2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

Effective 3D based Frontalization for Unconstrained Face Recognition

Claudio Ferrari, Giuseppe Lisanti, Stefano Berretti, Alberto Del Bimbo
Media Integration and Communication Center - MICC
University of Florence
Florence, Italy
Email: {claudio.ferrari,giuseppe.lisanti,stefano.berretti,alberto.delbimbo}@unifi.it

***Abstract*—In this paper, we propose a new and effective *frontalization* algorithm for frontal rendering of unconstrained face images, and experiment it for face recognition. Initially, a 3DMM is fit to the image, and an interpolating function maps each pixel inside the face region on the image to the 3D model's. Thus, we can render a frontal view without introducing artifacts in the final image thanks to the exact correspondence between each pixel and the 3D coordinate of the model. The 3D model is then back projected onto the frontalized image allowing us to localize image patches where to extract the feature descriptors, and thus enhancing the alignment between the same descriptor over different images. Our solution outperforms other frontalization techniques in terms of face verification. Results comparable to state-of-the-art on two challenging benchmark datasets are also reported, supporting our claim of effectiveness of the proposed face image representation[1].**

I. Introduction

Among biometric techniques, face recognition has a clear advantage in its non-intrusiveness that allows deployment also in unconstrained scenarios, without user cooperation. This latter capability is one of the main reasons for the success of face recognition, as also evidenced by the increasing demand for surveillance systems that can operate in real contexts, even under strong variations in the face pose, expression, illumination, resolution, etc. In the last few years, an impressive development has been registered in this research area, with results which have substantially closed the gap with the human-level performance in face verification, also thanks to the introduction of the *Deep Neural Net* (DNN) architectures and learning methods [1]. Despite these recent advancements, there are some aspects in the conventional face recognition pipeline (including *detection*, *alignment*, *representation*, *classification*) that require further investigation. In particular, the alignment step is of fundamental importance for the subsequent stages, as for many other face analysis applications such as facial expression recognition [2]. The alignment involves, amongst other things, the compensation for *in-plane* and *out-of-plane* rotations of the head. In most of the cases, this also demands for precise detectors of face landmarks, which is, by itself, a difficult problem, particularly in the presence of face occlusions due to pose variations.

[1]The frontalized images for the LFW dataset can be found at *https://www.micc.unifi.it/resources/datasets/frontalized-faces-in-the-wild/*

Based on the above considerations, in this paper we propose an effective face frontalization approach, which relies on a 3D *morphable model* (3DMM) at its core. Thanks to the features localization capability provided by the constructive properties of the 3DMM, the 2D-3D fitting, and the rendering approach, we will show face recognition results comparable to the state of the art, for an unsupervised scenario, on two benchmark datasets of unconstrained face images.

A. Related Work

In unconstrained face recognition, compensating out of plane rotations is one important issue. Since head rotations occur in the 3D space, pose normalization (also known as face *frontalization*) solutions require that some 3D information of the face is inferred. Methods that address this problem are usually classified as 2D or 3D. In general, effective results have been obtained with methods in both categories, but since pose variations occur in the 3D space, 3D methods are more promising in perspective [3].

2D methods usually cope with the lack of explicit depth information by relying on a training image database, which includes images with different pose (and thus different 3D views). Some 2D transformations (*e.g.*, piecewise affine, thin plate splines) are often used to approximate the 3D transformation, while the error is compensated by some statistical learning strategy. Following this general idea, Berg and Belhumeur [4] presented a face verification built upon a large and diverse collection of linear classifiers that distinguish between two people. Authors propose an identity-preserving alignment procedure based on the detection of 95 face parts that enforces a fairly strict correspondence across images. This alignment procedure uses a reference dataset to distinguish geometry differences due to pose and expression from those that pertain to identity. Ho and Chellappa [5] proposed a method for reconstructing the virtual frontal view of a non-frontal image by using Markov Random Field (MRF), and a variant of the belief propagation algorithm. In this approach, the input face image is divided into a grid of overlapping patches and a set of possible warps for each patch is obtained by aligning it with images from a training database of frontal faces. The problem of finding the optimal warps is then formulated as a discrete labeling problem using MRF. A statistical approach to face frontalization is also proposed by Sagonas et al. [6].

