Geometrical and Visual Feature Quantization for 3D Face Recognition

Walid Hariri[1,2], Hedi Tabia[1], Nadir Farah[2], David Declercq[1] and Abdallah Benouareth[2]
[1]*ETIS/ENSEA, University of Cergy-Pontoise, CNRS, UMR 8051, Cergy-Pontoise, France*
[2]*Labged Laboratory, Computer Science Department, Badji Mokhtar Annaba University, Annaba, Algeria*

Keywords: LBP, HoS, Bag-of-Features, Codebook, Depth Image, Term Vector.

Abstract: In this paper, we present an efficient method for 3D face recognition based on vector quantization of both geometrical and visual proprieties of the face. The method starts by describing each 3D face using a set of orderless features, and use then the Bag-of-Features paradigm to construct the face signature. We analyze the performance of three well-known classifiers: the Naïve Bayes, the Multilayer perceptron and the Random forests. The results reported on the FRGCv2 dataset show the effectiveness of our approach and prove that the method is robust to facial expression.

1 INTRODUCTION

Face recognition has recently gained a blooming attention and interest from the pattern recognition community (Jafri and Arabnia, 2009). Many applications use facial recognition including security oriented applications (access-control/verification systems, surveillance systems), computer entertainment and customized computer-human interaction. Most existing recognition methods are based on the 2D appearance of faces and discard their 3D shapes. This leads to a poor discrimination power when dealing with variation such as illumination, expressions, occlusion or head poses. The availability of 3D scanners and the rapid evolution in graphics hardware and software, have greatly facilitated a shift from 2D to 3D approaches. The advantage of the 3D representation is in having more discriminant information regarding the face's shape which is less sensitive to variations. Recent surveys of 3D face recognition advances can be found in (Luo et al., 2015; Mishra et al., 2015).

3D face data can be also combined with 2D face data to build multimodal approaches. Most efforts to date in this area use relatively simplistic approaches to fusing results obtained independently from the 3D data and the 2D data. A multimodal recognition system could perform better than any one of its individual components (Jyothi and Prabhakar, 2014; Bowyer et al., 2006; Chang et al., 2005).

To describe the visual proprieties of both 2D and 3D face data, Local Binary Patterns (LBPs) (Ojala et al., 2002) is a technique that has been widely applied for this purpose, especially in facial expression recognition (Shan et al., 2009) and face recognition (Ahonen et al., 2006; Huang et al., 2012). LBP has a several advantages such as its high discriminative power, its tolerance against illumination changes and its computational simplicity. A plenty of LBP based face recognition methods have been proposed in the literature (Huang et al., 2011b; Yang and Chen, 2013). Although, the various advantages of the LBP descriptor, few drawbacks, such as the lose of the global structure of the face, are still not handled. In this paper, we propose a compact face representation computed from two different modalities and aggregated using the conventional Bag-of-Features (BoF). The paradigm of BoF has successfully been applied in several domains such as shape classification, object detection and image retrieval (O'Hara and Draper, 2011). Here, we exploit both 3D and 2D face descriptions to get a robust 3D face recognition system. We firstly extract the Histogram of Shape index (HoS) which represent the surface planarity of each face region according to the values of its Shape index (*SI*). All shapes can be mapped into the interval: $SI \in [0,1]$, where each value represents a different shape (saddle, cup, dome, rut, ridge, etc.). Note that *SI* is mostly used to extract histogram as shape descriptor because it is independent of the scale variation. Next, once we have extracted 2D depth image from each 3D face mesh, we extract the LBP descriptor to describe the face depth variation. Note that in our method, instead of combining the extracted region histograms to get a global descriptor, we use BoF representation instead to get a global face description as orderless collections of local features, after concatenation of their term vectors. Finally, three different classifiers are applied to assess the performance of our method.

