<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Separating Directional Lighting Variability in Statistical Face Modelling Based on Texture Space Decomposition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Ionita, Mircea C.; Bacivarov, Ioana; Corcoran, Peter</td>
</tr>
<tr>
<td><strong>Publication Date</strong></td>
<td>2007-07-04</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>IEEE</td>
</tr>
<tr>
<td><strong>Link to publisher's version</strong></td>
<td><a href="http://ieeexplore.ieee.org/search/srchabstract.jsp?tp=&amp;arnumber=4288566">http://ieeexplore.ieee.org/search/srchabstract.jsp?tp=&amp;arnumber=4288566</a></td>
</tr>
<tr>
<td><strong>Item record</strong></td>
<td><a href="http://hdl.handle.net/10379/1615">http://hdl.handle.net/10379/1615</a></td>
</tr>
</tbody>
</table>

Some rights reserved. For more information, please see the item record link above.
SEPARATING DIRECTIONAL LIGHTING VARIABILITY IN STATISTICAL FACE MODELLING BASED ON TEXTURE SPACE DECOMPOSITION

Mirea C. Ionita, Ioana Bacivarov, Peter Corcoran
National University of Ireland, Galway
Electronic Engineering Department

ABSTRACT
In this paper we propose a simple method for decomposing the linear texture space of a facial appearance model into two linear subspaces, one for inter-individual variability and another for variations caused by directional changes of the lighting conditions. The approach used is to create one linear subspace from individuals with uniform illumination conditions and then filter a set of images with various directional lighting conditions by projecting corresponding textures on the previously built space; the residues are further used to build a second subspace for directional lighting. The resulted subspaces are orthogonal, so the overall texture model can be obtained by a simple concatenation of the two subspaces. The main advantage of this representation is that two sets of parameters are used to control inter-individual variation and separately intra-individual variation due to changes in illumination conditions.

Index Terms—statistical face models, eigenfaces, directional illumination, PCA, AAM

1. INTRODUCTION
The appearance of an object can be represented by statistical models trained using a set of annotated image examples. This is thus highly dependent on the way in which the model is trained. A new image can be interpreted by finding the best plausible match of the model to the image data. While there has been a great deal of literature in computer vision detailing methods for handling statistical models for human faces, there still exist some problems which need to be solved. For example, statistical models for human faces are sensitive to illumination changes, especially if the lighting in the test image differs significantly from the conditions learned from the training set. The appearance of a face can change dramatically as the lighting conditions change. Due to the 3D aspect of the face, direct lighting source can cast strong shadows and shading which affect certain facial features. It is well known that the variation due to illumination changes can be greater than the variation between individual faces.

Various methods have been proposed to overcome this challenge. Feature-based approach seeks to utilize features that are invariant to lighting variations. In [1], Hu et al. propose to replace the AAM texture by an active wavelet network for face alignment, while in [2] texture is replaced by distance maps, robust against lighting variations. Other methods rely on removing illumination component using lighting models. The linear subspace approach [3], [4], [5] approximates the human face surface with a Lambertian surface and computes a basis for a 3D illumination subspace, using images acquired under different lighting conditions. The illumination convex cone goes a step further with the model, taking into account shadows and multiple lighting sources [6], [7], [8]. More complex models have been proposed like the geodesic illumination basis model [9], or the 3D linear subspace model that segments the images into regions with directions of surface normals close to each other [10]. The canonical form approach appears as an alternative, where an attempt to normalize the variations in appearance by image transformations or by synthesizing a new image from the given image in a normalized form is undertaken. Recognition is then performed using this canonical form [11], [12].

In [13], [14], the decomposition of an eigenface into two orthogonal eigenspaces is proposed for realizing a general face recognition technique, under lighting changes. A somewhat similar approach is used in [15] for face tracking, where the face is represented by the addition of two approximately independent subspaces to describe facial expressions and illumination, respectively.

In [16], [17], facial appearance models of shape and texture are employed and non-orthogonal texture subspaces for lighting, pose, identity, and expression are extracted using appropriate image sets. An iterative expectation-maximization algorithm is then applied in order to maximize the efficiency of facial representation over the added subspaces. The projections on each subspace are then used to recalculate the subspaces. This approach is shown to improve the identity recognition results.