978-1-5090-4846-5/16/$31.00 ©2016 IEEE

The key observation of this work is that, for the facial images lying in a linear space, the rank of a frontal facial image, due to the approximate structure of human face, is much smaller than the rank of facial images in other poses. Based on this, a unified method is proposed for joint face frontalization (pose correction), landmark localization, and pose-invariant face recognition using a small set of frontal images only.

3D methods are based on a 3D face model, either *deformable* or *not-deformable*, used to precisely estimate the 3D face. Zhu et al. [7] presented a 3DMM [8] based pose and expression normalization method to recover the canonical-view, expression-free image, preserving the face appearance with little artifact and information loss. First, a landmark based [9] pose adaptive 3DMM fitting method is proposed, which exploits the "landmark marching" assumption to describe the movement of 3D landmarks across poses. Then, the whole image is meshed into a 3D object, and it is subsequently normalized with 3D transformations. To build a pose robust face recognition system, Yi et al. [10] used a 3DMM, but performing the transformation in the filter space. Differently from the other 3DMM based methods, this solution proposes a "Pose Adaptive Filter" method, which transforms the filters according to the pose and shape of face image retrieved by fitting a 3DMM to the face image, and then uses the pose adapted Gabor filters for feature extraction. Juefei-Xu et al. [11] proposed the *Spartans* framework, which uses a 3D Generic Elastic Model (3D-GEN) to generate virtual face images with various poses for the gallery, and then match the probe to the virtual face images. In particular, the 3D-GEN is used to derive the depth information from a single frontal image per subject of the training set. The high-dimensional Walsh LBP descriptor is uniformly sampled on periocular regions of facial images with robustness toward alignment. During the learning stage, subject-dependent correlation filters are learned for pose-tolerant non-linear subspace modeling in kernel feature space followed by a coupled max-pooling mechanism. Variants of 3D methods use a single, unmodified 3D reference model to estimate a rough approximation of the 3D face surface, and use this surface to generate the new views [1], [12], [13]. Recently, this idea has been followed by Hassner et al. [14]. First, a face is detected, cropped and rescaled to a standard coordinate system. Then, facial feature points are localized [9] in the query image, and used to align it to the feature points on a *reference* face photo. From the 2D coordinates on the query image and their corresponding 3D coordinates on the model, a projection matrix is estimated. An initial frontalized face is obtained by back-projecting the appearance (colors) of the query image to the reference coordinate system using the 3D surface as a proxy. A final result is produced by borrowing appearances from corresponding symmetric sides of the face wherever facial features are poorly visible due to the pose of the query.

B. Our Contribution and Paper Organization

The main contribution of this paper lies in an effective face frontalization approach, experimented in the task of unconstrained face recognition. In particular, we can show that performing a frontal rendering of an unconstrained face image using the proposed technique and a properly constructed 3DMM [16] capable of effectively adapting to faces with varying expression, ethnicity and gender, achieves state of the art results on two benchmark datasets, namely LFW [17] and IJB-A [18], even using baseline descriptors. Experiments revealed that our frontalization technique turned out to be more effective compared to other frontalization algorithms, under the same experimental setting. We also show that using our frontalized version of a face image, we get results in line with recent state-of-the-art methods.

The rest of the paper is organized as follows: In Sect. II, the proposed frontalization along with the 3D head pose estimation and 3DMM fitting techniques is presented. In Sect. III we describe the feature extraction and unsupervised learning process. A comparative evaluation of the proposed approach with respect to other methods for face frontalization is reported in Sect. IV, together with the application of the proposed frontalization to face recognition in comparison with state-of-the-art methods. Finally, discussion and future work are reported in Sect. V.