Hariri W., Tabia H., Farah N., Declercq D. and Benouareth A.
Geometrical and Visual Feature Quantization for 3D Face Recognition.
DOI: 10.5220/0006101701870193
In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017)*, pages 187-193
ISBN: 978-989-758-226-4

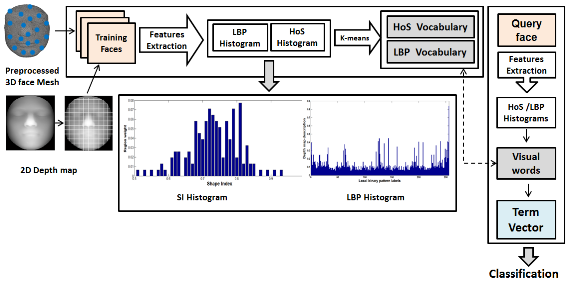

Figure 1: Overview of the proposed method.

The remainder of the paper is organized as follows, Section 2 presents the related works, the method is described in Section 3. We detail the face description in Section 4. In Section 5, we present the used classifiers. Experimental results and conclusion end the paper.

2 RELATED WORKS

The conventional design of recognition systems goes into two different steps; feature extraction and face comparison. Several feature extraction methods have been proposed in the literature; signature points (Chua et al., 2000), curves (Drira et al., 2013), landmarks and curvatures (Creusot et al., 2013), SIFT (Smeets et al., 2013), curvelet (Elaiwat et al., 2014), covariance (Tabia et al., 2014; Tabia and Laga, 2015; Hariri et al., 2016a; Hariri et al., 2016b). For face comparison, matching based technique such as Iterative Closest Point (ICP) registration (Besl and McKay, 1992) have been widely used. Machine learning based approaches have also been used for face comparison (SVM (Lei et al., 2013), Neural network (Sun et al., 2014; Ding and Tao, 2015), Random forests (Fanelli et al., 2013), Adaboost (Xu et al., 2009; Ballihi et al., 2012)). Most of these approaches construct a set of statistical rules which are then used to recognize unknown faces. Compared with matching methods, machine learning based approaches provide efficient solutions to deal with big size galleries.

In this paper, we have applied a supervised learning approach based on three well-known classifiers; the Naïve Bayes, the Multilayer perceptron and the Random forests.

3 METHOD

Figure 1 presents an overview of the proposed method. After the acquisition step, the input face surface is preprocessed to improve the quality of the input face which may contain some imperfections as holes, spikes and includes some undesired parts (clothes, neck, ears, hair, etc.) and so on. It consists of applying successively a set of filters. First, a smoothing filter is applied, which reduces spikes in the mesh surface, followed by a cropping filter which cuts and returns parts of the mesh inside an Euclidean sphere. Next a filling holes filter is applied, which identifies and fills holes in input meshes. Finally, to remove spikes, we apply a median filter on 3D face vertices. The filter starts by sorting the z coordinate within a neighborhood, finding then the median, and finally replacing the original z coordinate with the value of the median.

Once the 3D face mesh has been preprocessed, we firstly extract N feature points using a uniform sampling of the 3D face surface, the feature points are the center of N patches from a paving of the face. Next, from each 3D face, we extract the 2D depth image where the gray value of each image pixel represents the depth of the corresponding point on the 3D surface. All 2D faces are normalized to 150×150 pixels. This 2D representation is widely used in 3D facial analysis (Berretti et al., 2011; Huang et al., 2011a; Vretos et al., 2011). As an example, Figure 3 illustrates 2D depth map images derived from 3D face scans of the same subject. Then, we extract HoS and LBP descriptors from the 3D face mesh and its corresponding 2D depth image, respectively. Once the local descriptors have been extracted, we apply the BoF paradigm to build a final compact signature.

Finally, to describe the whole face using BoF representation, we build two different visual vocabularies (C_{Hos}, C_{Lbp}) using k-means clustering. Note that the number of visual words chosen for each vocabulary may be different. In our implementation, we use same size vocabularies ($k_h = k_l$, where k_h is the number of words in C_{Hos} and k_l is the number of words in C_{Lbp}). A term vector is then computed as the occurrence of each visual word in the face, and a global term vector of dimension $k = k_h + k_l$ is constructed. It is obtained by serially concatenating both HoS and LBP term vectors (see Figure 2).

