

Face Recognition using 3D Morphable Model based alignment

Vudit Vudit Radhakrishna Achanta Sabine Susstrunk
School of Computer and Communication Sciences
Ecole Polytechnique Federale de Lausanne(EPFL), Switzerland

Abstract

Face recognition in real-world images continues to be a challenge. Existing methods either require a large number of training examples or heavy computation. We present a face recognition approach that requires less than ten training samples and is computationally efficient. Our method fits a three-dimensional model using landmarks detected on a small number of face samples captured in uncontrolled environments. This allows us to simultaneously normalize the face pose and align the images. We then use Robust Sparse Coding (RSC) method for face recognition task. Our evaluations on standard dataset shows that our method outperforms the state-of-the-art method.

1. Introduction

Face recognition has always been challenging and researched topic in field of computer vision. It is mainly due to wide areas of its application from bio-metric applications to automatic tagging images on social media. There are upcoming applications in kiosk-style payment system. In order to have such applications work reliably we need to have a robust system. The unconstrained nature of face recognition task makes it difficult to generalise recognition algorithms.

The progress in this area in recent years [26] has shrunk the gap between machine and human level of accuracy. These approach rely on large amount of training data and computation power. Our work focuses on improving recognition with small amount of training data(atmost 10 images) and low computation.

Our approach is evaluated against difficult dataset having huge variability in the intra-class samples. We perform all of our face alignment and recognition task without use of specialised hardware and show improved accuracy rates.

The rest of the paper is organised as follows: Section 2 we review the state-of-the-art face recognition methods,Section 3 we present our approach. We present results in Section 4 and finally conclude in Section 5

2. Previous Works

Handling varying face poses for recognition task is a challenging task. This alone can cause high intra-class variability and affect recognition accuracy. The face image alignment in general helps in recognition. In past several methods have shown to tackle this problem. Based on whether pose normalisation is done in 2D or pose robust methods and 3D we have two categories.

2.1. Methods using general alignment

The pose robust methods use different learnt or hand-made features for face recognition. [24, 9, 19] are unsupervised techniques to learn the descriptors for the face. Recently, Deep Learning techniques[25, 26] have shown to perform well in the recognition task. They are supervised techniques and require huge amount of data to work. There are several other methods which try to approximate 3D transformation in 2D. Active Appearance Models [10] based methods statistically model the shape and grey level appearance of the face in the image. 2D warping [5], Markov Random Field [4] have shown to perform well when the pose changes are not dramatic. In [23] authors use statistical method to get the frontal image of face by knowing that they have minimum rank across different poses.

2.2. Methods using 3D models

3D model based methods generally have higher precision than 2D methods as pose variation can be modeled better using 3D transformations. Recognition methods rely on single-fixed model or morphable model. 3D Morphable Model was presented by Blanz and Vetter [6] by transforming 3D face scans in PCA based vector space representation. Since then different other representation have been developed [22, 8, 3, 21]. These all model incorporate additional information like facial expression, modeling ear regions in-order to make model robust. In [7] fit the 3d model to image by simulation image formation in 3D using computer graphics. They extract shape and texture features. This allows recognition to be pose and illumination invariant. An alternative method is the landmark based 3D face model fitting [15, 2], which estimates the model parameters with the

correspondence between 2D and 3D landmarks. The 3D models can be used to normalise pose and expressions in the 2D images. Hence when employed in that sense they can act as pre-processing step in general face recognition pipeline. Thus, different features like Gabor [20], Local Binary Pattern [1], can be used to describe the image. In [31] use a landmark marching technique to describe movement of 3D landmarks across poses and later generate 3D mesh incorporating this information. Yi *et al.* [30] propose a Pose Adaptive Filter method where face pose and shape is used to modify the Gabor which then extract features. In [16] a 3D Generic Elastic Model (3D-GEN) is used create a set of novel poses from a train image. Later for recognition test image is matched against these generated set of images. A dictionary based approach is used in [11] to create a 3D morphable model. They train images have varied facial poses and expressions hence model is robust to both pose and expressions.

