3D Face Recognition on Point Cloud Data
An Approaching based on Curvature Map Projection using Low Resolution Devices

Luis Felipe de Melo Nunes[2], Caue Zaghetto[2] and Flavio de Barros Vidal[1]
[1]*Department of Computer Science, University of Brasilia, Brazil*
[2]*Department of Mechanical Engineering, University of Brasilia, Brazil*

Keywords: Point Cloud, Face Recognition, Curvature Maps, Three-dimension Face Data, Low Resolution Device.

Abstract: Facial recognition is the most natural and common form of biometrics, routinely used by humans and one of the most promising areas in biometrics research. The majority of traditional researches and commercial use of facial recognition systems are focused on methods that explores 2D (two-dimensional) images of human faces. All of them are based on features extraction that does not use any 3D shape information from the faces, especially with regard to depth. This paper presents a method based on Point Cloud and Curvature Map Projection to perform a 3D face recognition. The achieved results are presented and divided in two test scenarios, composed by a biometric evaluation analysis applying the Equal Error Rate score, Receiver Operating Characteristic and an accuracy comparison with other related works. The proposed work presents an accuracy of about 98.92%, allowing it to be applied for 3D face recognition tasks.

1 INTRODUCTION

The increasing need to monitor and restrict access to information or environments has led to major efforts towards the development of a variety of security mechanisms, such as biometric systems (Jain et al., 2000). In addition to applications related to access control, there are also others associated with civil identification and criminal investigation. To properly identify a user, biometric systems must rely on traits that present sufficient levels of universality, distinctiveness, permanence, collectability, acceptability and circumvention (Jain et al., 2004).

Among the various ways of performing biometrics, it is possible to highlight the facial recognition. Undoubtedly, facial recognition is the most natural and common form of biometric routinely used by humans and one of the most promising areas in biometrics research (Soldera et al., 2017). In general, facial recognition algorithms uses facial shape and their spatial relationships to perform individuals recognition (Jain et al., 2000). Although a human being is able to recognize a human face in an unfamiliar environment in approximately 100-200 ms, to a computer, running the best yet existing algorithms, it is still a challenge to accomplish this kind of task (Haykin and Network, 2004). It is true that in the last decade the reliability of face recognition algorithms has been improved, but in unconstrained environments problems such as uncontrolled illumination, head pose, facial expression and partial occlusion are still a bottleneck to these algorithms to achieve higher efficiency (Soldera et al., 2017).

The majority of traditional research and commercial use of facial recognition systems are focused on methods that explores 2D (two-dimensional) images of human faces (Bowyer et al., 2006). These methods, in general, are based on features extraction that does not take into account the 3D shape of faces, especially with regard to depth.

This paper presents a method based on Point Cloud and curvature map projection in order to perform face recognition. To the best of our knowledge, although some works have already addressed the challenge of 3D facial recognition (Patil et al., 2015), the solution presented in this paper is the only one that uses Point Cloud data and FCM method, applied to a dataset of 3D face images acquired by a low cost sensor device, to perform the task.

The remainder of this paper is divided into five sections: In Section 2, related works are discussed; Background concepts and the proposed method are

Nunes, L., Zaghetto, C. and Vidal, F.
3D Face Recognition on Point Cloud Data - An Approaching based on Curvature Map Projection using Low Resolution Devices.
DOI: 10.5220/0006843702660273
In *Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2018) - Volume 2*, pages 266-273
ISBN: 978-989-758-321-6

presented and discussed in the Section 3; Experimental results are shown in Section 4; and, finally, Section 5 presents conclusions and future works.

2 RELATED WORKS

Although there is a considerable number of work related to facial recognition, few of them present a solution that explores 3D facial characteristics and morphology, in the way that this work does, as will be explained. Face recognition systems are among the most reliable biometric systems. They are totally unobtrusive and a natural mode of identification among humans (Jain et al., 2004). In well-behaved environments, the performance can be compared to fingerprints (Zaghetto et al., 2017). However, face recognition is still a challenge, since its accuracy is reduced due to a number of factors, such as illumination, pose, distance and many others (Burrows and Cohn, 2009).

Towards improvement of facial recognition systems, a variety of recent solutions have been proposed. The work of (Haghighat et al., 2016), for instance, present a fully automatic face recognition system robust to most common face variations in unconstrained environments. The system is capable of recognizing faces from non-frontal views and under different illumination conditions using only a single gallery sample for each subject.