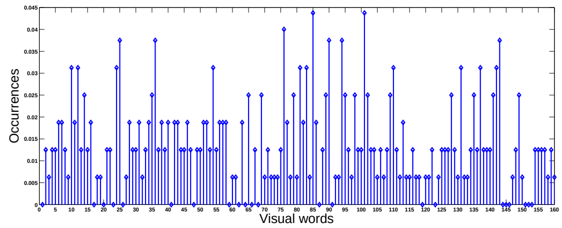

Figure 2: Global term vector from a query face. Items indicate the occurrence of each visual word in the face.

4 FACE DESCRIPTION AND FEATURES QUANTIZATION

In this section, we present the geometrical and visual features of faces, as well as their quantization using BoF paradigm.

4.1 Local Binary Pattern

LBP has originally been proposed by (Ojala et al., 2002) for texture classification, and then extended for various fields, including face recognition (Ahonen et al., 2006; Yang and Chen, 2013) and face detection (Sandbach et al., 2012). The LBP histogram compares each center pixel with its neighbors, encoding this relation into a binary word i.e. the ones whose intensities exceed the center pixel's are marked as (1), otherwise as (0). This allows detection of patterns, while being robust to contrast changes.

The 256-bin histogram of the labels computed over an image can be used as a texture descriptor. In this way we get a simple circular point features consisting of only binary bits. Typically the feature ring is unfolded as a row vector; and then with a binomial weight assigned to each bit, the row vector is transformed into decimal code for further use. Local primitives which are codified by these bins include different types of curved e.g. edges, spots, flat areas, etc. Figure 4 shows some regions detected by the uniform patterns of LBP.

Some LBP histogram-based methods change the neighborhood of the LBP operator for improved performance. By varying the value of radius, the LBP of different resolutions is thus obtained. For example, the operator $LBP_{4,1}$ uses 4 neighbors while $LBP_{16,2}$ considers the 16 neighbors on a circle of radius 2. In the following, we refer to the neighborhood size by p, where r is the circle radius that forms a circularly symmetric neighbor set. d_{lbp} is the LBP dimension (number of bins) which is obtained by: $d_{lbp} = 2^p$.

In our method, we extract LBP histogram for each face patch of 2D depth image, therefore, we compute 8-by-8 neighbors to get 256 bins histogram.

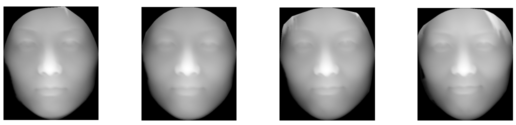

Figure 3: Facial depth images derived from 3D face scans.

4.2 Histogram of Shape Index

Shape index (SI) expresses different shape classes by a single number ranging from 0 to 1, it can be estimated by the following equation:

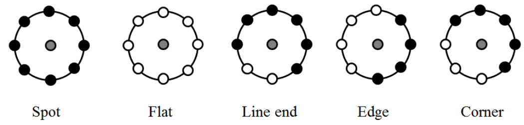

Figure 4: Regions detected by the uniform patterns of LBP. White circles represent ones and black circles zeros.

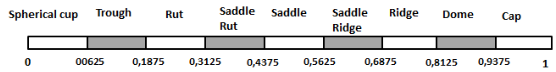

Figure 5: Nine shape types and their locations on the Shape index scale.

$$SI = \frac{1}{2} - \frac{1}{\pi} arctan \left(\frac{k_{max} + k_{min}}{k_{max} - k_{min}} \right) \quad (1)$$

Where k_{max} and k_{min} are the principal curvatures of the surface. In our method, the values of SI are quantized to 50 bins to get a histogram (HoS) of dimension: $d_{hos} = 50$. It will be then normalized to form a feature vector for each face region. Every distinct surface shape corresponds to a unique value of SI, except the planar shape. Points on a planar surface have an indeterminate SI, since $k_{max} = k_{min} = 0$.

The SI captures eight basic shape types of a surface which are based on the signs of Gaussian and mean curvatures that was employed by (Besl, 2012). Figure 5 presents these shapes according to their SI value.

