DISTANCE MAPS: A ROBUST ILLUMINATION PREPROCESSING FOR ACTIVE APPEARANCE MODELS

Sylvain Le Gallou*, Gaspard Breton*, Christophe Garcia*, Renaud Séguier**

** France Telecom R&D - TECH/IRIS*
4 rue du clos courtel, BP 91226, 35 512 Cesson Sévigné
sylvain.legallou@francetelecom.com, gaspard.breton@francetelecom.com, christophe.garcia@francetelecom.com

*** Supélec - IETR, SCEE Team*
Avenue de la boulaie, BP 81127, 35 511 Cesson Sévigné
renaud.seguier@supelec.fr

Keywords: Illumination, lighting robustness, AAM, deformable models, face analysis.

Abstract: Methods of deformable appearance models are useful for realistically modelling shapes and textures of visual objects for reconstruction. A first application can be the fine analysis of face gestures and expressions from videos, as deformable appearance models make it possible to automatically and robustly locate several points of interest in face images. That opens development prospects of technologies in many applications like video coding of faces for videophony, animation of synthetic faces, word visual recognition, expressions and emotions analysis, tracking and recognition of faces. However, these methods are not very robust to variations in the illumination conditions, which are expectable in non constrained conditions.

This article describes a robust preprocessing method designed to enhance the performances of deformable models methods in the case of lighting variations. The proposed preprocessing is applied to the Active Appearance Models (AAM). More precisely, the contribution consists in replacing texture images (pixels) by distance maps as input of the deformable appearance models methods. The distance maps are images containing information about the distance between edges in the original object images, which enhance the robustness of the AAMs models against lighting variations.

1 INTRODUCTION

Due to illumination changes which can considerably modify the texture of images, the Active Appearance Models method (AAM) (Cootes et al., 1998) is not very robust in general situations, but only under constrained lighting conditions. On the other hand, AAMs are very impressive in precisely locating several points of interest of an object. This method allows the automatic and robust localization of several points of interest in facial images, which opens development prospects of technologies in many applications like video coding of faces for videophony, animation of synthetic faces, word visual recognition, expressions and emotions analysis, tracking and recognition of faces.

Many methods have been proposed to overcome the problem of illumination variations. Some works use wavelet-based methods, such as the Active Wavelet Networks for Face Alignment (Hu et al., 2003), which proposes to replace the AAM texture by a wavelet network representation. Other works rely on edge-based approaches, or patch filtering, in which illumination component is removed thanks to lighting models. Huang et al. (Huang et al., 2004) compares these two approaches for Active Shape Models (ASM) (Cootes et al., 1995).

In this paper, we propose a novel low-cost method designed to enhance the performances of deformable models methods in the case of lighting variations, and apply it to the Active Appearance Models (AAM). The contribution consists in replacing texture images (pixels) as input to the AAM method by distance maps. These distance maps are images, containing information about the distances between the edges of objects in the original textured images.

This paper is organized as follows. Section 2 briefly presents the Active Appearance Models (AAM) method. Section 3 describes our method of distance map creation used as AAM method preprocessing. In Section 4, we present our experimental results, the images used to apply the AAM method, the AAM model creation and the comparisons between AAM applications with and without the distance maps preprocessing. Finally, Section 5 concludes the paper with final remarks.

2 AAM: ACTIVE APPEARANCE MODEL

The Active Appearance Model method is a deformable model method which allows shapes and textures to be conjointly synthesized. AAMs, proposed by Edwards, Cootes and Taylor in 1998, are based on a priori knowledge of shapes (points of interests connected to each other) and shape-free textures of a training database. AAMs can thus be used to generate a set of plausible representations of shapes and textures of the learned objects. They also allow the search for objects in images by jointly using shape and texture information. This research is performed by an optimization process on model parameters, in order to match the model as well as possible on the image zone containing the object. This method proceeds in three steps (briefly explained):