3. Our Approach

We do 3D Morphable Model(3DMM) Fitting of face images and later perform pose normalisation to generate novel front face which are aligned. The aligned images are then used to create dictionary of training examples for RSC algorithm [29].

3.1. Face Landmarks

One of the essential components for fitting 3DMM on the face images are facial landmarks. These are the points present on most of the faces around eyes, nose and mouth region. The face landmarks for images are estimated according to Kazemi *et al.* [17] using Dlib [18]. In total 68 points are detected as shown in figure 2a.

3.2. Face Alignment

In following sections we describe our alignment process. We start by explaining the 3DMM, then present our single and multiple image based fitting. The result of fitting step is 3D model and textured isomap [28] of the face. These two are together are used for pose normalisation. The unseen regions are filled with image inpainting technique [27]. Figure 1 illustrates this pipeline

3.2.1 3D Morphable Model

A 3DMM for face consists of mean shape and albedo model along with set of eigenvectors. The different linear combinations of eigenvectors are responsible for the variation in the 3D model. In our approach we only consider shape and ignore the albedo part. A face shape model can be repre-

sented as:

$$\mathbf{s} = \bar{\mathbf{s}} + \sum_{i=0}^{K-1} \alpha_i \hat{\mathbf{s}}_i \quad (1)$$

where $\bar{\mathbf{s}} \in \mathbb{R}^{3n \times 1}$ is the mean shape, $\hat{\mathbf{s}}_i \in \mathbb{R}^{3n \times 1}$ is the shape eigenvector ($i \in [0, K - 1]$), and α_i are shape coefficients. n is the number of vertices in the model.

In our work we use Surrey Face Model [15] with total $n = 3448$ vertices. There total $K = 63$ shape eigenvectors accounting for 99% of original data variation.

We require a 3D model point for each 2D image landmark in fitting methods described later. We have 50 such correspondences for Surrey Face Model as shown in figure 2b.

3.2.2 Single Image Based 3DMM Fitting

For a given face image \mathbf{I} , we want to fit the 3DMM to it. This problem involves first determining the pose of the face in the image and then estimating the shape coefficients.

We assume an affine camera model and determine the pose of face in image using 2D-3D landmark points correspondence. The camera matrix is estimated using *Gold Standard Algorithm* [14], which minimises the error between 2D points and projected 3D points. The $m = 50$ landmark points discussed in the previous section are used for camera matrix estimation.

We create matrix \mathbf{P} of size $3m \times 4m$ by substituting 3×4 camera matrix on the diagonals. Hence 2D projections of 3D model points can be represented as

$$\mathbf{x}_{\text{proj}} = \mathbf{Ps} \quad (2)$$

$$\mathbf{x}_{\text{proj}} = \mathbf{P}\bar{\mathbf{s}} + \mathbf{P}\hat{\mathbf{s}}\alpha \quad (3)$$

As we consider only landmark points for our approach so different matrices dimensions are, $\mathbf{s}, \bar{\mathbf{s}} \in \mathbb{R}^{4m \times 1}$, $\hat{\mathbf{s}} \in \mathbb{R}^{4m \times K}$, $\alpha \in \mathbb{R}^{K \times 1}$ and $\mathbf{x}_{\text{proj}} \in \mathbb{R}^{3m \times 1}$. The shape coefficients are estimated by minimization of the 3D landmark points projection on the image plane.

$$\ell(\alpha) = \frac{1}{2m} \|\mathbf{x}_{\text{2d}} - \mathbf{x}_{\text{proj}}\|^2 \quad (4)$$