Another approach is proposed by (Borgi et al., 2015), which addresses the problem using a multi-scale directional framework called Shearlet Network (SN) to extract facial features, and, a refinement of the Multi-Task Sparse Learning (MTSL) framework. One should note that, although previously mentioned works address the problem of facial recognition, they do not explore the morphological three-dimensional (3D) characteristics of faces.

As it is known, 2D images are very sensitive to illumination changes (Papatheodorou and Rueckert, 2007). Given this fact and the fact that controlling light in real scenarios are not an easy task, 2D face recognition biometric systems will never be free from this weakness. Training algorithms using different illumination scenarios as well as illumination normalization of 2D images has been used, but with limited success (Papatheodorou and Rueckert, 2007).

In 3D images, however, variations in illumination do only affect the texture of the image, preserving its shape and morphological three-dimensional (3D) characteristics intact (Hesher et al., 2003). To overcome the problem of illumination on 2D images, the work of (Chen et al., 2017) proposes that 3D features may be extracted from 2D images. Based on making full use of advantages of Sparse Preserving Projection (SPP) on feature extraction, the discriminant information was introduced into SPP to arrive at a novel supervised feather extraction method, that named Uncorrelated Discriminant SPP (UDSPP) algorithm.

Although (Chen et al., 2017) uses 3D features, it does not work with real 3D images arguing that "although the 3D model method can achieve satisfactory recognition rate, it needs to pay higher computation cost". The work of (Hu et al., 2014) proposed likewise a facial recognition method which only uses 3D images in the face detection process.

Another recent work that addresses the problem of 3D facial recognition was proposed by (Kim et al., 2017). In this work, a novel 3D face recognition algorithm using a deep convolutional neural network (DCNN) and a 3D augmentation technique is proposed. It is mentioned that "training discriminative deep features for 3D face recognition is very difficult due to the lack of large-scale 3D face datasets", so a CNN is trained on 2D face dataset and then applied to 3D face recognition. Here it should be remarked that 3D facial recognition is still an open field to improvement, either because it demands high computational power or because it lacks of a large dataset to train algorithms or validate results. The work of (Zhou et al., 2015) proposes a real time 3D face recognition utilizing a trained two-level cascade classifier and preprocessing the RGB and depth data.

A recent work of (Goswami et al., 2014) proposes the unification of 2D and 3D information in order to accomplish a hybrid face recognition, applying techniques of entropy and saliency to construct a descriptor and utilizing geometrical analysis of 3D fiducial points. A complete survey of 3D facial recognition is presented in (Patil et al., 2015).

Table 1 summarizes the comparison between the related work and the proposed method in this paper.

3 PROPOSED METHOD

In this section, to describe our proposed methodology all stages are presented in the fluxogram of the Figure 1. The methodology is summarized into four main steps: *3D Cloud points preprocessing*, *Face Curvature Maps*, *Features Extraction* and *Similarity Matching*, all of them described below in Subsections 3.1 up to 3.4, respectively.

3.1 3D Cloud Points Preprocessing

Firstly, all available 3D data captured contains three-dimensional information of a full human upper-body.

Table 1: Comparison between methods using the characteristics: 3D recognition, real 3D dataset, low Cost sensor, no training, Low computational power, Point Cloud e Face Curvature Maps (FCM).

Method	**Characteristics** 3D Recog.	Real 3D dataset	Low cost sensor	No training	Low comp. power	Point Cloud	FCM
Haghighat, Abdel-Mottaleb, and Alhalabi (2016)			X		X		
Borgi, Labate, Elarbi, and Amar (2015)			X		X		
Hu N. et al. (2014)			X		X		
Chen Zhanwei, Huang Wei, and Zhihan Lv (2017)	X		X				
Kim D. et al. (2017)	X	X					
Guswami, Vatsa and Singh (2014)	X	X	X				
Zhou, Chen and Wang (2015)	X	X	X		X		
Proposed method	**X**	**X**	**X**	**X**	**X**	**X**	**X**

