

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/220929432>

# Face Recognition Across Pose with Automatic Estimation of Pose Parameters through AAM-Based Landmarking

Conference Paper · August 2010

DOI: 10.1109/ICPR.2010.332 · Source: DBLP

CITATIONS

8

READS

106

3 authors:



**Lucía Teijeiro-Mosquera**

University of Vigo

7 PUBLICATIONS 185 CITATIONS

[SEE PROFILE](#)



**Jose Luis Alba-Castro**

University of Vigo

121 PUBLICATIONS 1,632 CITATIONS

[SEE PROFILE](#)



**Daniel González Jiménez**

Gradiant - Centro Tecnológico de Telecomunicaciones de Galicia

38 PUBLICATIONS 718 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Sign Language Recognition [View project](#)



ARTFIBIO [View project](#)

# Face recognition across pose with automatic estimation of pose parameters through AAM-based landmarking

Lucía Teijeiro-Mosquera, José Luis Alba-Castro  
*University of Vigo, Spain*  
 {lteijeir, jalba}@gts.tsc.uvigo.es

Daniel González-Jiménez  
*GRADIANT, Spain*  
 dgonzalez@gradient.org

## Abstract

*In this paper we present a fully automatic system for face recognition across pose where no frontal view is needed in enrollment or test. The system uses three Active Appearance Models(AAMs): the first one is a generic multiresolution AAM, while the remaining ones are trained to cope with left/right variations (i.e. pose-dependent AAMs). During the fitting stage, pose is automatically estimated using eigenvector analysis, and a synthetic face is generated through texture warping. Results over CMU PIE Database show promising results compared to the performance achieved with manually landmarked faces.*

## 1. Introduction

One of the pending challenges in face recognition under uncontrolled acquisition conditions still consist of the pose invariance problem. Appearance difference caused by intra-personal face rotation is often larger than inter-personal differences under the same pose. There have been multiple research efforts to deal with face recognition across pose ([15] is a recent review). The straightforward solution to this problem is to store different views for each registered subject. But in many real applications, we can not afford to have more than one image per subject, and even, frequently, the stored face image is not frontal, as in video-surveillance. In these cases the approaches can be divided into those based on fully redesigning the recognizer [2][14][10] in order to match pose-invariant face features, and those that rely on creating virtual views of rotated faces and use a general purpose recognizer for matching the real image and the synthetic one under the same pose. In this last approach, we distinguish between methods based on 3D modeling, [16][1] and methods based on 2D, [7][11][3].

This work builds on our previous approach [7] for creating virtual views from 2D rotated faces by modifying the pose parameters of a trained PDM. The advantage of just changing pose parameters resides in keeping the identity information intact. In this paper we extend the study to include automatic estimation of pose parameters through AAM-based landmarking and design a fully automatic system for face recognition across pose.

The rest of the sections are organized as follows: Active Appearance Models [6] are reviewed in section 2, and our approach to landmark faces is explained. In section 3, we review the method to create virtual views of the faces to be matched. The complete automatic system is explained in section 4. The system is tested in section 5 over the CMU Pie database [12], and some conclusions are drawn in section 6.

## 2. Pose-Dependent Active Appearance Models

Active Appearance Models combine a powerful model of joint shape and texture with a gradient-descent fitting algorithm. AAM was first introduced by Cootes et al. in 1998 [6] and, since then, this modeling has been widely used in face and medical image analysis.

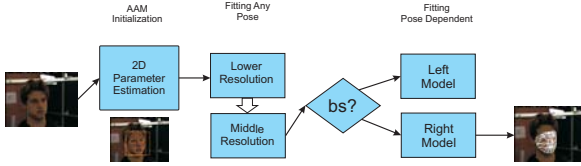
During a training stage, a set of landmarked face-shapes  $s_i = (x_{0i}, y_{0i}, \dots, x_{ni}, y_{ni})$  are aligned using Procrustes analysis, in order to get invariance against 2D rigid changes (scale, translation, roll rotation). A shape model is created through PCA of the aligned training shapes (1). In the same way, textures  $g$  are warped to a frame reference set, and intensity normalized before being combined in a PCA texture model (2). The joint shape-texture model is built by applying a third PCA to the suitably combined shape and texture coefficients. At this point we have a model that can represent one face in a few set of parameters, known as the appearance parameters  $c_i$  (3).