In our paper we propose a simple algorithm that permits to handle the illumination changes, obtaining a general and robust facial appearance model. The main idea is to decompose the texture space into two orthogonal subspaces, one of uniform illumination and the second for illumination variability. The advantage of our algorithm is that two separate sets
of parameters are used to control the variation between individuals and the variation in illumination conditions. Another advantage is that an exhaustive image database for training the statistical model is no longer needed. The outline of the paper is as follows. Section 2 presents a short overview of the statistical models of appearance. Section 3 describes the method for separating directional lighting variation, while Section 4 shows how the two separate texture models can be fused. Section 5 shows how the proposed model can be fitted and presents also some experimental results. Finally, the conclusions and future work are presented in Section 6.

2. STATISTICAL FACE MODELS OF SHAPE AND TEXTURE

Shapes are defined as a number of landmarks used to best describe the contour of the object of interest (i.e. the face). A shape vector is given by the concatenated coordinates of all landmark points, as \( (x_1, x_2, \ldots, x_L, y_1, y_2, \ldots, y_L)^T \), where \( L \) is the number of landmark points.

The shape model is obtained by applying PCA on the set of aligned shapes

\[
s = \bar{s} + \Phi_s b_s, \tag{1}
\]

where \( \bar{s} = \frac{1}{N_s} \sum_{i=1}^{N_s} s_i \) is the mean shape vector, with \( N_s \) the number of shape observations; \( \Phi_s \) is the matrix having as its columns the eigenvectors; \( b_s \) defines the set of parameters of the shape model.

Texture, defined as the pixel values across the object of interest, is also statistically modelled. Face patches are first warped into the mean shape based on a triangulation algorithm. Then a texture vector \( (t_1, t_2, \ldots, t_P)^T \) is built for each training image by sampling the values across the warped (shape-normalized) patches with \( P \) being the number of texture samples.

The texture model is also derived by means of PCA on the texture vectors

\[
t = \bar{t} + \Phi_t b_t, \tag{2}
\]

where \( \bar{t} = \frac{1}{N_t} \sum_{i=1}^{N_t} t_i \) is the mean texture vector, with \( N_t \) being the number of texture observations; \( \Phi_t \) is again the matrix of eigenvectors, and \( b_t \) the parameters for the texture model.

The sets of shape and texture parameters \( c = (W_s b_s \ b_t)^T \) are used to describe the overall appearance variability of the modelled object, where \( W_s \) is a vector of weights used to compensate for the differences in units between shape and texture parameters.

3. TEXTURE SPACE DECOMPOSITION

To construct our texture models we used images from the Yale Face Database B [8]. First, a shape model and a texture model are built from different individuals at constant frontal illumination (see Fig. 1 and Fig. 2). The shape and texture models are built as described in (1) and (2), keeping also the same notations. This type of illumination stands as an approximation for uniform illumination conditions. We further refer to the resulted texture eigenspace as the Uniform Lighting Subspace (ULS) of the individuals.

![Fig. 1. Variation between individuals.](image1)

![Fig. 2. Textures in the shape-normalized space.](image2)

For each individual we now consider images with various directional lighting conditions (see Fig. 3). The same reference shape is used to obtain the new texture vectors \( g \), which ensures that the previous and new texture vectors have all equal lengths.

![Fig. 3. Variation both between individuals and in illumination conditions.](image3)

We then filter these vectors by projecting them on ULS ((3),(4)).

\[
\mathbf{b}_{opt} = \Phi_t^T (\mathbf{g} - \bar{t}), \tag{3}
\]

\[
\mathbf{g}_{filt} = \bar{t} + \Phi_t \mathbf{b}_{opt}. \tag{4}
\]

The residual texture is given by

\[
\mathbf{g}_{res} = \mathbf{g} - \mathbf{g}_{filt} = \mathbf{g} - \bar{t} - \Phi_t \mathbf{b}_{opt}. \tag{5}
\]

The residues are used to create an orthogonal texture subspace of directional lighting. We call this the Directional Lighting...
Subspace (DLS). The directional lighting texture model is described, similar to (2), by

\[ \mathbf{g}_{\text{res}} = \mathbf{g}_{\text{res}} + \Phi_{g} \mathbf{b}_{g} \]  

(6)

4. TEXTURE MODELS FUSION

DLS is built from the residual (difference) images subsequent to a projection on ULS. Thus, DLS is orthogonal to ULS.