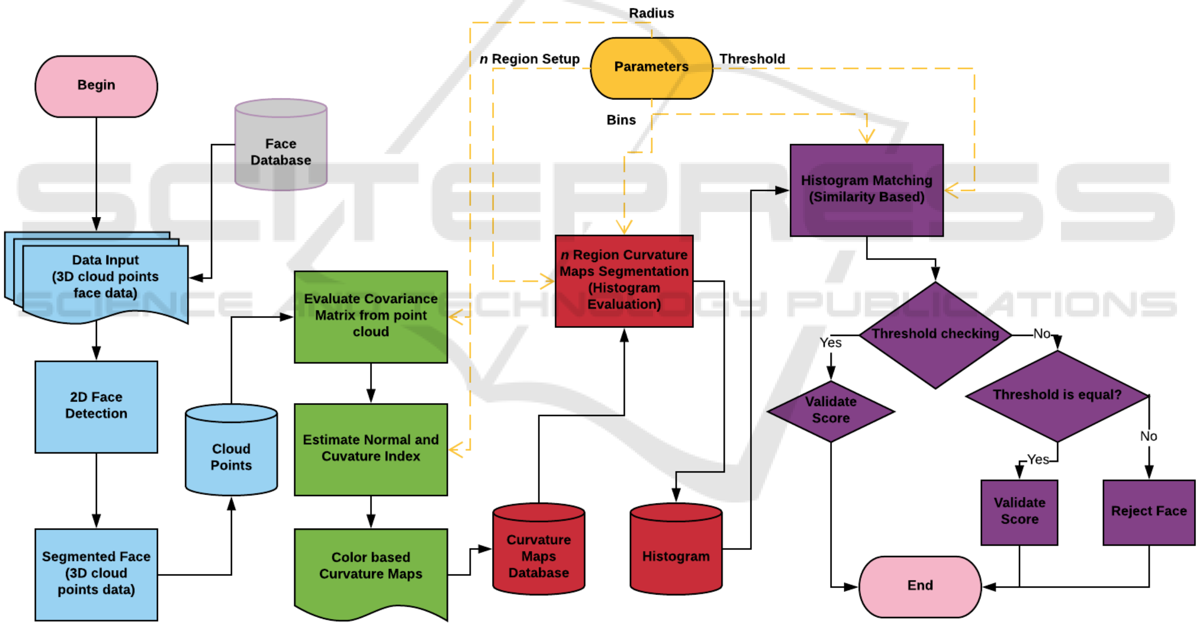


Figure 1: Fluxogram of the Proposed Methodology.

In this case, the first preprocessing is to define a bounded area which includes the face using the *Viola-Jones'* algorithm (Viola and Jones, 2004). Once this technique was developed for a two dimensional data, we developed a simple adapted technique that allows extract the three dimensional data using the two dimensional information (vertical and horizontal positions) from the face, due to a spacial correlation between 2D (RGB) and 3D (Depth) data. In Figure 1 these steps are described by light-blue color. Samples of this preprocessing step are described in Figure 2.

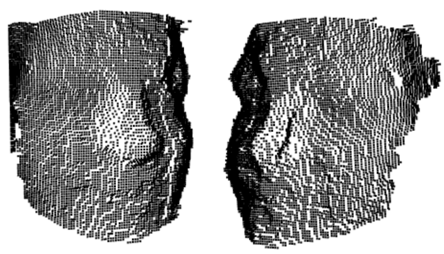


Figure 2: Samples results of the 3D preprocessing step - Subset S_{face}.

All results from this preprocessing step are defined as a subset(S_{face}) of three-dimensional data points delimited by the horizontal and vertical spatial dimensions containing a face of the individual/subject to be recognized/verified and described by Equation 1.

$$S_{face} \subset \Omega \in \mathbb{R}^3 \tag{1}$$

where Ω is a set of three dimensional data in spacial domain $\mathbb{R}^3$. The subset S_{face} consists of points $p_k(x,y,z)$ where $k = 1 \ldots w \ldots m$.

3.2 Face Curvature Maps

Defining M_{cov} as the Covariance Matrix of points $p_k(x,y,z)$, described in Equation 2, which evaluates covariance around of w three dimensional neighbor points, as follows:

$$M_{cov} = \frac{1}{w_n} \sum_{i=1}^{w_n} (p_i - \bar{p}) \cdot (p_i - \bar{p})^T, \tag{2}$$

where $\bar{p}$ is the centroid position of region bounded by a defined radius r with w_n neighbors of set points. The relationship among the set of p_i points and the radius r is described on Figure 3.

