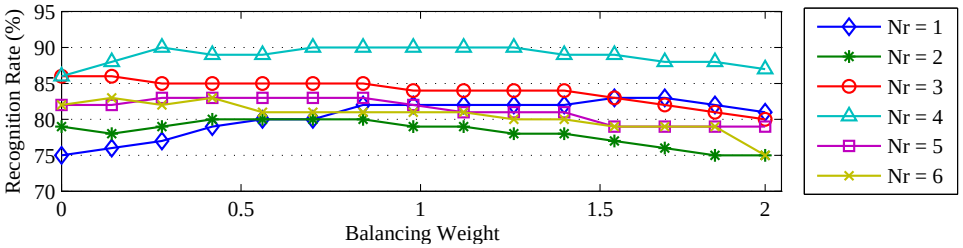


Fig. 4. Recognition rate versus the balancing weight (w_b) and different number of sections in the radial direction (N_r).

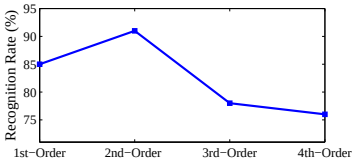


Fig. 5. The recognition rates for different THD orders.

databases. The face images are resampled to 160 by 160 with 20 level strips.

We select the number of sections in the radial direction N_r, the balancing weight w_b, and the best THD order to optimize the performance of the proposed method on selected training data from the Bosphorus database, one image per subject with a random pose. The system's rank-1 recognition rate with different values of N_r and w_b is displayed in Figure 4. From Figure 4, the system has the best performance when N_r=4 and w_b=1. In the next experiment, the system's performance versus the THD's order is evaluated (see Figure 5). Results show that the second-order THD is the best for the proposed method.

We evaluate the robustness of the proposed method using faces with pose variations from Bosphorus database. in this paper, we used only one frontal image per enrolled subject is used as the gallery. There are three different poses (yaw, pitch and mixture) for each subject, and we used a total of 1155 images for all 105 subjects as probes. The rank-1 recognition rates of the proposed method are shown in Table 1. In yaw rotations, our method achieves 85.1% recognition rate, while the benchmarks [18] and [11] have 38.3% and 67.6% recognition rates, respectively. Again for pitch rotations, recognition rate of the proposed method is 92.5% compared to 76.2% and 80.9% for the benchmarks. The rank-1 recognition rate of the proposed method for faces with mixture rotation is 75.2% compared to 26.7% and 49.1% for the benchmarks.

We compared the proposed algorithm using the FRGC v2 database with methods used in [19] and [20] as the benchmark. We selected one neutral face for each class (person) to create the gallery; 466 training face images for 466 classes in total. The rank-1 recognition rate of the proposed method is shown in Table 2. As can be seen, the proposed method has achieved a recognition rate of 98.1% and 92% for the neutral versus neutral and for the neutral versus non- neutral scenarios, respectively. The proposed method performs better than [20], and is comparable to [19]. THD maintains a high level of accuracy even on the faces having expressions, due to the robustness of THD to surface deformations. Note that the above results are obtained on the FRGC v2 dataset while the parameters were trained on a subset of the Bosphorus dataset. This proves that our approach has not been over-fitted to any database.

Table 1. Comparing the rank-1 recognition rate of THD and the benchmarks in faces with pose varations.

Considered Rotation	Malassiotis and Strintzis [18]	Hajati et al. [11]	THD
Yaw	38.3%	67.6%	**85.1**%
Pitch	76.2%	80.9%	**92.5**%
Mixture	26.7%	49.1%	**75.2**%
Average	47.1%	65.9%	**84.3**%

Table 2. Comparing the rank-1 recognition rate of THD and the benchmarks in faces with expression variations.