II. 3D BASED FACE FRONTALIZATION

Our face frontalization grounds on two steps: *(i)* 3D head pose estimation and 3DMM fitting; *(ii)* estimation of the transformation used to back-project the image texture to the 3D model's space and render the frontal image.

A. 3D Head Pose Estimation and Model Fitting

In order to get an estimate of the 3D head pose, we establish a correspondence between a set of facial landmarks detected both in 2D and 3D. We employ the detector defined in [9] to get a set of 2D landmarks, while the same landmarks set is manually annotated once on the vertices of the 3D average model obtained from [16]. The pose that permits us to map each vertex of the 3DMM onto the image is retrieved from the landmark correspondences under and affine camera model as in [15], [16].

As we will expound in detail in Sect. III, we localize the feature descriptors of the face image in correspondence of the projected vertices locations. For this reason, we want the projected vertices to match consistently across all the face images even in presence of non-rigid deformations (*e.g.*, expressions). To this aim, before performing the frontal rendering, we proceed to deform the 3D model in order to fit the face image exploiting the method in [16]. This latter method performs the model deformation minimizing the displacement between the projection of the 3D landmarks on the image plane and their detected position. Points in neighbouring areas will move accordingly by construction of the 3DMM. Assumed that 2D landmarks are detected with sufficient precision, the 3D model is then efficiently fit to the face image.

B. Frontal View Rendering

Once the 3D model is fitted and projected onto the image, a straightforward way to perform an image rendering consists

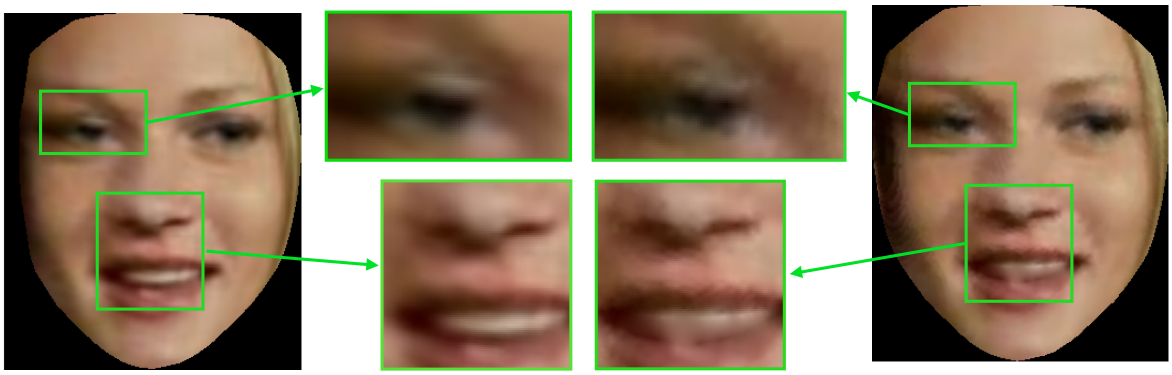

Fig. 1. Difference between our frontalization approach (left) and the one used in [15] (right). The reader can appreciate how the rendering artifacts are removed. Note that the original image size is 250×250, but the face bounding box is approximately 90×90. Better seen in digital form.

in associating to each projected vertex of the model the RGB value of the pixel it falls onto, as in [15]. In this manner, we get a full correspondence between 3D vertices and RGB values. Even though we are now able to build a rendering at arbitrary poses, the original coordinate frame of the 3D model is constructed such that the model faces the z axis; thus, we can easily build a frontal view by just dropping the z value and construct the image by defining a dense regular grid and by putting the RGB values in correspondence of the (x, y) coordinates of the model on the grid. Values corresponding to points in the grid where any vertex falls are interpolated. This approach is easy, but the quality of the rendering is not optimal and many artifacts are introduced in the final image. This can happen for many reasons; for instance, depending on the 3D rotation, some vertices can fall on the same pixel once projected onto the image plane or, on the contrary, pixels can be missed resulting in additional interpolations. The same happens also depending on the scale factor induced by the image resolution; in low resolution images many vertices will be projected onto the same pixel, while the opposite happens in high resolution images.