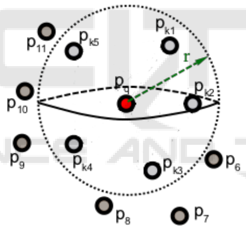


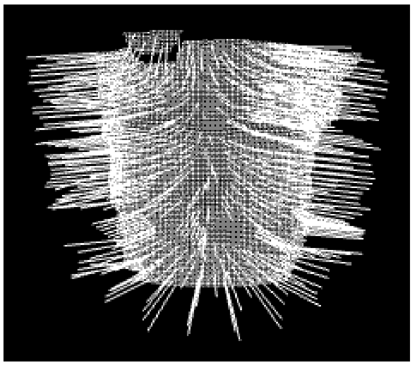
Figure 3: Relationship among the set of p_i points and the radius r.

All face normal curvature indexes (C_v) are evaluated by Equations 3 and 4.

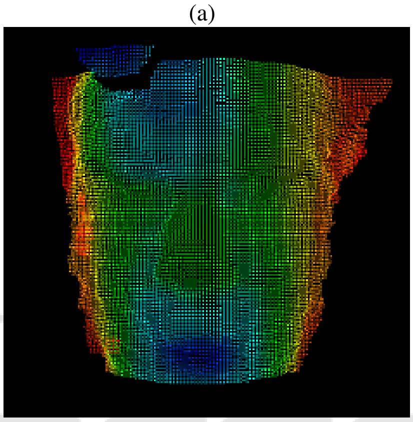
$$M_{cov} \cdot \vec{v}_j = \sigma_j \cdot \vec{v}_j, \quad j \in \{0,1,2\}, \tag{3}$$

$$C_v = \frac{\sigma_0}{\sigma_0 + \sigma_1 + \sigma_2} \tag{4}$$

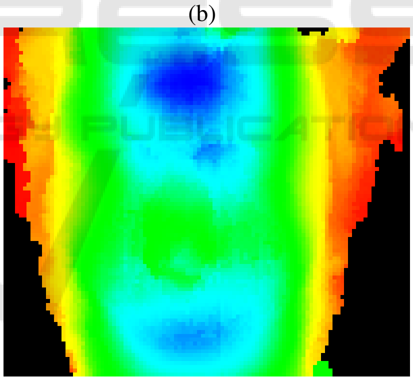
where σ_j and $\vec{v}_j$ are the eigenvalues and eigenvector of matrix M_{cov} respectively for each p_i and $\sigma_0 < \sigma 1 < \sigma 2$. In the next step, all extracted C_v indexes are normalized to values between 0(minimal value) up to 1(maximum value) and described as a face intensity color map, as described in Figure 4-(a) and (b) as follows. Note that using this projection process, all three-dimensional information are included in a two-dimensional image, allowing to use any classic face recognition technique, for example the *EigenFaces*(Belhumeur et al., 1997).



(a)



(b)



(c)

Figure 4: In (a) the Extracted Normal indexes and (b) curvature color maps. For (c) is the normalized to 2D curvature color map information.

3.3 Features Extraction

From curvature indexes C_v all the geometric information are carried out of each set of points from a subset of 3D points clouds. The used features in the matching stage are defined in according to each evaluated histogram H in Equation 5.

$$m = \sum_{i=1}^{b} H_i \tag{5}$$

where b is the previously defined number of bins and m, as previously defined, is the number of p_k points in a face subset S_{face}. The variable b defines how many features sets will be formed and for high values of variable b, it implies in the addition of curvatures information. Otherwise, for reduced values of b, less curvature details are used in the feature vectors.

3.4 Similarity Matching

In according to (Jain et al., 2004), a biometric system is essentially a pattern recognition system that operates by obtaining biometric data of an individual, extracting a set of features of the acquired data and comparing this set of features with the ones already stored in a database. Depending on the context, a biometric system can work in a verification or identification mode. In the verification mode, the system must validate the identity of an individual comparing the captured biometric data with the previously captured and stored data in the database. In the identification mode, the system must recognize an individual by comparing the biometric data with all others previously stored in the database, searching for the most similar one.