- A training phase in which the model and his deformation parameters are created.
  A Principal Component Analysis (PCA) on a shape training base and a PCA on a shape-free texture training base are applied respectively in order to create the statistical shape and texture models given by the formulas:
  $$x_i = x_{moy} + \Phi_x * b_x \quad (1)$$
  $$g_i = g_{moy} + \Phi_g * b_g \quad (2)$$
  with x_i and g_i are respectively the synthesized shape and texture, x_{moy} and g_{moy} the mean shape and the mean texture, Φ_x and Φ_g the matrices of eigenvectors of shape and texture covariance matrices and b_x and b_g the controlling vectors of the synthesized shape and texture.
  Another PCA is then applied on several examples of b which is the concatenation of b_x and b_g in order to obtain the appearance parameter c:
  $$b = \Phi * c \quad (3)$$
  with Φ the matrix of PCA eigenvectors. c is a vector controlling b_x and b_g (equation 3) at the same time, that is to say the shape (equation 1) and texture (equation 2) of the model.
- An experience matrix creation phase in which a relation between the variations of the model control parameter (c) and the adjustments of the model in images is created thanks to several experiences.
  Indeed, each image from the training base contains a synthesized object by a value of the parameter c. Let us note c_0 the value of c in the image i of the training base. By modifying the parameter c_0 by δc ($c = c_0 + \delta c$), we synthesize a new shape x_m and a new texture g_m (equation 3). Let us consider now the texture g_i of the original image i which is inside the shape x_m. The difference of pixels $\delta g = g_i - g_m$ and a linear regression with multiple variables on a certain number of experiments (modification of the training base images by δc), will give us a relation between δc and δg:
  $$\delta c = R_c * \delta g \quad (4)$$
  R_c is called experiment matrix.
- A searching phase which allows the model to be adjusted on objects in new images (using the relation found in Experience matrix creation phase).
  This phase is used to search for a particular texture and shape in new images. The modifications of the appearance parameter c from equation 4 allow the model on the searched object to be adjusted in new images. The algorithm of object search in a new image is as follows:
  1- Generate g_m and x from the c parameters (initially set to 0).
  2- Calculate g_i, the texture of the image in which is the searched object, which is inside x shape.
  3- Evaluate $\delta g_0 = g_i - g_m$ and $E_0 = |\delta g_0|$.
  4- Predict $\delta c_0 = R_c * \delta g_0$.
  5- Find the 1st attenuation coefficient k (among [1.5, 0.5, 0.25, 0.125, 0.0625]) giving $E_j < E_0$, with $E_j = |\delta g_j| = |g_{ij} - g_{mj}|$, and g_{mj} is the texture given by $c_j = c - k * \delta c_0$ and g_{ij} is the texture of the image which is inside x_{ij} (shape given by c_j).
  6- While error E_j is not stable, restart at stage 1 with $c = c_j$.

When convergence of the third phase is reached, representations of texture and shape of the searched object are respectively synthesized through the model in g_m and x. Figure 1 gives an example of a face search with the AAM method.

3 A NEW PREPROCESSING: DISTANCE MAPS

In the proposed approach, we consider distance relations between different edges of a searched texture. We do not directly consider colour or grey levels in the original image, so that the approach is more robust against illumination changes.

The preprocessing of AAMs that we present here is the transformation of the original images into distance maps. Distance map creation associated with an original texture image (Figure 2-A) is obtained in 4 steps as follows.

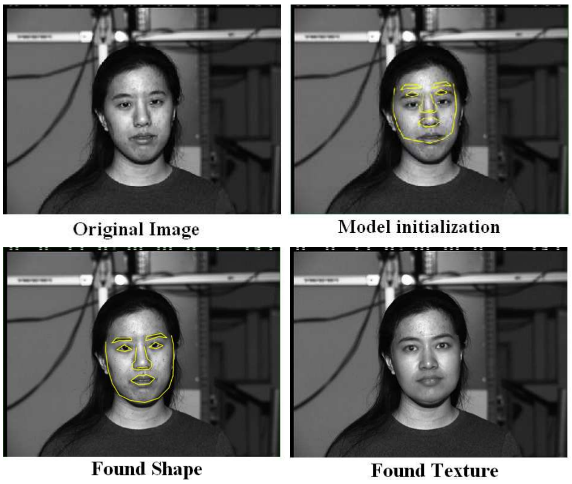

Figure 1: Example of a face search with the AAM method. In the model initialization image, only the mean shape is displayed, not the mean texture.