Here we present ridge regression method

$$\ell'(\alpha) = \frac{1}{2m} \|\mathbf{x}_{\text{2d}} - \mathbf{x}_{\text{proj}}\|^2 + \lambda \|\alpha\|^2 \quad (5)$$

$$\ell'(\alpha) = \frac{1}{2m} \|\mathbf{x}_{\text{2d}} - \mathbf{P}\bar{\mathbf{s}} - \mathbf{P}\hat{\mathbf{s}}\alpha\|^2 + \lambda \|\alpha\|^2 \quad (6)$$

$$\nabla \ell'(\alpha) = -\frac{1}{m} (\mathbf{P}\hat{\mathbf{s}})^T (\mathbf{x}_{\text{2d}} - \mathbf{P}\bar{\mathbf{s}} - \mathbf{P}\hat{\mathbf{s}}\alpha) + 2\lambda\alpha \quad (7)$$

where λ is the regularization parameter. Equating the gradient to zero we obtain,

$$[(\mathbf{P}\hat{\mathbf{s}})^T \mathbf{P}\hat{\mathbf{s}} + 2m\lambda\mathbf{I}] \alpha = (\mathbf{P}\hat{\mathbf{s}})^T (\mathbf{x}_{\text{2d}} - \mathbf{P}\bar{\mathbf{s}}) \quad (8)$$

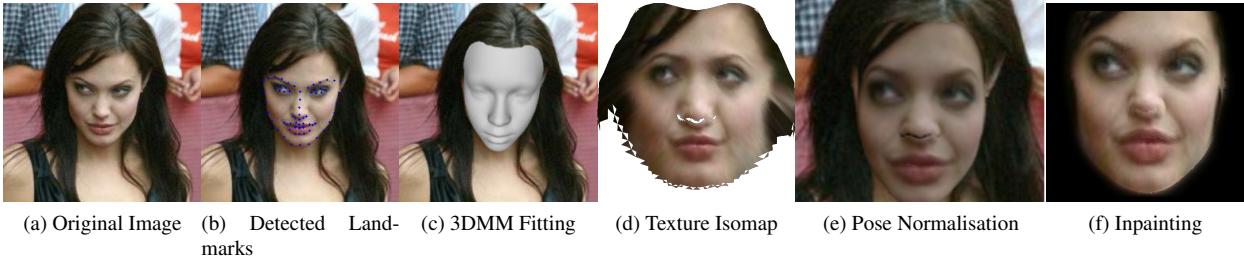
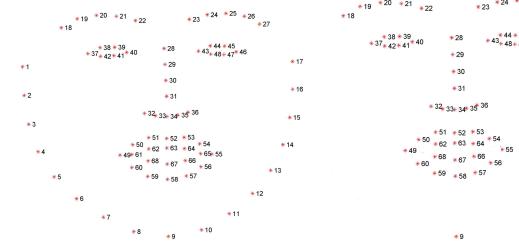


Figure 1: Following sequence of images represent our face alignment processing using 3DMM



(a) 68 landmark points from Dlib [18] (b) 50 landmark points for which we have correspondence in 2D image and Surrey Face Model.

Figure 2: Landmarks Points. Labels are according to Multi-PIE [13] 68 points mark-up

Solving this linear equation 8 we get the our shape coefficients α . The result for the fitting are shown in figure 3 second row. The black region correspond to invisible face region for given face pose.

3.2.3 Multiple Image Based 3DMM Fitting

When we have single with high yaw angle(see figure 6), some part of face is not seen. Hence landmarks detected are improper. We can mitigate this short-coming by using other available images in train set for estimating 3DMM fitting. We can estimate shape coefficient which optimizes the cost function 4 for all the images. For a set of face images $i = 0, 1, \dots, N - 1$ we estimate poses individually and create P_i matrix for each image. The new cost function can be formulated as

$$L(\boldsymbol{\alpha}) = \frac{1}{N} \sum_{i=0}^{N-1} \ell_i(\boldsymbol{\alpha}) + \lambda' \|\boldsymbol{\alpha}\|^2 \quad (9)$$

$$\nabla L(\boldsymbol{\alpha}) = \frac{1}{N} \sum_{i=0}^{N-1} \nabla \ell_i(\boldsymbol{\alpha}) + 2\lambda' \boldsymbol{\alpha} \quad (10)$$