In this case we are focused on to solve a face verification problem using a Similarity Matching schema presented by (Jain et al., 2004), that can be formally defined as: Given an input vector of curvature indexes features C_v extracted from the 3D face data and an alleged identity I, determine if (I, C_v) belongs to the class f_1 or f_2, where f_1 indicates that the alleged identity is true and f_2 that it's false. C_v is compared with C_I, as the vector of biometric features of the individual I, to determine its class. Thus

$$(I, C_v) \in \begin{cases} f_1, & \text{if } S(C_v, C_I) \geq t \\ f_2, & \text{otherwise} \end{cases} \quad (6)$$

where S is a function that measures the similarity score between the vectors C_v and C_I, and t is the predefined threshold. $S(C_v, C_I)$ is called similarity matching score between the biometric features of the individual and the alleged identity. The identification problem can be formally defined as: given as entry a vector of features C_v, determine if the identity I_k, where $k \in \{1, 2, ..., N, N+1\}$. Here $I_1, I_2, ..., I_N$ are the identities already in the system and I_{N+1} indicates the rejected case, where no identity is compatible with the users. Thus

$$C_v \in \begin{cases} I_k, & \text{if } \max_k \{S(C_v, C_{Ik})\} \geq t, k = 1, 2, ..., N \\ I_{N+1}, & \text{otherwise} \end{cases} \quad (7)$$

where C_{Ik} is the vector of biometric features corresponding to the identity I_k, and t is a predefined threshold.

4 RESULTS

In this Section are presented the evaluation of the proposed methodology, describing details about the used database and showing results in two specific tests scenarios as follows.

4.1 Database

As previously presented by the fluxogram on Figure 1, the VAP RGB-D Face Database by (Hg et al., 2012) was used in order to perform the tests scenarios to evaluate the proposed methodology. This database provides 31 different subjects, each of them containing 17 sets of RGB-D images on different poses and facial expressions (13 poses and 4 facial expressions). Each pose or expression presents 3 images samples composed by a RGB color space and a Depth data, both registered (allowing the correlation of features between them easily by a translation transform). The sensor used for data acquisition of this database was the *Microsoft Kinect* first generation (Zhang, 2012). All acquired depth images were filtered in order to treat occlusions and spikes to obtain a smoother and hole-free Point Cloud data representation of each subject as described in (Hg et al., 2012).

The main objective of this proposed methodology is to develop a face recognition algorithm in tests scenarios using only faces with the absence of rotation and occlusion (generated by the face position in relation to the camera). Specifically in this case, only the subject's frontal face pose was used to allow evaluate the influence of all parameters used to adjust the algorithm's behavior on the facial recognition process.

4.2 Evaluation Process

The evaluation process is composed by two tests scenarios, as described as follows:

4.2.1 First Scenario - Biometric Evaluation

In this scenario each subject - in a recognition (classification application) system - can be treated as a class. The used method to provide a possible identification (intra-personal score minimization) is analogue to a classifier described in Section 3.4. The Equal Error Rate (EER)(Trentin and Gori, 2001) can be defined as an objective, threshold-independent measure of the classifier's performance for statistical pattern recognition, which is used to evaluate this classifier and commonly used to evaluate biometric systems.

To evaluate and minimize the errors of the proposed methodology, the most influential variables were selected (radius, bins), and then EER was applied to each one of these variables, defining the optimal value for them.

The rejection criterion established to the ERR was based in the score obtained from the minimization function. The first criteria defined is the maximum threshold, avoiding the false recognition from unknown subjects and refusing badly acquired sensor outputs. The second criteria is defined by a threshold interval limitation between the two best enroll matches if they do not belong to the same subject (since each subject has 3 images samples), allowing the system to presume doubt between two subjects, and reject the input.

The radius, responsible for the vicinity description of a point and, consequently, the curvature intensity was analyzed in step intervals of 5, generating different curvature maps and score for each radius value from minimization function. In Figure 5 it is possible to visualize the rates values of false acceptance and false rejection obtained for each radius values, obtaining an interception at the radius of 26.67, resulting in an ERR of 3,58% for acceptance and rejection.

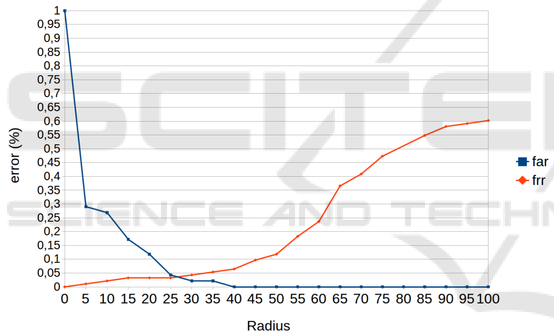


Figure 5: Equal Error Rate of Vicinity Radius.