4.3 Bag of Features

The BoF representation aims to aggregate local descriptors into a compact signature (Tabia et al., 2012; Tabia et al., 2011). We first learn a codebook $C = (w_1, w_2, ..., w_k)$ of k visual words (terms) obtained by k-means clustering. Note that in our method, we have two codebooks (C_{Hos}, C_{Lbp}) of dimension k_h and k_l respectively, where $k = k_h + k_l$ is the dimension of the global codebook C. The next step is to quantize local descriptors of dimension d_{hos}, d_{lbp} into a set of visual words. Hence, it produces a k-dimensional vector, which is subsequently normalized.

To build a BoF of an image, the following steps are required:

- *Build Vocabulary*: by clustering features from all faces in a training set.
- *Assign Terms*: by assigning the features from each face to the closest terms in the vocabulary.
- *Generate Term Vector*: obtained by counting the frequency of each term in the face. In this paper, we extract two term vectors; V_{hos} and V_{lbp} respectively for HoS and LBP descriptors. These

vectors are then grouped into one single vector $V = \{V_{hos}, V_{lbp}\}$.

5 CLASSIFIERS

In this section, we presents the three supervised classifiers used in our method, where each face is represented by a term vector. Note that we use gallery faces as training set, where probe faces are used for the test. In the following, we note $V = [v1, \ldots, v_k]$ the term vector of each face, where each v_i refers to the occurrence of the term i in the given face. k is the number of attributes, and m is the number of classes.

- *Naïve Bayes classifier*: are probabilistic classifiers based on the Bayesian theorem (Lewis, 1998). Given a face image to be classified to m possible outcomes of face subjects in the dataset. Each face image is represented by a term vector V of dimension k. The classifier assigns to this face image the probability $p(C_i|v_1, \ldots, v_k)$. The conditional probability can be decomposed as: $p(C_i|V) = \frac{p(C_i)\ p(\mathbf{v}|C_i)}{p(\mathbf{v})}$. In practice, the naïve conditional independence assumptions assume that each term occurrence v_i in the face is conditionally independent of every other term occurrence v_j in the same face, for $j \neq i$, given the subject identity C_i.
- *Random forest classifier*: consisting of a collection of tree-structured classifiers developed by (Breiman, 2001). In our experiments we used Random Forest algorithm by considering 50 trees. Classification of a new face from an input term vector V is performed by putting it down each of the trees $(t_1, \ldots, t_{10})$ in the forest F. Each tree t_i gives a classification decision d_i by voting for the face subject class C_i. According to the global decision $D = (d_1, \ldots, d_{10})$, the forest chooses the classification d_j having the most votes over all the trees in the forest.
- *Multilayer perceptron classifier*: to classify a probe face using its term vector V, it uses a set of term occurrence as input values (v_i) and associated weights (w_i) and a sigmoid function (g) that sums the weights and maps the results to an output (y). Note that the number of hidden layers used in our experiments is given by: $\frac{m+k}{2}$.

6 EXPERIMENTAL RESULTS

6.1 Experiments on FRGCv2 Dataset

FRGCv2 database (Phillips et al., 2005) is one of the most comprehensive and popular datasets, containing 4007 3D face scans of 466 different persons, the data were acquired using a minolta 910 laser scanner that produces range images with a resolution of 640×480. The scans were acquired in a controlled environment and exhibit large variations in facial expression and illumination conditions but limited pose variations. The subjects are 57% male and 43% female, with the following age distribution : 65% 18-22 years old, 18% 23-27and 17% 28 years or over. The database contains annotation information, such as gender and type of facial expression. This dataset was firstly preprocessed as described in the Section 3. In this experiment, we use neutral images from each subject as galleries and the rest images are used as probes.

First, we uniformly sample $N = 200$ feature points from the preprocessed 3D face surface. The N feature points are the center of N patches. Next, for each patch, we have extracted two histogram descriptors to describe both geometrical and visual proprieties of the face region; the HoS is of $d_{hos} = 50$ dimension computed on the 3D mesh as described in Section 4.2, and the $\mathrm{LBP}_{8,1}$ based descriptor ($p = 8$, $r = 1$), which is a 256-dimensional histogram computed as described in Section 4.1.