Using the gradient formulation 7 without regularization term and equating 10 to zero we have

$$\left[\sum_{i=0}^{N-1} \left[(\mathbf{P}_i \hat{\mathbf{s}})^T \mathbf{P}_i \hat{\mathbf{s}} \right] + 2Nm\lambda' \mathbf{I} \right] \boldsymbol{\alpha} = \sum_{i=0}^{N-1} (\mathbf{P}_i \hat{\mathbf{s}})^T (\mathbf{x}_{2d}^i - \mathbf{P}_i \bar{\mathbf{s}}) \quad (11)$$

3.2.4 Handling Unseen Landmarks

Due to large pose variations in uncontrolled environment there can be situation when all the face landmarks are not seen. Kazemi *et al.* [17] still estimate all the landmark points irrespective of pose. This will result in improper 3DMM fitting as shape coefficient estimations(equation 8) will be inaccurate.

In our experiments we observed that yaw angle affects the visibility of the landmarks the most. Hence we drop landmarks points affected by large yaw angle in our computation. We don't employ this strategy in single image based fitting as dropping landmarks will lead inaccurate fitting. Whereas, in multiple image scenario we can get unseen landmarks in image from the other image. We drop landmark by replacing the rows corresponding to these points in s, \hat{s} with zeros. Using this we are able to keep our previous equation 8 same but making $(P\hat{s})^T P\hat{s}$ ill-conditioned. The regularisation parameter λ is tuned accordingly to remove ill-conditioned matrix. The chosen value of λ is 3.0.

For yaw angle greater than 35 degree we drop points labeled in figure 2 as 27,46,47,35,36,54,55,56. Corresponding landmarks dropped are on the other side when yaw angle is less than -35. Figure 6 shows the above stated problem.

3.2.5 Textured Isomaps

In most of face recognition algorithms it is important to have different faces to be aligned i.e. same facial point should occur at same coordinates in all the images. Using our 3D model we can achieve that by generating isomaps [28] of the mesh. This representation of 3D mesh preserves the geodesic distance of vertices in 2D plane. We map the 2D image color information to this representation



Figure 3: 3DMM Fitting and Corresponding Isomaps. Top Row: Original Image from LFW dataset. Middle Row: Corresponding Isomaps. White region in the last row images are invisible to given pose in the image. Same region is black in middle row Last Row: .Output of 3DMM Fitting and novel pose are shown.



Figure 4: 3DMM Fitting and Corresponding Isomaps. Top Row: Original Image from LFW dataset. Middle Row: Corresponding Isomaps. White region in the last row images are invisible to given pose in the image. Same region is black in middle row Last Row: .Output of 3DMM Fitting and novel pose are shown.



Figure 5: 3DMM Fitting and Corresponding Isomaps. Top Row: Original Image from LFW dataset. Middle Row: Corresponding Isomaps. White region in the last row images are invisible to given pose in the image. Same region is black in middle row Last Row: .Output of 3DMM Fitting and novel pose are shown.



Figure 6: 68 landmark points are shown for the image. It can be seen that the points near nose and mouth region are not well aligned to true landmark points. The yaw angle for face in image is 41.3 degrees

of the mesh. The implementation is same as by Huber *et al.* [15]. Though they mention that these maps can directly be used for face recognition but few large yaw angles isomaps generated can be highly distorted. Figure 3 last row shows the generated isomaps.

3.3. Pose Normalisation and Inpainting

Using the 3D model and texture map we can generate novel poses for the given face image. As we are interested in face recognition task we frontalise the face in the image.

Using Model View Projection (MVP) matrix we can get to 2D plane from any 3D model. Model matrix transforms model's local coordinates to World coordinates. View matrix from World coordinate to camera and finally Projection matrix brings everything to screen space.