Once the radius is the most independent and the first required variable to be defined in the fluxogram of the proposed methodology, all remaining variables will use this optimal radius value fixed as reference to define their optimal values from the EER method.

The subsequent analyzed variable is the number of bins of the intensity histograms obtained from the curvature maps. The ERR was applied with the same previously criterion, with radius value fixed during this analysis, changing only the number of bins used to represent the histogram as described in Figure 6.

Although it is expected to obtain a better classifier result using a higher number of bins representation (and consequently intensity distribution), that doesn't guarantee a best discriminative value among subjects. In Figure 6, values from 8 up to 16 bins achieved an error of 2,5%, but the number of bins are restricted to be a power of 2 (to guarantee the equal numerical distribution of the intensity values [ranged from 0 to 255]) and located in the exact average of 8 and 16, both of these values are optimal for the number of bins in this application. The use of values greater than 16 bins provided a higher disparity between the images of a same subject, possibly due to the information loss (caused by mainly by filtering process) and the occlusions filling estimation done by (Hg et al., 2012) preprocessing, which causes a higher rejection rate and consequently false rejection as well.

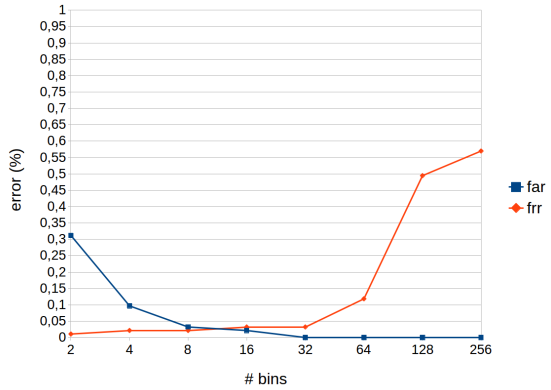


Figure 6: Equal Error Rate of the number of bins.

4.2.2 Second Scenario - Accuracy Comparison

In order to obtain a real performance evaluation, this scenario was developed to compare the performance of the proposed methodology with others state-of-the-art techniques related to the used data type and the face recognition task. This evaluation is focused on the Rank-1 Accuracy (Lathauwer et al., 2000) of the facial recognition process, presented as the most usual evaluation method found in the techniques used in to comparison. Also, the True Positive Rates (TPR) and False Positive Rates (FPR) were computed for both variables Radius and Bins. Receiver Operating Characteristic (ROC) Curves and Area Under Curves (AUC) values for accuracy estimates are shown in Figure 7.

The Rank-1 Accuracy is obtained by a ratio of the relevant samples from the recognition process (true positives and true negatives) and the total subject's enrollments in the data base. From extensive tests used to define parameters of the proposed methodology, presented in previously test scenario, the best recognition results achieved an accuracy of 98,92% and used to be compared with the best performance of others state-of-the-art techniques as described in Table 2.

To obtain a fair comparison, the selected techniques were obtained related to the database presenting the identification task in facial recognition as well,

Table 2: The best Rank 1 Accuracy of the face recognition algorithms related to the VAP RGB-D database.

Data Type	**Method**	**Rank 1 (%)**
RGB Images + Depth Map	Goswami et al.(Goswami et al., 2014)	80.6
RGB Images + Depth Image	Hu et al. (Hu et al., 2014)	90.0
RGB Images + Depth Image	Bormann et al. (Bormann et al., 2013)	96.0
RGB Images + Depth Map	Zhou et al. (Zhou et al., 2015)	95.9
Depth Map	Saleh and Edirisinghe (Saleh and Edirisinghe, 2016)	96.67
RGB Images + Depth Map	Chowdhury et al. (Chowdhury et al., 2016)	98.71
Point Cloud	**Proposed Methodology**	**98.92**

based on different techniques. In Goswami *et al.* (Goswami et al., 2013) a method is proposed to extract an entropy map from the depth map and the RGB image of a person and a saliency map from the RGB image, computing a histogram of gradient (HOG) from these maps and classifying them by a Random Forest (RDF). Other work from Goswami *et al.* (Goswami et al., 2014) has presented improvements adding a geometric attribute computation from depth map fiducial points, creating the called RISE (entropy and saliency maps) and ADM (geometric attribute relation) descriptors.