- The original texture image is divided into a grid of rectangular regions in which histogram equalization is performed. The adaptive histogram equalization will enhance edges contrast. We have implemented the Contrast Limited Adaptive Histogram Equalization (CLAHE) method (Zuiderveld, 1994). More precisely, images are divided into 8x8 contextual regions (i.e. 64 contextual regions in one image), and in each region we applied the CLAHE method according to the Rayleigh distribution (Figure 2-B).

- A smoothed image is obtained by applying a low-pass filter (Figure 2-C).

- Edge extraction is performed in the smoothed image blocks (in a grid of rectangular regions of the smoothed image). This adaptive edge extraction allows edge filtering threshold to be adapted to the local context of the image. The adaptive edge extraction is performed by a sobel filter applied both in x and y axes, in the same 8x8 contextual regions as in the first step. This step produces the edge image (Figure 2-D).

- Finally, for each pixel of the edge image, the Euclidean distance from this pixel to the nearest edge pixel is computed. This last step gives the Distance map (Figure 2-E), associated with an original texture image (Figure 2-A), which is a texture that can be used by AAMs.

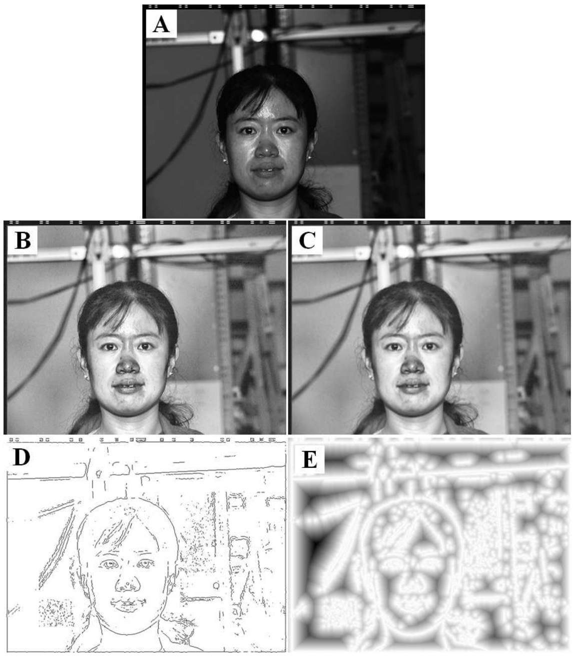

Figure 2: Example of a Distance map creation : A - The original image, B - Image after an adaptive histogram equalization, C - Smoothed image, D - Edge image, E - The distance map.

4 EXPERIMENTAL SYSTEM

In order to make a comparison between results of AAMs applied with and without the preprocessing, we have implemented the experimental system described in this section.

4.1 Images Database

The images used for our tests come from the CMU Pose, Illumination, and Expression (PIE) Database (Sim et al., 2002). It contains facial images of 68 people. Each person is recorded under 21 different illuminations created by a ”flash system” laid out from the left to right of faces.
Figure 3 illustrates the CMU acquisition system, with positions of the 21 flashes and the camera used for creating facial images.

In order to applied the AAM method on this database, we use the AAM reference software made available on line gracefully by T. Cootes on this web site: http://www.isbe.man.ac.uk/~bim/software (Cootes, 2005). We selected 8 persons from the 68 persons of the PIE database. 4 faces will be used in the training phase of AAM method (the 4 top faces in Figure 4) and the 4 remaining faces will be used in the searching phase of AAM method (the 4 bottom faces in Figure 4). The 4 faces used in the

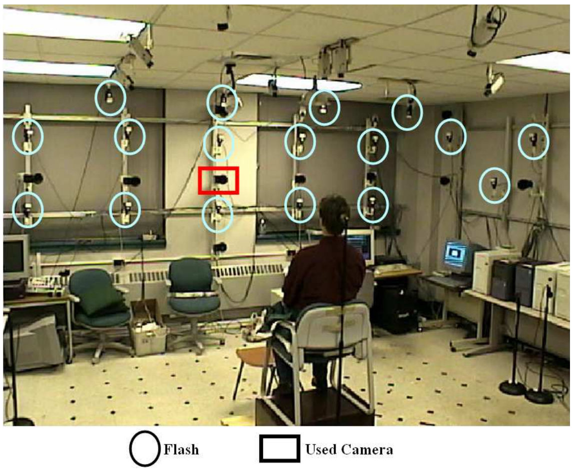

Figure 3: The CMU system of acquisition: positions of 17 of 21 flashes (4 left flashes are not visible in this view) and the camera.

training phase of AAM method have a specific illumination : a full-frontal lighting (illumination number 11). The 20 remaining illuminations on these 4 faces will be used in the searching phase of AAM method with the 4 remaining faces and their 21 illuminations.