The proposed frontalization approach overcomes such issues by exploiting the prior knowledge of the face 3D shape. Basically, instead of interpolating the RGB values of pixels where multiple projected vertices fall or are missing, we interpolate the 3D position of each image coordinate inside the region $\mho$ defined by the convex hull of the projected 3D model. This can be practically done since for each vertex i in the 3D shape, we know the 2D position on the image, $P_i = (X_i, Y_i, Z_i) \mapsto (x_i, y_i) = p_i$. We can use these correspondences to fit a surface of the form $P = F(x, y)$, being P_i a vertex in the 3D model and p_i the corresponding projection on the image plane, *i.e.*, a pixel coordinate. We can then evaluate the surface values for each pixel (u, v) inside the face region $\mho$. In doing so, a new 3D shape can be built:

$$\forall (u, v) \in \mho,\ P_{(u,v)} = F(x_{(u,v)}, y_{(u,v)})\ . \tag{1}$$

The resulting 3D model's vertices perfectly fall on each pixel of the image regardless the resolution or the viewpoint. We can now use the new 3D model to sample the RGB values and build the frontalized image in the same way as in [15], but in a clean and more accurate way.

An issue arising here is that out-of-plane rotations will eventually make some points to be self occluded; once projected onto the image, self-occluded points will have (approximately) the same (x, y) coordinates of the ones in the visible side, but a much different z coordinate in 3D. The interpolating function will then estimate ambiguous values and fail. To overcome this problem, first an estimate of the visible 3D vertices given the 3D rotation is obtained, then the surface $F(x, y)$ is computed considering those visible points only. Evidently, the resulting 3D model and, accordingly, the rendered image will have some missing values, substituted by black pixels. Nevertheless, they correspond to points where the information is actually missing. These missing areas, generated by self-occlusions, can be filled with the symmetric visible part. An illustrative example is shown in Fig. 2. It is worth to notice in Fig. 2(c) how the reconstructed 3D model is not uniform, but shows a sort of parametrization imposed by the actual appearance of the particular image. In the example of Fig. 2, the face image undergoes a yaw rotation of $\approx -45°$; even if the density of the 3D model is higher in the visible part of the face, we can notice that it depends also on the orientation of the point's normals with respect to the image plane: the more the projected surface patches normals show orientation parallel to the image plane, the less the surface will be dense.

III. Descriptors of the Face

Usual face recognition approaches perform the identification/verification step by computing image descriptors on the whole sub-image defined by the face bounding box, process them in some way, and feed the resulting early fusion to some classifier. Other than building an accurate frontal rendering, we exploit our frontalization method to precisely localize the image coordinates, where LBP feature descriptors will be computed. As described in Sect. II-B, the rendered image is built based upon the 3D model; In doing so, we can easily back-project the 3D points in the frontal image. Such points define the coordinates in the image where descriptors are going to be extracted, as done in [15]. This has a two