In Hu *et al.* (Hu et al., 2014) was proposed a face recognition for a user tracking robotics application, using the depth map from head detection and the RGB image for recognition by illumination normalization, head pose correction and face space projection. Bormann *et al.* (Bormann et al., 2013) implements a similar algorithm to Hu *et al.* algorithm, Fisherfaces(Belhumeur et al., 1997) space parameterization, a Support Vector Machine (SVM) and Nearest Neighbor techniques for classification. Zhou *et al.* (Zhou et al., 2015) proposed a three-dimensional face recognition using 7 feature points and a two-level Cascade Classifier, formed by a Decision Tree Classifier in the first level, and an improved Euclidian Distance classifier in the second level. Saleh and Edirisinghe proposed an Eigenface-based method, training models with eigenfaces applied to the normal images and depth images, under different illumination conditions. Chowdhury *et al.* (Chowdhury et al., 2016) proposed a method based on machine learning, that trains a Neural Network to reconstruct the depth map from a color image, using the color image and the real depth map as input elements, and classifying the reconstructed depth map through another multi class neural network.

Although close in accuracy performance results (Figure 7), all considered values above the limits of excellent (i.e. AUC $\approx$ 0.90), the estimate for Bin values (AUC = 0.92) is a 0.02 better than for Radius vicinity values (AUC = 0.90). All curves are well over the chance value (i.e. AUC = 0.50).

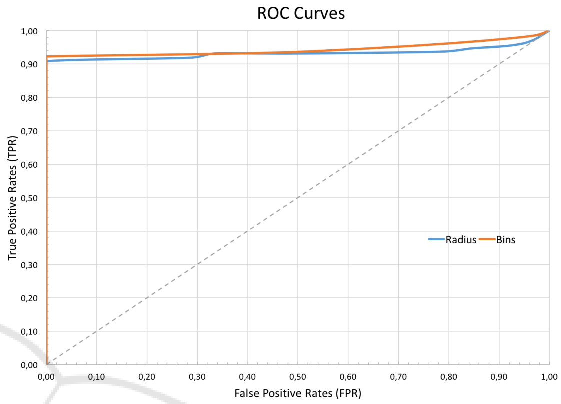


Figure 7: ROC Curves of Radius and Bins.

5 CONCLUSIONS

This work proposes a methodology based on Point Cloud face data using curvature map projection for face recognition. The proposed methodology was evaluate on two scenarios for biometrics and accuracy that describe all features and details about the parameters setup and performance influence on face recognition process.

The Table 2 presents all the Rank-1 Accuracy obtained in each state-of-the-art technique and concludes that the results obtained by the proposed methodology outperform the best achieved results of all the mentioned techniques. This information demonstrates that the proposed methodology is qualified to be applied in three-dimensional face recognition problems. In the experiments, the accuracy was also evaluated plotting Receiver Operating Characteristic (ROC) Curves and achieving excellent result for both analyzed variables.

All analysis must be carefully accomplished, since the implemented techniques in each work are different and there are specifics variations in relation to the used data base, the used data type and their experimentations. This allows to archive the conclusion that our proposed methodology presents competitive scores compared to all those techniques. In the end, the proposed test scenarios relied on the use of the VAP RGB-D database (Hg et al., 2012) (due to its public availability), and the full comparison to others techniques were impaired due to their experimentations in private data bases. Future works seek to experiment the proposed methodology in private databases to obtain broader results.

REFERENCES

Belhumeur, P. N., Hespanha, J. a. P., and Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(7):711–720.

Borgi, M. A., Labate, D., El Arbi, M., and Amar, C. B. (2015). Sparse multi-stage regularized feature learning for robust face recognition. *Expert Systems with Applications*, 42(1):269–279.

Bormann, R., Zwölfer, T., Fischer, J., Hampp, J., and Hägele, M. (2013). Person recognition for service robotics applications. In *Humanoid Robots (Humanoids), 2013 13th IEEE-RAS International Conference on*, pages 260–267. IEEE.

Bowyer, K. W., Chang, K., and Flynn, P. (2006). A survey of approaches and challenges in 3d and multi-modal 3d+ 2d face recognition. *Computer vision and image understanding*, 101(1):1–15.