To build the codebook from both descriptors, we used k-means algorithm for descriptor clustering. We denote C_{hos} and C_{lbp} the codebooks constructed from respectively HoS and LBP descriptors. Both codebook sizes are fixed to be $k_h = k_l = 80$ terms. Once the codebooks are constructed, each face is then described using two term vectors ($V_{hos} = [h_1, h_2 ..., h_{k_h}]$, $V_{lbp} = [l_1, l_2 ..., l_{k_l}]$) which represent the number of times the each term appears in the face (Section 4.3). The concatenated term vector $V = [V_{hos}, V_{lbp}]$ used for classification is of k dimension ($k = k_h + h_l = 160$). It is subsequently normalized (see Figure 2).

Finally, to assess the classification performance of our method on FRGCv2 dataset, we have applied three different classifiers as presented in Section 5. Note that in our experiment, we use neutral gallery faces for the training, while the rest are used for the test.

Table 1 presents the classification performance of the BoF representation using HoS and LBP histograms separately, and the performance of the concatenated signature. We can clearly see that the combination of the BoF of the extracted HoS and LBP his-

Table 1: Classification performance on FRGCv2 dataset using three different classifiers.

Method	Naïve Bayes	Random forests	MLP
Bag-of-HoS	91.2%	93.9%	96.1%
Bag-of-LBP	93.8%	94.8%	95.4%
Bag-of-(HoS+LBP)	95.4%	96.1%	97.9%

Table 2: Comparison with state-of-art methods on the FRGCv2 dataset. The reported results are performed using ”Neutral versus All” identification protocol.

Method	Neutral Versus. All
(Alyüz et al., 2010)	97.5%
(Elaiwat et al., 2014)	94.4%
(Drira et al., 2013)	97.0%
(Huang et al., 2012)	97.6%
(Faltemier et al., 2008)	97.2%
Our method	**97.9%**

tograms gives higher classification performance compared to the separate use of term vectors from single representation. This can be explained by the fact that both descriptors are complementary and thus their concatenation increases the classification performance.

From Table 1, we can also see that BoF of LBP representation performs better compared to the BoF of HoS representation when using Random forests and Naïve Bayes classifiers. After combination, the Multilayer perceptron classifier gives the highest classification rate (97.9%) due to its deep representation, followed by Random forests and Naïve Bayes respectively. The high classification performance obtained by Multilayer perceptron classifier is achieved thanks to its high tolerance to noisy data, as well as its ability to classify patterns on which they have not been trained. Note that the performance of our method can further be improved by combining the three used classifiers using different combination approaches (e.g. linear combination, majority vote, highest confidence vote, etc.).

Table 2 presents a comparison of rank-1 recognition rate of our method to other state-of-art methods using the same protocol. In this experiment, we evaluate ”Neutral versus All” identification experiment. It can be observed that our method achieves the best face identification performance among other state-of-art methods with 97.9%. Particularly, (Elaiwat et al., 2014) used curvelet local features and achieved a rank-1 identification rate of 94.4%. (Drira et al., 2013) used radial curves to represent the face surfaces, they achieved 97.0% rank-1 identification rate. Furthermore, (Huang et al., 2012) used eLBP for 3D facial representation and achieved 97.6% rank-1 identification rate. Besides, (Alyüz et al., 2010) utilized curvature-based 3-D shape descriptors and achieved 97.5% rank-1 identification rate. Finally, (Faltemier et al., 2008) achieved 97.2% rank-1 identification rate using the matching of the best committee of local regions. This comparison shows the efficiency of our proposed method against expression variation.

6.2 Effect of Vocabulary Size

In this experiment, we analyzed how the classification performance varies with respect to the visual vocabulary size i.e. the number of visual words (k_h, k_l) in each codebook. We set $N = 200$ the number of extracted HoS and LBP features, and vary the number of visual words k_h, k_l of each codebook between 60 to 180. In our implementation, we use same values of both k_h and k_l.

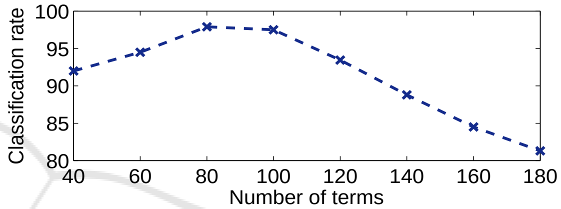

Figure 6: Effect of the codebook size on the classification performance using Multilayer perceptron classifier.