We keep MV matrix as 4x4 identity matrix and Projection as 4x4 ortho camera. Using texture map we can get back the pixel values in 2D plane by standard Opengl rendering method. Final frontalised image is of size 250×250 pixels.

As show in figure 3, there are regions not visible for some face poses. We use inpainting technique [27] to fill in those missing regions. In this method we start from the boundary of the region to be inpainted. A normalised weighted sum of all known neighbourhood pixel values replaces the boundary pixel. The algorithm proceeds by inpainting next nearest unknown pixel. Figure 7 shows the application of pose normalisation and inpainting.

3.4. Face Recognition

We use Robust Sparse Coding(RSC) [29] algorithm for the recognition task. It creates a dictionary matrix D of all training images. For a given test images y it tries to represent it as a linear combination of train images. Additionally, there is a diagonal weight matrix W to count for occlusions and lighting changes. The regularised least-square formulation of RSC is

$$\min_x \|W^{\frac{1}{2}}(y - Dx)\|_2^2 + \lambda\|x\|_2^2 \quad (12)$$

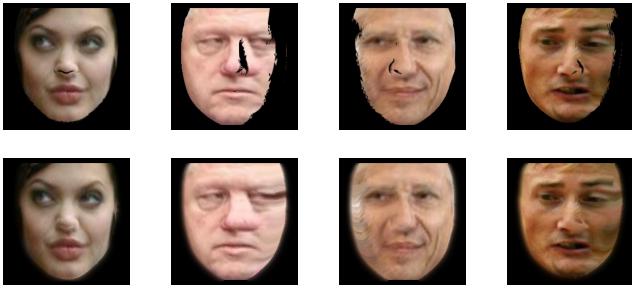


Figure 7: Top row: Front face image generated using isomaps and 3D model. Bottom row: Images after inpainting is done. For the last two images change is not drastic because those regions are again invisible for this pose.

#Classes	#Train Images per Class
538	1
366	2
192	5
140	7
102	10

Table 1: No. of train image per class

where x is the vector of weights.

4. Results

We create train and test set for evaluating our approach. All the images in both sets are frontalised after 3DMM fitting. A dictionary of processed train images is used for RSC algorithm. The baseline for comparison with our approach is Fontaine *et al.* [12]. They have shown improved accuracy rates with ℓ^2 -norm implementation of RSC using very few training examples.

4.1. LFW Database

Labeled Faces in the Wild dataset contains more than 13000 images of 5749 individuals. These images are not aligned or front facing, hence it is necessary to do alignment. We keep 3 images for testing and vary the training set size. Table 1 shows the no. of classes and different train set size used in our experiment.

The frontalised image from our approach is of size 250×250 pixels. The recognition time for image of this resolution is over 2 minutes. Hence we resize image to 50×50 pixels and 100×100 pixels for our experiments. In [12], they use alignment process which crops smaller face region of 30×30 size. We used there cropping method (we call it FontaineCrop proceeding further) on top of our frontalised image. Our approach already aligns image but we wanted to see the effect, as in [12] they gain in performance without

loss much in accuracy. Table 2 summarises different configurations tested. We also compare the multiple image based fitting(M-5,M-6) with single image based fitting.

Methods	Face Alignment Method	Resolution
M-1	Fontaine <i>et al.</i> [12]	30×25
M-2	Our Approach + FontaineCrop	30×25
M-3	Our Approach	50×50
M-4	Our Approach	100×100
M-5	Our MultiView Approach	50×50
M-6	Our MultiView Approach	100×100

Table 2: Different methods configurations

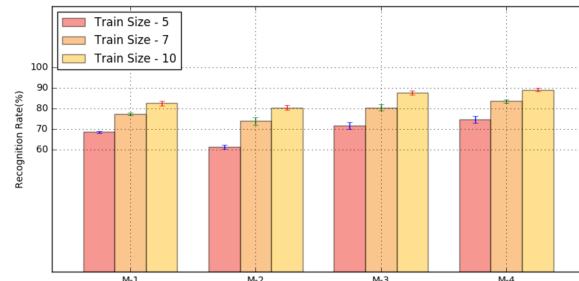


Figure 8: Recognition rate for first four methods listed in table 2 for different train set size.