The fused texture model is given by

\[ \mathbf{t}_{\text{fused}} = \mathbf{i} + \Phi_{g} \mathbf{b}_{t} + \Phi_{g} \mathbf{b}_{p} \]  

(7)

The fusion between the two texture models is realized by a simple (weighted) concatenation of parameters. A vector of weighted shape parameters concatenated with the texture parameters,

\[ \mathbf{c} = \begin{pmatrix} W_{b} \mathbf{b}_{s} \\ W_{t} \mathbf{b}_{t} \\ W_{p} \mathbf{b}_{p} \end{pmatrix} \]  

(8)

where \( W_{b} \) and \( W_{d} \) are two vectors of weights used to compensate for the differences in units between the two sets of texture parameters, and for the differences in units between shape and texture parameters, respectively.

5. MODEL FITTING & RESULTS

The Active Appearance Model (AAM) is a common technique used to optimize the parameters of a statistical model of appearance [18]. As demonstrated by Batur et al. [19], the standard AAM algorithm, which uses a gradient estimate built from training images, cannot be successfully applied on images when important variations in the illumination conditions are present. This is because the estimated gradient specializes around the mean of the dataset it is built from.

The solution proposed by Batur et al. is based on using an adaptive gradient AAM [19]. The gradient matrix is linearly adapted according to texture composition of the target image, in order to generate a better estimate of the actual gradient. This technique represents a tradeoff between using a fixed gradient (AAM) and numerically computing the gradient matrix at each iteration (the standard optimization technique).

Although more computationally expensive, we decided to use at this stage of our research the standard optimization technique, as it also provided us with a more accurate solution. We are also currently working to develop a fast optimization algorithm for our model, trying to make use of the AAM technique for optimizing all model parameters except the illumination parameters.

We tested the model convergence on a new set of images with some variations of the illumination conditions. We found no important differences in the quality of the fit between our fused model and a standard model based on a single texture eigenspace built from the combined set of training images.

The important advantage of fitting a fused model is that it offers control on illumination, based on the separate set of illumination parameters. Thus, after fitting the model, one can obtain an estimate of the illumination conditions present, normalize the illumination (e.g., to uniform illumination), or generate different illumination conditions.

![Fig. 5. (a) Input images; (b) Synthesized image/face patches; (c) The synthesized texture is replaced with the original (real) texture; (d) Illumination normalized image/face patch.](image)

In Fig. 5 we show an example where we fit the model on a new image in which a spotlight is present at the person’s left side. As it can be seen from Fig. 5(b), the face patch is correctly segmented, yet the person’s identity is not so accurately reproduced in the synthesized image. This is due to the limitations of the ULS model which was built using a small number of observations. Yet one can now extract the real texture of the original image from inside the fitted shape (see Fig. 5(c)). The real texture vector can be viewed as the projection of the individual on the directional lighting eigenspace plus a constant vector representing the individual under uniform lighting. This can be written as

\[ \mathbf{t}_{\text{real}} = \mathbf{i}_{\text{unif}} + \Phi_{g} \mathbf{b}_{\text{light}} \]  

(9)

where \( \mathbf{b}_{\text{light}} \) are the illumination parameters which were estimated during the overall optimization stage. By altering \( \mathbf{b}_{\text{light}} \) new illuminations can be generated. Fig. 5(d) shows the results for setting all parameters to zero and obtaining illumination normalization.
6. CONCLUSIONS

We described a statistical face model based on texture space decomposition which enabled the separation of illumination variability and intra-individual variability. The model can be useful in applications where the control over the illumination conditions is desired.

The separation of the sets of parameters can also be useful as one does no longer need an exhaustive image database for all the modelled components. Thus, an appropriate (separate) database can be used for modelling components like shape (including pose) variability, inter-individuals variability, or directional lighting.

An improved shape model, designed to include also pose variability, can be used to further enhance the capabilities of the overall appearance model.

The technique can also be extended to color images using a color training database for modelling individual variability under uniform lighting.

Thus, such models can offer solutions in applications like face recognition or face tracking for dealing with variations in pose or illumination conditions, where many of the current techniques still show weak points.

7. REFERENCES