Burrows, A. M. and Cohn, J. F. (2009). Anatomy of face. In *Encyclopedia of Biometrics*, pages 16–23. Springer.

Chen, Z., Huang, W., and Lv, Z. (2017). Towards a face recognition method based on uncorrelated discriminant sparse preserving projection. *Multimedia Tools and Applications*, 76(17):17669–17683.

Chowdhury, A., Ghosh, S., Singh, R., and Vatsa, M. (2016). Rgb-d face recognition via learning-based reconstruction. In *Biometrics Theory, Applications and Systems (BTAS), 2016 IEEE 8th International Conference on*, pages 1–7. IEEE.

Goswami, G., Bharadwaj, S., Vatsa, M., and Singh, R. (2013). On rgb-d face recognition using kinect. In *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*, pages 1–6. IEEE.

Goswami, G., Vatsa, M., and Singh, R. (2014). Rgb-d face recognition with texture and attribute features. *IEEE Transactions on Information Forensics and Security*, 9(10):1629–1640.

Haghighat, M., Abdel-Mottaleb, M., and Alhalabi, W. (2016). Fully automatic face normalization and single sample face recognition in unconstrained environments. *Expert Systems with Applications*, 47:23–34.

Haykin, S. and Network, N. (2004). A comprehensive foundation. *Neural Networks*, 2(2004):41.

Hesher, C., Srivastava, A., and Erlebacher, G. (2003). A novel technique for face recognition using range imaging. In *Signal processing and its applications, 2003. Proceedings. Seventh international symposium on*, volume 2, pages 201–204. IEEE.

Hg, R., Jasek, P., Rofidal, C., Nasrollahi, K., Moeslund, T. B., and Tranchet, G. (2012). An rgb-d database using microsoft's kinect for windows for face detection. In *Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on*, pages 42–46. IEEE.

Hu, N., Bormann, R., Zwölfer, T., and Kröse, B. (2014). Multi-user identification and efficient user approaching by fusing robot and ambient sensors. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 5299–5306. IEEE.

Jain, A., Hong, L., and Pankanti, S. (2000). Biometric identification. *Communications of the ACM*, 43(2):90–98.

Jain, A. K., Ross, A., and Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1):4–20.

Kim, D., Hernandez, M., Choi, J., and Medioni, G. (2017). Deep 3d face identification. *arXiv preprint arXiv:1703.10714*.

Lathauwer, L. D., Moor, B. D., and Vandewalle, J. (2000). On the best rank-1 and rank-(r1,r2,. . .,rn) approximation of higher-order tensors. *SIAM J. Matrix Anal. Appl.*, 21(4):1324–1342.

Papatheodorou, T. and Rueckert, D. (2007). 3d face recognition. In *Face Recognition*. InTech.

Patil, H., Kothari, A., and Bhurchandi, K. (2015). 3-d face recognition: features, databases, algorithms and challenges. *Artificial Intelligence Review*, 44(3):393–441.

Saleh, Y. and Edirisinghe, E. (2016). Novel approach to enhance face recognition using depth maps. In *Systems, Signals and Image Processing (IWSSIP), 2016 International Conference on*, pages 1–4. IEEE.

Soldera, J., Schu, G., Schardosim, L. R., and Beltrao, E. T. (2017). Facial biometrics and applications. *IEEE Instrumentation & Measurement Magazine*, 20(2):4–10.

Trentin, E. and Gori, M. (2001). A survey of hybrid ann/hmm models for automatic speech recognition. *Neurocomputing*, 37(1):91–126.

Viola, P. and Jones, M. J. (2004). Robust real-time face detection. *Int. J. Comput. Vision*, 57(2):137–154.

Zaghetto, C., Aguiar, L. H. M., Zaghetto, A., Ralha, C. G., and de Barros Vidal, F. (2017). Agent-based framework to individual tracking in unconstrained environments. *Expert Systems with Applications*, 87:118–128.

Zhang, Z. (2012). Microsoft kinect sensor and its effect. *IEEE MultiMedia*, 19(2):4–10.

Zhou, W., Chen, J.-x., and Wang, L. (2015). A rgb-d face recognition approach without confronting the camera. In *Computer and Communications (ICCC), 2015 IEEE International Conference on*, pages 109–114. IEEE.