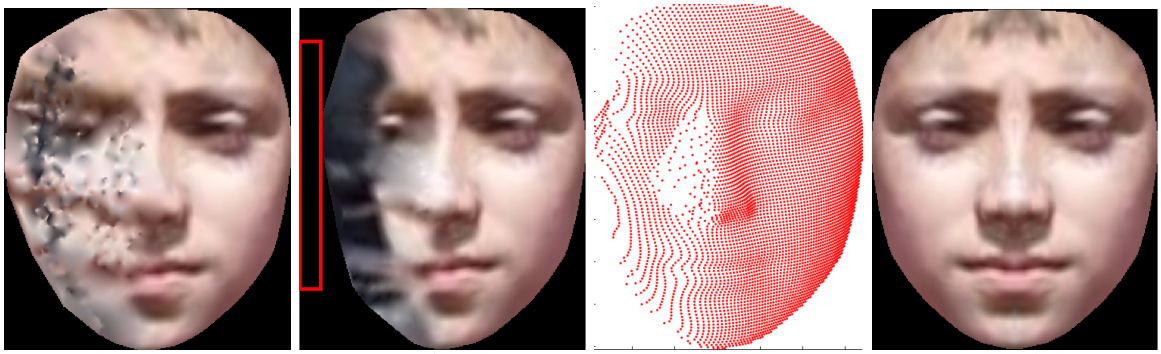

Fig. 2. Frontalization result for heavy out-of-plane rotation ($>$ 40 degrees). Without self occlusion handling (a), many errors due to ambiguities in the 3D coordinates interpolation arise. Excluding the occluded points eliminates the errors and introduces black areas (red rectangle in b). In (c), the 3D model resulting from Eq. (1) is reported. Final result, obtained by symmetrizing the visible part is shown in (d).

fold advantage: (1) We can choose where to extract our descriptors; it is well known that some facial areas are more discriminative than others; (2) Thanks to the 3DMM, we get an intrinsic alignment between single descriptors across the different images. The final face descriptors will always have the same length, regardless of the image size.

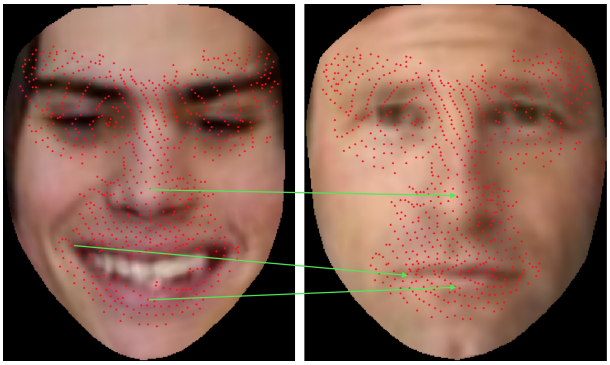

Fig. 3. Localizing the feature descriptors. Our localization strategy permits us to more accurately match descriptors with the same semantic meaning in spite of the location on the image (green arrows).

We chose to extract the feature descriptors selecting a subset of vertices of the 3D model located around the landmarks; we computed for each landmark the k-nearest neighbours (k=370) and selected the intersection of these subsets, resulting finally in 732 vertices. We argued that the most of the information of a human face is retained in such areas (namely, eyes, eyebrows, nose, mouth), and those are the parts less prone to be self occluded. The latter points have been chosen pretty densely so as to generate redundancy based on the result in [19], where it is demonstrated that face recognition benefits from high dimensional feature vectors. Moreover, our 3DMM is capable of moving the vertices to fit the face image, resulting in a more accurate alignment. As shown in Fig. 3, this allows us to precisely match descriptors related to points with the same semantic meaning (see in particular the mouth area). The final face descriptor is built by concatenating the single descriptors. A further dimensionality reduction step is performed by computing PCA on the final descriptors of the training set and applying the projection to the data under test. This allows us to reduce the descriptor from 42456 to 2500 dimensions. The metric used to match the face image descriptors is the *cosine* distance.

IV. Experimental Results

In this section, we report experiments on our approach. First, we describe the protocols and datasets used; we then report the results of the comparison with other face frontalization algorithms. Finally, we report results obtained with our full method, comprising frontalization and feature descriptors localization and compare them with the state-of-the-art.

A. Datasets and Protocols

Tests have been performed on the Labeled Faces in the Wild (LFW) benchmark [17], and on the IARPA Janus Benchmark A (IJB-A) dataset [18]. The LFW dataset represents a challenging benchmark for face verification algorithms including about 13000 face images of 5749 subjects taken under spontaneous conditions, with variabilities in terms of expressions, occlusions and partial pose variations. The recent IJB-A dataset pushes these challenges to the limit including images and videos taken under extreme conditions of illumination, resolution and including full pose variations (i.e., full profiles).

For the LFW dataset, we designed our solution following the *View*-1 protocol defined in [17] and used the *View*-2 protocol to produce our final results. *View*-2 provides 10 sets of 600 image pairs, each set including 300 pairs of the same subject and 300 pairs of different subjects. Ten-fold cross validation is used. We followed the "Unsupervised" protocol and report the results in terms of *Area under the ROC curve* (AUC). More details on the above mentioned protocols can be found in [20].