$$s_i = \bar{s} + b_{s_i} P_s \quad (1)$$

$$g_i = \bar{g} + b_{g_i} P_g \quad (2)$$

$$\begin{aligned} s_i &= \bar{s} + c_i Q_s \\ g_i &= \bar{g} + c_i Q_g \end{aligned} \quad (3)$$

During the fitting stage, starting from a reasonable good initialization, the AAM algorithm iteratively corrects the appearance parameters, using a gradient descent algorithm that tries to minimize the squared error between the texture of the face normalized and warped to the reference frame,  $g_s$ , and the texture of the face reconstructed by the model,  $g_m$ , using the appearance parameters. This way, given a target face, with an unknown shape, we can figure out the shape of the face by fitting it to the model and recovering the shape from the appearance parameters. One of the drawbacks of the AAM is that the performance of the fitting algorithm decreases if the model has to explain high variations, like yaw and pitch rotations of the face. In order to handle these situations Cootes et al. [5] trained different AAM models to fit different small ranges in pose. The *view-based* AAM can cope with large pose changes but at the price of making the searching of the best-fit model more complicated and sometimes unreliable.



**Figure 1. Landmarking Process using the Pose-Dependent AAM.**

In this work, we change the searching for the best-fitting model. First we use a multiresolution AAM [4] that includes both left and right rotation to have a coarse approximation of the face shape, and hence, a coarse approximation of the shape parameters. Before the higher level of resolution, we decide which is the best model for our image, based on the value of the shape PCA parameter linked to the pose eigenvector. In the following section we justify the selection of this parameter. The higher resolution of the generic multiresolution model is not used for fitting, being replaced by the pose dependent model.

The advantage of our approach is that we save the extra-cost of landmarking the face with several pose-dependent models. Also, estimating the pose using the shape parameters instead of the appearance parameters,

makes our decision more robust to variable lighting conditions.

### 3. Face Frontalization

Our goal now is to synthesize frontal views of the face (gallery and probe) only knowing one rotated view of the face. González-Jiménez et al. [7] state that, as rigid movement constitute the largest variation in shape, this information will be decoupled and concentrated in the first few eigenvectors of the PCA model of 2D shapes. As Procrustes analysis is performed prior to PCA, only rigid variations, like pitch and yaw rotations (not eliminated with a 2D alignment) are reflected in the first shape eigenvectors. In our case, we use a PCA model of both frontal and yaw rotated faces from CMU Pie database[12], therefore, all the yaw rotation information concentrate in the first shape parameter, the so-called pose-parameter. This pose-parameter has, approximately, a linear variation with the rotation angle  $\theta$  between  $-45^\circ$  and  $+45^\circ$  ( $b_s \simeq (b_{s(pose)} \simeq K\theta, b_{s(oth)})$ ), and therefore, the rotation angle can be estimated from any landmarked face, just by projecting the aligned landmarks into the shape model. We are going to use the higher resolution level of the generic model with this objective.

Once we have a face represented in our shape model  $b_{s_i} = (s_i - \bar{s})P_s^T$ , the frontalization process consist of setting to zero the rotation parameter  $b_{s_i}^{Fr} = (b_{s_i(pose)} = 0, b_{s_i(oth)})$ , and reconstruct the shape using the frontalized parameters  $s_i^{Fr} = \bar{s} + b_{s_i}^{Fr} P_s$ . The frontalized shape is filled with the texture from the image. In cases where the rotation is large and self-occlusions appear, symmetry is applied, and the texture of the visible half side of the face is used to fill both left and right half side of the frontalized shape.