We compare the standard AAM method with these 168 original images, which will be called ”Standard experience”, to the standard AAM method with the distance maps preprocessing, which will be called ”Distance experience”. The ”Distance experience” is the standard AAM method applied to the 168 distance maps (associated with the 168 original images) instead of the 168 original images.
Figure 5 shows the training base: the 4 top images are original texture images used in the ”Standard experience” and the 4 bottom images are the corresponding distance maps (associated with the 4 top images) used in the ”Distance experience”.

4.2 Results

In figure 6, one can see a result of the ”Standard experience” (on the left) and a result of the ”Distance experience” (on the right), both obtained for an unknown face. On the top, the shape is displayed and on the bottom, the texture is displayed. In order to have a better representation, in the next figures, shapes found in the ”Distance experience” will be overlaid on the original textures. It should be noted that in all experiments, initialization of the AAM is manually performed and is identical for the ”Standard experience” and the ”Distance experience”.

Figure 7 presents some obtained results (results of the ”Standard experience” in the left column and re-

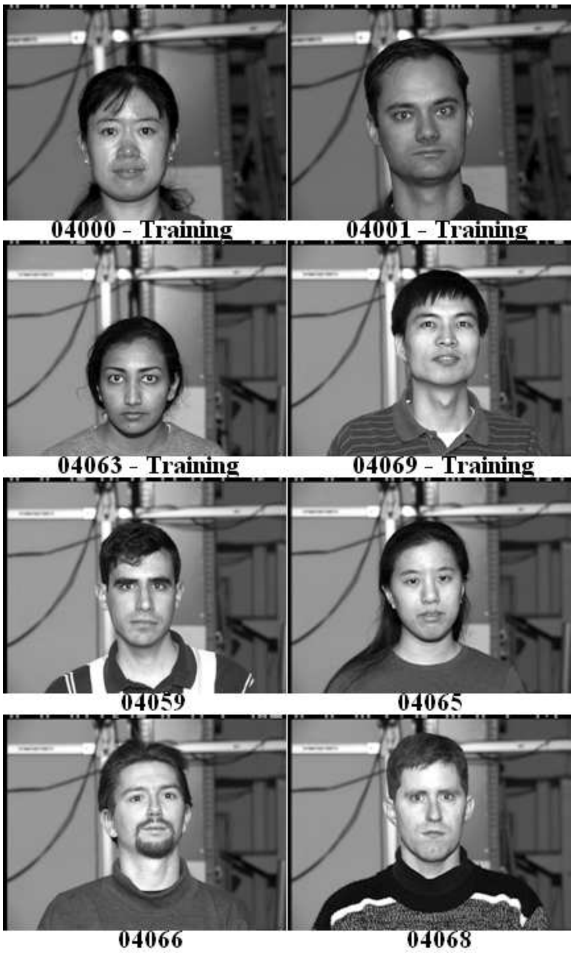

Figure 4: The 8 faces used (4, on the top, for training and searching phases of AAM, 4, on the bottom, for searching phase of AAM).

sults of the ”Distance experience” in the right column) for 4 unknown different faces under 4 different illuminations (from a right strong illumination: images on the top, to a left strong illumination: images on the bottom).

Figure 8 shows error curves obtained in the ”Standard experience” (square curve) and in the ”Distance experience” (round curve). Illuminations from numbers 2 to 8 are lightings more or less strong and more or less high on the left of faces, illuminations from 16 to 22 are lightings more or less strong and more or less high on the right of faces and illuminations from 9 to 15 are lightings in front of faces. Errors are expressed as a percentage distance between the eyes by point, i.e. an error of 1 corresponds to an error made in each point of the model equal to the distance between the eyes. Each curve point in figure 8 is the mean error made by the model in the 8 different face images un-