Experimental results displayed in the Figure 6 shows that good results can be obtained even with a relatively small number of visual words. The classification rate remains stable when the number of terms varies between 80 and 100, it starts to drop when choosing values outside this interval, especially when it exceeds 100 terms. Note that this performance depends also on N (the number of the extracted feature points from the face).

7 CONCLUSION

In this paper, we presented an efficient 3D face recognition approach. We firstly extract the Histogram of Shape index and Local binary pattern as region descriptors from 3D mesh and 2D depth image respectively. Next, we build two different visual vocabularies using k-means clustering. A term vector is then computed as the occurrence of each visual word in the face, and a global term vector is constructed by serially concatenating both HoS and LBP term vectors. Our experiments on FRGCv2 dataset show that the proposed method is robust against expression variation and gives challenging results compared to other state-of-art methods.

REFERENCES

Ahonen, T., Hadid, A., and Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(12):2037–2041.

Alyüz, N., Gökberk, B., and Akarun, L. (2010). Regional registration for expression resistant 3-d face recognition. *Information Forensics and Security, IEEE Transactions on*, 5(3):425–440.

Ballihi, L., Ben Amor, B., Daoudi, M., Srivastava, A., and Aboutajdine, D. (2012). Boosting 3-d-geometric features for efficient face recognition and gender classification. *Information Forensics and Security, IEEE Transactions on*, 7(6):1766–1779.

Berretti, S., Amor, B. B., Daoudi, M., and Del Bimbo, A. (2011). 3d facial expression recognition using sift descriptors of automatically detected keypoints. *The Visual Computer*, 27(11):1021–1036.

Besl, P. J. (2012). *Surfaces in range image understanding*. Springer Science & Business Media.

Besl, P. J. and McKay, N. D. (1992). Method for registration of 3-d shapes. In *Robotics-DL tentative*, pages 586–606. International Society for Optics and Photonics.

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.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Chang, K. I., Bowyer, K. W., and Flynn, P. J. (2005). An evaluation of multimodal 2d+ 3d face biometrics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(4):619–624.

Chua, C.-S., Han, F., and Ho, Y.-K. (2000). 3d human face recognition using point signature. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 233–238. IEEE.

Creusot, C., Pears, N., and Austin, J. (2013). A machine-learning approach to keypoint detection and landmarking on 3d meshes. *International journal of computer vision*, 102(1-3):146–179.

Ding, C. and Tao, D. (2015). Robust face recognition via multimodal deep face representation. *arXiv preprint arXiv:1509.00244*.

Drira, H., Ben Amor, B., Srivastava, A., Daoudi, M., and Slama, R. (2013). 3d face recognition under expressions, occlusions, and pose variations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(9):2270–2283.

Elaiwat, S., Bennamoun, M., Boussaid, F., and El-Sallam, A. (2014). 3-d face recognition using curvelet local features. *Signal Processing Letters, IEEE*, 21(2):172–175.

Faltemier, T. C., Bowyer, K. W., and Flynn, P. J. (2008). A region ensemble for 3-d face recognition. *Information Forensics and Security, IEEE Transactions on*, 3(1):62–73.

Fanelli, G., Dantone, M., Gall, J., Fossati, A., and Van Gool, L. (2013). Random forests for real time 3d face analysis. *International Journal of Computer Vision*, 101(3):437–458.

Hariri, W., Tabia, H., Farah, N., Benouareth, A., and Declercq, D. (2016a). 3d face recognition using covariance based descriptors. *Pattern Recognition Letters*, 78:1–7.

Hariri, W., Tabia, H., Farah, N., Declercq, D., and Benouareth, A. (2016b). Hierarchical covariance description for 3d face matching and recognition under expression variation. In *2016 International Conference on 3D Imaging (IC3D)*, pages 1–7. IEEE.

Huang, D., Ardabilian, M., Wang, Y., and Chen, L. (2011a). A novel geometric facial representation based on multi-scale extended local binary patterns. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, pages 1–7. IEEE.