### 4. Automatic face recognition across pose

In this section, we explain the whole process of face recognition robust to pose changes using the CMU PIE database as experimental setup. We split the 68 subjects on the CMU-Pie database into training and test subjects. We use the training subjects to build 3 different models, one generic multiresolution model (3 levels of resolution) including all poses, as explained before, one standard AAM model for the left rotated faces and another one for the right rotated faces. The shape model of the third level of resolution is used to frontalize the shape. Both training and testing images are preprocessed using [9], in order to cope with illumination variations. The

AAM in the lower level of resolution is initialized using an estimation of the scale and position of the face. This estimation is performed using four Viola-Jones detectors (face, eyes, mouth and nose). After fitting in two levels of resolution, the pose parameter is checked to decide which is the pose of the face and, therefore, which one of the pose dependent models needs to be applied for the final fitting (see figure 1). Then, we project the image into the higher resolution AAM to recover a better estimation of the shape parameters. Setting to zero the pose parameter, and reconstructing the shape, we get a frontalized shape with minimum identity lost. The frontalized virtual images are generated by warping the free-shape texture of the input image to the frontalized shape. As we have projected the face into the higher resolution AAM, we actually have two free-shape textures:  $g_s$  the texture from the original image normalized and warped to reference frame, and the texture of the face reconstructed by the model,  $g_m$ , using the appearance parameters. Also, we could have frontalized the shape by modifying only the pose parameter (our main approach) or taking directly the mean shape of the model with no identity information. We will perform recognition experiments with these four approaches to get virtual frontal faces:

- FoT:  $g_s$  warped to the frontalized shape by setting pose parameter to zero
- MoT:  $g_s$  warped to the mean shape of the model.
- FmT:  $g_m$  warped to the frontalized shape by setting pose parameter to zero
- MmT:  $g_m$  warped to the mean shape of the model.

In the experiments section, we will show the improvement in recognition when keeping the shape identity information and the original texture (FoT). In those cases that the rotation angle of the face is large, we need to apply symmetry to avoid artifacts from self occlusion. As we said in section 3, there is an approximate linear correspondence between the shape parameter and the rotation angle, so we decide to symmetrize or not according to the value of pose parameter.

With this technique we frontalize both the gallery and the probe image and perform recognition with standard frontal face matchers independently on the pose of both gallery and probe poses. For our experiments we use the recognition system described in [8] based on Gabor filters. Figure 2 shows the whole recognition process.

## 5. Experiments

We test our system on 34 subjects from CMU Pie database. For each subject, we use five different views between  $-45^\circ$  and  $+45^\circ$ , each view is used as gallery and probe. Both gallery and probe images are automatically landmarked using the system explained before.

In Table 1, we show pose-independent average results using the four approaches to synthesize virtual frontal views defined before. In this Table, we show that keeping the identity information, using each person frontalized shape instead of the mean shape, improves the recognition results, both if we use the original texture from the image (FoT better than MoT), or the texture recovered from the model (FmT better than MmT). Recognition results get worse using the automatic landmarked points, nevertheless, keeping the identity information still leads us to better recognition results than using the mean shape. It is important to highlight that due to the use of automatic landmarks the frontalization of shape through pose parameter estimation can be more noisy than using the mean shape, always constant. This explain the minor performance decrease on MoT.

When we use the model texture (FmT,MmT) instead of the original texture from the image (FoT,MoT), recognition results decrease because our test subjects are not included in the model and, consequently, the representation of test subjects texture is less accurate.

Tables 2 and 3 show the recognition results using only the frontalization option FoT, both for manually landmarked faces (72 points) and for fully automatic landmarking and pose estimation. The results in recognition decrease for most of the probe-gallery pose combinations, having more dramatic decrement for pose differences of  $45^\circ$  due to the propagation of larger landmarking errors in these poses.

## 6. Conclusions

In this paper we proposed a fully automatic face recognition system able to handle pose variation between probe and gallery images. We showed the importance of keeping the shape identity information through the proper handle of shape subspace coefficients, and

**Figure 2. Recognition Diagram.**

we demonstrate how the method proposed in [7] can be made more robust to automatic landmarking and automatic pose detection.