Considered Scenario	Berretti et al. [20]	Mian et al. [19]	THD
Neutral vs. Neutral	96%	99%	**98.1**%
Non-neutral vs. Neutral	91.4%	95.4%	**92**%
All vs. Neutral	94%	97.7%	**94.7**%

4. CONCLUSION

In this paper, a novel algorithm called Topographic High-order Derivatives (THD) was presented for 3D face recognition. Topography map of the face was partitioned into a predefined number of patches in both the texture and the height maps and high-order micro-pattern features were extracted from patches. To model the distribution of THD, an ensemble of spatial histograms is extracted as the representation of the input face image. We evaluated THD for pose and expression invariant face recognition task. Considerable improvements in the face recognition performance were achieved compared to the benchmark methods. Experimental results demonstrated that the proposed algorithm performs well on large databases such as the FRGC v2.

ACKNOWLEDGEMENT

Ajmal S. Mian was supported by ARC Discovery grant DP110102399.

REFERENCES

[1] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1042–1052, 1993.

[2] L. Nanni and D. Maio, "Weighted sub-gabor for face recognition," *Pattern Recognition Letters*, vol. 28, pp. 487–492, 2007.

[3] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 2037–2041, 2006.

[4] H. R. Kanan, Y. Gao, and K. Faez, "Face recognition using adaptively weighted patch pzm array from a single exemplar image per person," *Pattern Recognition*, vol. 41, pp. 3799–3812, 2008.

[5] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor," *IEEE Trans. on Image Processing*, vol. 19, pp. 553–544, 2010.

[6] K. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3d and multi-modal 3d + 2d face recognition," *Computer Vision and Image Understanding*, vol. 101, pp. 1–15, 2006.

[7] J. Cook, V. Chandran, and C. Fookes, "3d face recognition using log-gabor templates," in *Proc. of the International Conference on British Machine Vision*, 2006, pp. 769–778.

[8] K.I. Chang, K.W. Bowyer, and P.J. Flynn, "Multiple nose region matching for 3d face recognition under varying facial expression," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 1695–1700, 2006.

[9] N. Alyza, B. Gkberk, and L. Akarun, "Regional registration for expression resistant 3d face recognition," *IEEE Trans. on Information Forensics and Security*, vol. 5, pp. 425–440, 2010.

[10] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," *J. Optical Soc. Am*, vol. 14, pp. 1724–1733, 1997.

[11] F. Hajati, A. Raie, and Y. Gao, "2.5d face recognition using patch geodesic moments," *Pattern Recognition*, vol. 45, pp. 969–982, 2011.

[12] P. D. A. Harvey, *The History of Topographical Maps: Symbols, Pictures and Surveys*, Thames and Hudson, 1980.

[13] C. Richard, H. Robbins, and I. Stewart, *What Is Mathematics?: An Elementary Approach to Ideas and Methods*, Oxford University Press, 1996.

[14] H. Zhang, W. Gao, X. Chen, and D. Zhao, "Learning informative features for spatial histogram-based object detection," in *Proc. Of the International Conference on Neural Networks*, 2005, pp. 1806–1811.

[15] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, pp. 11–32, 1991.

[16] A. Savran, N. Alyuz, H. Dibeklioglu, O. Celiktutan, B. Gokberk, B. Sankur, and L. Akarun, "Bosphorus database for 3d face analysis," in *Proc. of the First COST 2101 Workshop on Biometrics and Identity Management (BIOID)*, 2008, pp. 47–56.

[17] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, and W. Worek, "Preliminary face recognition grand challenge results," in *Proc. International Conference of Automatic Face and Gesture Recognition*, 2006, pp. 15–24.

[18] S. Malassiotis and M. G. Strintzis, "Robust face recognition using 2d and 3d data: Pose and illumination compensation," *Pattern Recognition*, vol. 38, pp. 2537–2548, 2005.

[19] A. S. Mian, M. Bennamoun, and R. A. Owens, "An efficient multimodal 2d-3d hybrid approach to automatic face recognition," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 29, pp. 1927–1943, 2007.

[20] S. Berretti, A. D. Bimbo, and P. Pala, "3d face recognition using isogeodesic stripes," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 32, pp. 2162–2177, 2010.