Huang, D., Ardabilian, M., Wang, Y., and Chen, L. (2012). 3-d face recognition using elbp-based facial description and local feature hybrid matching. *Information Forensics and Security, IEEE Transactions on*, 7(5):1551–1565.

Huang, D., Shan, C., Ardabilian, M., Wang, Y., and Chen, L. (2011b). Local binary patterns and its application to facial image analysis: a survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 41(6):765–781.

Jafri, R. and Arabnia, H. R. (2009). A survey of face recognition techniques. *JIPS*, 5(2):41–68.

Jyothi, K. and Prabhakar, C. (2014). Multi modal face recognition using block based curvelet features. *International Journal of Computer Graphics & Animation*, 4(2):21.

Lei, Y., Bennamoun, M., and El-Sallam, A. A. (2013). An efficient 3d face recognition approach based on the fusion of novel local low-level features. *Pattern Recognition*, 46(1):24–37.

Lewis, D. D. (1998). Naive (bayes) at forty: The independence assumption in information retrieval. In *Machine learning: ECML-98*, pages 4–15. Springer.

Luo, J., Geng, S., Xiao, Z., and Xiu, C. (2015). A review of recent advances in 3d face recognition. In *Sixth International Conference on Graphic and Image Processing (ICGIP 2014)*, pages 944303–944303. International Society for Optics and Photonics.

Mishra, B., Fernandes, S. L., Abhishek, K., Alva, A., Shetty, C., Ajila, C. V., Shetty, D., Rao, H., and Shetty, P. (2015). Facial expression recognition using feature based techniques and model based techniques: A survey. In *Electronics and Communication Systems (ICECS), 2015 2nd International Conference on*, pages 589–594. IEEE.

O'Hara, S. and Draper, B. A. (2011). Introduction to the bag of features paradigm for image classification and retrieval. *arXiv preprint arXiv:1101.3354*.

Ojala, T., Pietikäinen, M., and Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987.

Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., Marques, J., Min, J., and Worek, W. (2005). Overview of the face recognition grand challenge. In *Computer vision and pattern recognition, 2005. CVPR 2005. IEEE computer society conference on*, volume 1, pages 947–954. IEEE.

Sandbach, G., Zafeiriou, S., and Pantic, M. (2012). Local normal binary patterns for 3d facial action unit detection. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 1813–1816. IEEE.

Shan, C., Gong, S., and McOwan, P. W. (2009). Facial expression recognition based on local binary patterns: A comprehensive study. *Image and Vision Computing*, 27(6):803–816.

Smeets, D., Keustermans, J., Vandermeulen, D., and Suetens, P. (2013). meshsift: Local surface features for 3d face recognition under expression variations and partial data. *Computer Vision and Image Understanding*, 117(2):158–169.

Sun, Y., Wang, X., and Tang, X. (2014). Deep learning face representation from predicting 10,000 classes. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 1891–1898. IEEE.

Tabia, H., Daoudi, M., Colot, O., and Vandeborre, J.-P. (2012). Three-dimensional object retrieval based on vector quantization of invariant descriptors. *Journal of Electronic Imaging*, 21(2):023011–1.

Tabia, H., Daoudi, M., Vandeborre, J.-P., and Colot, O. (2011). A new 3d-matching method of nonrigid and partially similar models using curve analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4):852–858.

Tabia, H. and Laga, H. (2015). Covariance-based descriptors for efficient 3d shape matching, retrieval, and classification. *Multimedia, IEEE Transactions on*, 17(9):1591–1603.

Tabia, H., Laga, H., Picard, D., and Gosselin, P.-H. (2014). Covariance descriptors for 3d shape matching and retrieval. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4185–4192.

Vretos, N., Nikolaidis, N., and Pitas, I. (2011). 3d facial expression recognition using zernike moments on depth images. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 773–776. IEEE.

Xu, C., Li, S., Tan, T., and Quan, L. (2009). Automatic 3d face recognition from depth and intensity gabor features. *Pattern Recognition*, 42(9):1895–1905.

Yang, B. and Chen, S. (2013). A comparative study on local binary pattern (lbp) based face recognition: Lbp histogram versus lbp image. *Neurocomputing*, 120:365–379.