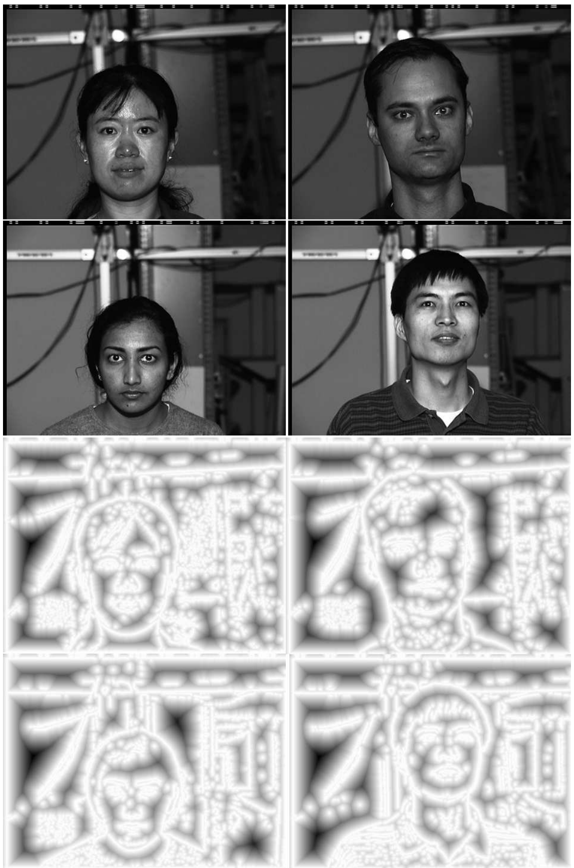

Figure 5: The 4 images used in the ”Standard experience” (on the top) and the 4 distance maps used in the ”Distance experience” (on the bottom).

der the same illumination. This error curve depicts the robustness of the preprocessing used for the distance maps since it makes it possible to find facial features knowing that only 4 face images with frontal lighting were learned by the model. We can see that with preprocessing of the distance maps, some errors are still made, but they are not as strong as with the standard method where errors are made upon facial features searching when lighting is on the sides. The left column in figure 7 illustrates these errors (A is illumination 20 of 21, B is illumination 15 of 21, C is illumination 6 of 21, D is illumination 2 of 21). The error curve with distance maps preprocessing is under the error curve of the ”Standard experience”, which shows the interest of the method. Moreover, the error curves shows that when the distance maps preprocessing is applied, facial features searching is less dependent on the direction of the illumination than in the standard method case. Indeed, we can see a very clear rise of the ”Standard experience” curve error in illuminations 2 to 8 and 16 to 22, i.e. when lighting is on the sides, the error becomes weak when lighting is in front of faces (illuminations 9 to 15), while the ”Distance experience” curve error is low and very stable for all illuminations.

Figure 6: A search result in the ”Standard experience” (on the left) and in the ”Distance experience” (on the right). On the top, the shape is displayed and on the bottom, the texture is displayed.

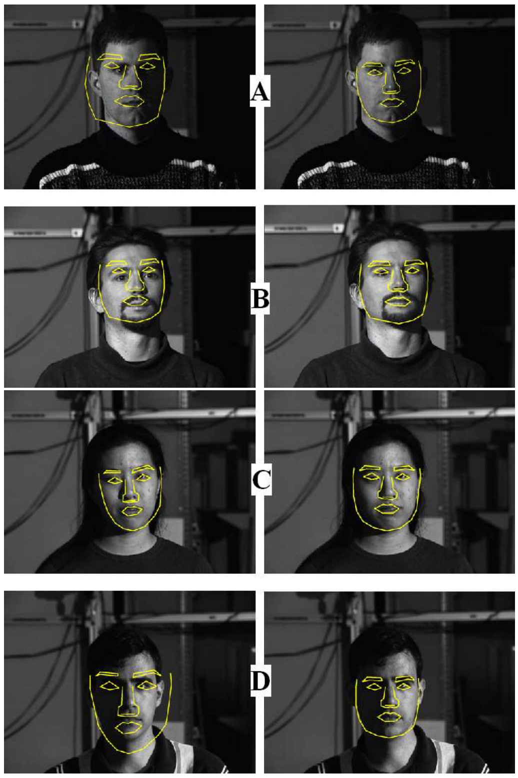

Figure 7: Examples of face searching on PIE facial images with the ”Standard experience” on the left and with the ”Distance experience” on the right. These are 4 examples of the 21 illuminations from the right to the left of a face (from images on the top to images on the bottom).