Figure 8 shows recognition rate for the different experiments on three independent runs. In all the methods we see increase in recognition rate as training examples increases. M-2 performs worse than M-1 probably. The drop can be due to the fact that in [12] they first align each image to reference front face image and fix distance between key landmarks. Later they crop facial region accordingly, which may not be optimal with our approach. M-3 and M-4 perform significantly better than M-1 achieving highest rate of $89.21 \pm 0.71\%$ with 10 training example of 100×100 resolution.

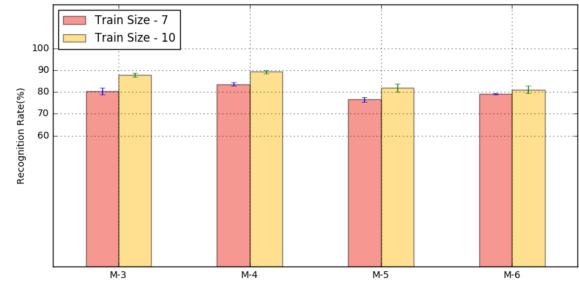


Figure 9: Recognition rate for single image based fitting vs multiple image based fitting for different train set size.

Method	Recognition Rate
Fontaine <i>et al.</i> [12]	$27.52 \pm 0.42\%$
Our Approach 50×50	$31.52 \pm 1.45\%$
Our Approach 100×100	$33.849 \pm 1.05\%$

Table 3: Comparison for train set size = 1

Method	Recognition Rate
Fontaine <i>et al.</i> [12]	$44.83 \pm 0.50\%$
Our Approach 50×50	$54.48 \pm 0.44\%$
Our Approach 100×100	$56.42 \pm 1.27\%$

Table 4: Comparison for train set size = 2

Steps	Time in seconds/image
Pose Estimation	0.036
Shape Coefficient	0.011
Texture Isomap	0.95
Pose Normalisation	
Inpainting	0.14
Face Recognition 50×50	0.5
Face Recognition 100×100	1.53

Table 5: Timings for steps in our approach

As shown in figure 9, we see multiple image based fitting results in poorer performance which is counter-intuitive to our assumptions. It can be due to the fact that dropping of unseen landmark points gives poorer 3DMM fitting.

In case of very small train set size like 1 and 2 train images per class, [12] have shown improved recognition rates. We see that in this case too our method performs better than [12]. Table 3 and 4 show the recognition rates for the two train set sizes. The values are obtain for 3 test images averages over 3 separate runs.

We also look into the timing of our entire pipeline. In [12] authors have shown that their entire pipeline runs under 1s. Our 3DMM fitting task takes roughly 1.30s, in which 80% amount of time is for isomap generation. The recognition part takes around 0.5s for 50×50 image size and 1.53s for 100×100 . There is still scope for optimising isomap generation. Table 5 summarises the time taken by each step.

From the results we can see that our approach provides improvement on the previous best method [12] in similar setting at the cost of increase in the computation time.

5. Conclusion

We present a recognition system which provides improved recognition rate using small number of training samples. We present our results on LFW dataset which has faces taken in uncontrolled environment with high pose variations and lighting conditions. We present face alignment technique using 3D morphable model fitting, which generates pose normalised front faces. We use RSC algorithm for face recognition task.