The IJB-A dataset [18] provides for two types of protocols, namely, *search* and *compare*. The *search* protocol is intended to measure the accuracy of search among N gallery templates, each of which including one or more images of a subject, in terms of the *true acceptance rate* (TAR) at various *false acceptance rates* (FAR). The *compare* protocol, instead, aims at evaluating the verification accuracy between two templates.

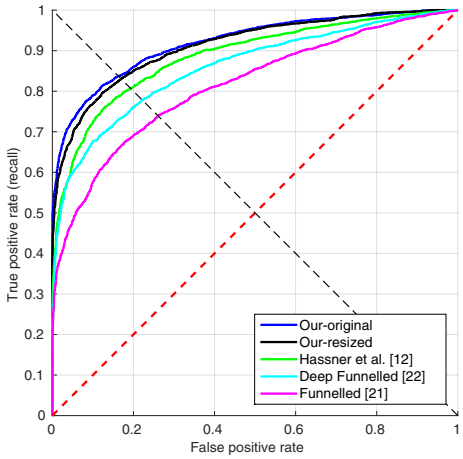

Fig. 4. ROC curves obtained on the LFW dataset using dense sampling on the frontalized face images.

TABLE I
AREA UNDER CURVE (AUC) AND EQUAL ERROR RATE (EER) VALUES OBTAINED ON LFW USING DENSE SAMPLING ON THE FRONTALIZED FACE IMAGES. VALUES ARE IN PERCENTAGE

Method	AUC	EER
Funneled [21]	81.36	26.07
Deep-funneled [22]	85.91	22.20
Hassner et. al. [14]	88.69	19.30
Our-resized	91.28	16.97
Our-original	**92.00**	**16.27**

The metrics used are the TAR corresponding to FAR equal to .1 and .01, and the rank-1 and rank-5 accuracy. The IJB-A contains 10 splits of data. A detailed descriptions of the protocols and metrics for the evaluation can be found in [18].

B. Comparison with Other Frontalization Algorithms

We compared our frontalizations with the ones obtained with the *funneling* [21] and *deep-funneling* [22] algorithms, and with the solution proposed by Hassner et al. [14]. For this experiment, which aims at evaluating the quality of the frontalizations, we considered a slightly different version of our pipeline, identical for all the above mentioned methods; instead of localizing the descriptors exploiting the re-projected 3D model, we densely sampled LBP features on a regular grid with cells of size 10×10 on the whole image. The face images obtained with funnelling, deep-funneling and with the Hassner's technique have size around 100×100 pixels, while our solution generates bigger images ($\approx 200 \times 200$). For a fair comparison, we also report results obtained with a rescaled version of our frontalizations (*Our-resized*), in order to approximately match the size of the others. Results reported in Fig. 4 and Table I show that our method produces a more effective frontal rendering inasmuch as the same verification algorithm is used. Even halving the size of the images does not significantly undermine the performance.

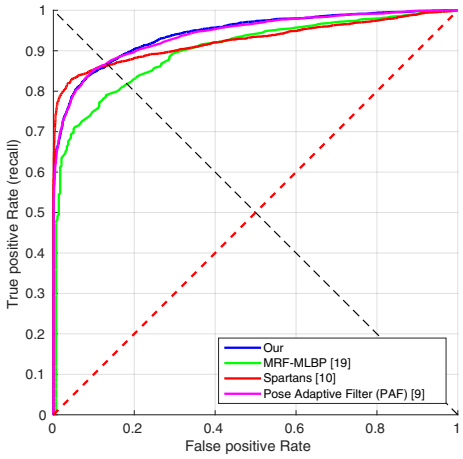

Fig. 5. ROC curves for our method and the state of the art on LFW.