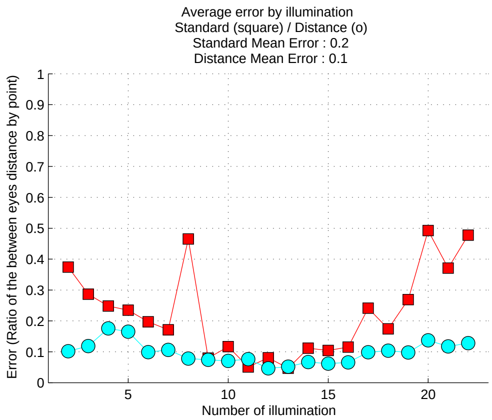

Figure 8: Average error per illumination for the 8 faces.

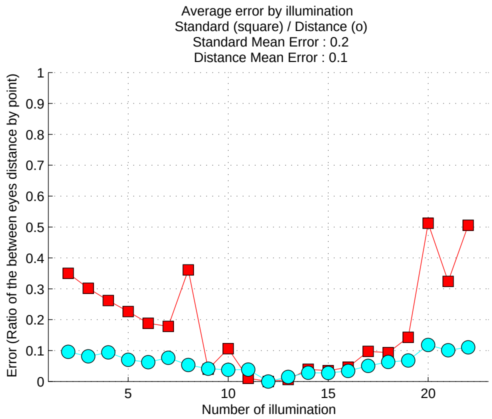

Figure 9: Average error per illumination for the 4 known faces.

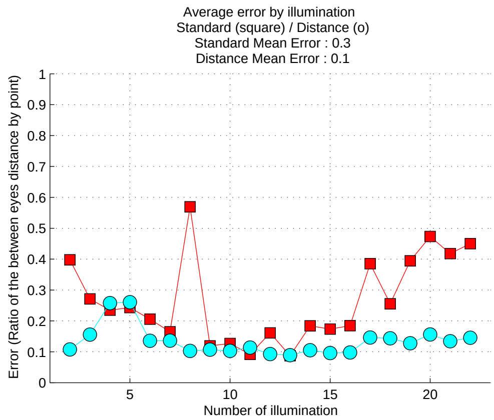

Figure 10: Average error per illumination for the 4 unknown faces.

In figures 9 and 10 we separated the 4 known faces and the 4 unknown faces to create error curves. We can remark that for both method, errors are smaller with known faces than with unknown faces, which is a logical outcome of AAMs.

5 CONCLUSION

We have described the use of a new preprocessing with the Active Appearance Model in facial features searching under variable illumination. The method is an edge-based approach with information concerning distances between edges gathered in images called ”Distance maps”. This contribution allows distance relations between different edges of a searched shape in textures images to be considered. Experiments demonstrated the robustness of this method with several images from the CMU PIE database. Indeed, experiments show that when distance maps preprocessing is applied, that is to say when distance maps textures are used as input of AAM method instead of original images textures, facial features searching is much less dependent upon the direction of illumination than using the standard method.

REFERENCES

Cootes, T. F. (2005). http://www.isbe.man.ac.uk/˜bim/software.

Cootes, T. F., Cooper, D., Taylor, C. J., and Graham, J. (1995). Active shape models - their training and application. In *Computer Vision and Image Understanding*. Vol.61, No. 1, p.38-59.

Cootes, T. F., Edwards, G. J., and Taylor, C. J. (1998). Active appearance models. In *ECCV'98, European Conference on Computer Vision*. Vol.2, p.484-498.

Hu, C., Feris, R., and Turk, M. (2003). Active wavelet networks for face alignment. In *BMVC'03, British Machine Vision Conference*.

Huang, Y., Lin, S., Li, S. Z., Lu, H., and Shum, H. Y. (2004). Face alignment under variable illumination. In *FGR'04, Automatic Face and Gesture Recognition*.

Sim, T., Baker, S., and Bsat, M. (2002). The cmu pose, illumination, and expression (pie) database. In *FG'02, IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE.

Zuiderveld, K. (1994). *Contrast Limited Adaptive Histogram Equalization*. Graphics Gems IV, p.474-485 edition.