There is still scope for improvement. We believe that multiple image based 3DMM fitting will provide improved result by better handling of unseen of landmarks. The optimisation of isomap generation will bring down alignment pipeline time, making our approach viable for real-time applications

References

- [1] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 28(12):2037–2041, 2006.
- [2] O. Aldrian and W. A. Smith. Inverse rendering of faces with a 3d morphable model. *IEEE transactions on pattern analysis and machine intelligence*, 35(5):1080–1093, 2013.
- [3] B. Amberg, R. Knothe, and T. Vetter. Expression invariant 3d face recognition with a morphable model. In *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on*, pages 1–6. IEEE, 2008.
- [4] S. R. Arashloo and J. Kittler. Pose-invariant face matching using mrf energy minimization framework. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 56–69. Springer, 2009.
- [5] A. Asthana, M. J. Jones, T. K. Marks, K. H. Tieu, R. Goecke, et al. Pose normalization via learned 2d warping for fully automatic face recognition. In *BMVC*, pages 1–11. Citeseer, 2011.
- [6] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 187–194. ACM Press/Addison-Wesley Publishing Co., 1999.
- [7] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Transactions on pattern analysis and machine intelligence*, 25(9):1063–1074, 2003.
- [8] J. D. Bustard and M. S. Nixon. 3d morphable model construction for robust ear and face recognition. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2582–2589. IEEE, 2010.
- [9] Z. Cao, Q. Yin, X. Tang, and J. Sun. Face recognition with learning-based descriptor. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2707–2714. IEEE, 2010.
- [10] T. F. Cootes, G. J. Edwards, C. J. Taylor, et al. Active appearance models. *IEEE Transactions on pattern analysis and machine intelligence*, 23(6):681–685, 2001.

- [11] 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 *3D Vision (3DV), 2015 International Conference on*, pages 509–517. IEEE, 2015.
- [12] X. Fontaine, R. Achanta, and S. Susstrunk. Face recognition in real-world images. In *International Conference on Acoustics, Speech, and Signal Processing*, 2017.
- [13] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Multi-pie. *Image and Vision Computing*, 28(5):807–813, 2010.
- [14] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [15] P. Huber, G. Hu, R. Tena, P. Mortazavian, W. P. Koppen, W. Christmas, M. Rätsch, and J. Kittler. A multiresolution 3d morphable face model and fitting framework. In *Proceedings of the 11th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2016.
- [16] F. Juefei-Xu, K. Luu, and M. Savvides. Spartans: single-sample periocular-based alignment-robust recognition technique applied to non-frontal scenarios. *IEEE Transactions on Image Processing*, 24(12):4780–4795, 2015.
- [17] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1867–1874, 2014.
- [18] D. E. King. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758, 2009.
- [19] H. Li, G. Hua, Z. Lin, J. Brandt, and J. Yang. Probabilistic elastic matching for pose variant face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3499–3506, 2013.
- [20] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image processing*, 11(4):467–476, 2002.
- [21] A. Patel and W. A. Smith. 3d morphable face models revisited. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1327–1334. IEEE, 2009.
- [22] P. Paysan, R. Knothe, B. Amberg, S. Romdhani, and T. Vetter. A 3d face model for pose and illumination invariant face recognition. In *Advanced video and signal based surveillance, 2009. AVSS’09. Sixth IEEE International Conference on*, pages 296–301. IEEE, 2009.
- [23] C. Sagonas, Y. Panagakis, S. Zafeiriou, and M. Pantic. Robust statistical face frontalization. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3871–3879, 2015.
- [24] K. Simonyan, O. M. Parkhi, A. Vedaldi, and A. Zisserman. Fisher vector faces in the wild. In *BMVC*, volume 2, page 4, 2013.
- [25] Y. Sun, X. Wang, and X. Tang. Deep learning face representation from predicting 10,000 classes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1891–1898, 2014.
- [26] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701–1708, 2014.
- [27] A. Telea. An image inpainting technique based on the fast marching method. *Journal of graphics tools*, 9(1):23–34, 2004.
- [28] J. B. Tenenbaum, V. De Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.
- [29] M. Yang, L. Zhang, J. Yang, and D. Zhang. Robust sparse coding for face recognition. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 625–632. IEEE, 2011.
- [30] D. Yi, Z. Lei, and S. Z. Li. Towards pose robust face recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3539–3545, 2013.
- [31] 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 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 787–796, 2015.