TABLE II
AREA UNDER THE CURVE (AUC) VALUES FOR OUR METHOD AND THE STATE OF THE ART ON LFW. VALUES ARE IN PERCENTAGE

Method	AUC
MRF-MLBP [23]	89.94
Spartans [11]	94.28
Pose Adaptive Filter (PAF) [10]	94.05
MRF-fusion-CSKDA [24]	**98.94**
Our	94.29

TABLE III
RESULTS ON THE IJB-A DATASET

Metric	GOTS	OpenBR [25]	Our
	1:N (Search Protocol)		
TAR@FAR=0.01	$.406 \pm .014$	$.236 \pm .009$	$\mathbf{.609 \pm .015}$
TAR@FAR=0.10	$.627 \pm .012$	$.433 \pm .006$	$\mathbf{.801 \pm .013}$
RANK@1	$.443 \pm .021$	$.246 \pm .011$	$\mathbf{.608 \pm .023}$
RANK@5	$.595 \pm .020$	$.375 \pm .008$	$\mathbf{.767 \pm .014}$

C. Comparison with State of the Art

In the following, we report the results obtained on the LFW dataset and the IJB-A dataset using our full pipeline, and compare them with the state-of-the-art.

We report the results on the LFW dataset obtained following the "Unsupervised" protocol in comparison with the four best performing state of the art techniques[2], namely: MRF-MLBP [23], Spartans [11], MRF-fusion-CSKDA [24] and Pose Adaptive Filter (PAF) [10]. It is possible to appreciate from Fig. 5 and Table II that we obtain comparable performance with respect to the state of the art. It is also worth to note that our method and the PAF technique, both based on fitting a 3DMM, show the same trend in Fig. 5, while the other methods are based on different algorithms.

In Table III, we report the results on the IJB-A dataset in comparison with two baselines: a *government-off-the-shelf* (GOTS) algorithm, and the open source face recognition algorithm OpenBR [25]. We obtain higher performance with respect to the two baselines. However, it is difficult to analyse

[2]We do not report the curve for MRF-Fusion-CSKDA [24] since the relative data are not available.

our performance in comparison to the GOTS algorithm due to missing details about this solution. For OpenBR, instead, despite the similar pipeline, which comprises a step of dimensionality reduction via PCA followed by the application of Linear Discriminant Analysis (LDA), we show largely improved performance. We can argue that our solution generates a frontal rendering of the face image, which is more effective than unprocessed images if applied to recognition.

V. Discussion and Future Work

In this paper, we have proposed: *(i)* An effective algorithm able to generate an artifact-free frontal rendering of a face image based on fitting a 3DMM; *(ii)* An unsupervised face recognition pipeline that relies on the 3D model used to generate the frontalized image and thus enhance the alignment between the feature descriptors. Experimental results obtained on two challenging benchmark datasets support our claim of effectiveness of the proposed approach.

However, the method is not exempt from limitations. First of all, it heavily relies on the accuracy of the landmark detector. Moreover, the 3DMM fitting, besides being lightly affected by the accuracy of the landmark detection as well, is conditioned by the image resolution since it indirectly determines the magnitude of the deformation applied to the 3DMM. Finally, for extreme poses ($>\approx 60^\circ$ in yaw rotation), the method introduces some artifacts in the final image due to a wrong estimation of the projected model's convex hull. Some future developments will regard finding solutions to the latter issues.

Acknowledgment

This paper makes use of the following data made available by the Intelligence Advanced Research Projects Activity (IARPA): IARPA Janus Benchmark A (IJB-A) data detailed at http://www.nist.gov/itl/iad/ig/facechallenges.cfm.

This research is based upon work supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via IARPA contract number 2014-14071600011. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright annotation thereon.

References

[1] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2013, pp. 1701–1708.

[2] F. De la Torre, W.-S. Chu, X. Xiong, F. Vicente, X. Ding, and J. F. Cohn, "Intraface," in *IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 2015, pp. 1–8.

[3] X. Zhao, W. Zhang, G. Evangelopoulos, D. Huang, S. K. Shah, Y. Wang, I. A. Kakadiaris, and L. Chen, "Benchmarking asymmetric 3d-2d face recognition systems," in *IEEE Int. Conf. and Work. on Automatic Face and Gesture Recognition*, 2013, pp. 1–8.