	FoT	MoT	FmT	MmT	No Pose Correction
Manual Points	98,68	91,32	62,65	55,88	59,56
Automatic Points	96,18	94,70	73,38	73,82	

**Table 1. Average recognition results.**

Probe Pose Gallery Pose	-45°	-22,5°	0°	+22,5°	+45°	Mean
-45°	-	100	100	100	91,18	97,79
-22,5°	100	-	100	100	97,06	99,26
0°	100	100	-	100	100	100
+22,5°	100	100	100	-	100	100
+45°	91,18	94,12	100	100	-	96,68
Mean	97,79	98,53	100	100	97,06	<b>98,68</b>

**Table 2. Results using manual landmarks.**

Probe Pose Gallery Pose	-45°	-22,5°	0°	+22,5°	+45°	Mean
-45°	-	97,06	97,06	94,12	91,18	94,85
-22,5°	94,12	-	100	100	85,29	94,85
0°	94,12	100	-	100	100	98,53
+22,5°	91,18	100	100	-	100	97,85
+45°	88,23	94,12	100	97,06	-	94,85
Mean	91,91	97,79	99,26	97,79	94,12	<b>96,18</b>

**Table 3. Results automatic landmarks.**

Our results using manual landmarks are state of the art [16, 2, 14, 3]. The performance of the fully automatic system decreases due to the propagation of the landmarking error. Our future work will be focus on improving the landmarking through a better initialization of the generic multiresolution model. This will also let us increase the number of pose-dependent AAM, having each of them less variation to handle.

## 7. Acknowledgments

We would like to thank B. Sankur for his fruitful comments on this work, M. B. Stegmann for the availability of his AAM code [13], E. Iglesias-Agulla and E. Argones-Rúa for the implementation of the feature detectors and the frontal face matcher. This work has been supported by the Spanish Ministry of Science and Innovation through project TEC2008-05894.

## References

[1] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 187–194, 1999.

[2] C. Castillo and D. Jacobs. Using stereo matching for 2-d face recognition across pose. *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1–8, June 2007.

[3] X. Chai, S. Shan, X. Chen, and W. Gao. Locally linear regression for pose-invariant face recognition. *Image Processing, IEEE Transactions on*, 16(7):1716–1725, July 2007.

[4] T. Cootes, C. Taylor, and A. Lanitis. Multi-resolution search with active shape models. *Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision and Image Processing., Proceedings of the 12th IAPR International Conference on*, 1:610–612 vol.1, Oct 1994.

[5] T. Cootes, K. Walker, and C. Taylor. View-based active appearance models. *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 227–232, 2000.

[6] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *Lecture Notes in Computer Science*, 1407:484–??, 1998.

[7] D. Gonzalez-Jimenez and J. Alba-Castro. Toward pose-invariant 2-d face recognition through point distribution models and facial symmetry. *Information Forensics and Security, IEEE Transactions on*, 2(3):413–429, Sept. 2007.

[8] D. González-Jiménez and J. L. Alba-Castro. Shape-driven gabor jets for face description and authentication. *IEEE Transactions on Information Forensics and Security*, 2(4):769–780, 2007.

[9] R. Gross and V. Brajovic. An image preprocessing algorithm for illumination invariant face recognition. *4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, June 2003.

[10] R. Gross, I. Matthews, and S. Baker. Appearance-based face recognition and light-fields. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(4):449–465, April 2004.

[11] T. Shan, B. Lovell, and S. Chen. Face recognition robust to head pose from one sample image. *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 1:515–518, 0-0 2006.

[12] T. Sim, S. Baker, and M. Bsat. The cmu pose, illumination, and expression (pie) database of human faces. Technical report.

[13] M. Stegmann, B. Ersboll, and R. Larsen. Fame-a flexible appearance modeling environment. *Medical Imaging, IEEE Transactions on*, 22(10):1319–1331, Oct. 2003.

[14] Z. Wang, X. Ding, and C. Fang. Pose adaptive lda based face recognition. *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4, Dec. 2008.

[15] X. Zhang and Y. Gao. Face recognition across pose: A review. *Pattern Recognition*, 42(11):2876 – 2896, 2009.

[16] X. Zhang, Y. Gao, and M. Leung. Recognizing rotated faces from frontal and side views: An approach toward effective use of mugshot databases. *Information Forensics and Security, IEEE Transactions on*, 3(4):684–697, Dec. 2008.