[4] T. Berg and P. N. Belhumeur, "Tom-vs-Pete classifiers and identity-preserving alignment for face verification," in *British Machine Vision Conference*, 2012, pp. 1–11.

[5] H. T. Ho and R. Chellappa, "Pose-invariant face recognition using Markov Random Fields," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 4, pp. 1573–1584, Apr. 2013.

[6] C. Sagonas, Y. Panagakis, S. Zafeiriou, and M. Pantic, "Robust statistical face frontalization," in *IEEE Int. Conf. on Computer Vision*, 2015.

[7] X. Zhu, Z. Lei, J. Yan, D. Yi, and S. Z. Li, "High-fidelity pose and expression normalization for face recognition in the wild," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2015, pp. 787–796.

[8] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," in *ACM Conf. on Computer Graphics and Interactive Techniques*, 1999.

[9] X. Xiong and F. De la Torre, "Supervised descent method and its applications to face alignment," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2013.

[10] D. Yi, Z. Lei, and S. Z. Li, "Towards pose robust face recognition," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2013.

[11] F. Juefei-Xu, K. Luu, and M. Savvides, "Spartans: Single-sample periocular-based alignment-robust recognition technique applied to non-frontal scenarios," *IEEE Trans. on Image Processing*, vol. 24, no. 12, pp. 4780–4795, Dec. 2015.

[12] I. Masi, G. Lisanti, A. D. Bagdanov, P. Pala, and A. D. Bimbo, "Using 3d models to recognize 2d faces in the wild," in *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2013.

[13] T. Hassner, "Viewing real-world faces in 3D," in *Int. Conf. on Computer Vision*, 2013.

[14] T. Hassner, S. Harel, E. Paz, and R. Enbar, "Effective face frontalization in unconstrained images," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2015.

[15] I. Masi, C. Ferrari, A. Del Bimbo, and G. Medioni, "Pose independent face recognition by localizing local binary patterns via deformation components," in *Int. Conf. on Pattern Recognition*, 2014.

[16] C. Ferrari, G. Lisanti, S. Berretti, and A. Del Bimbo, "Dictionary learning based 3D morphable model construction for face recognition with varying expression and pose," in *Int. Conf. on 3D Vision*, 2015.

[17] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Tech. Rep. 07-49, Oct. 2007.

[18] B. F. Klare, B. Klein, E. Taborsky, A. Blanton, J. Cheney, K. Allen, P. Grother, A. Mah, and A. K. Jain, "Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2015.

[19] D. Chen, X. Cao, F. Wen, and J. Sun, "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2013, pp. 3025–3032.

[20] G. B. Huang and E. Learned-Miller, "Labeled faces in the wild: Updates and new reporting procedures," University of Massachusetts, Amherst, Tech. Rep. UM-CS-2014-003, May 2014.

[21] G. B. Huang, V. Jain, and E. Learned-Miller, "Unsupervised joint alignment of complex images," in *IEEE Int. Conf. on Computer Vision*, 2007.

[22] G. B. Huang, M. Mattar, H. Lee, and E. Learned-Miller, "Learning to align from scratch," in *Advances in Neural Information Processing Systems*, 2012.

[23] S. R. Arashloo and J. Kittler, "Efficient processing of MRFs for unconstrained-pose face recognition," in *IEEE Int. Conf. on Biometrics: Theory, Applications and Systems*, 2013.

[24] R. Arashloo, Shervin and J. Kittler, "Class-specific kernel fusion of multiple descriptors for face verification using multiscale binarised statistical image features," *IEEE Trans. on Information Forensics and Security*, vol. 9, no. 12, pp. 2100–2109, Dec. 2014.

[25] J. Klontz, B. Klare, S. Klum, A. Jain, and M. Burge, "Open source biometric recognition," in *IEEE Int. Conf. on Biometrics: Theory, Applications and Systems*, 2013.