Look at the Driver, Look at the Road: No Distraction! No Accident!

Cheng Guan

July 4, 2018

Figure 1. 64 key point landmarks (left). Symmetric Delauney triangulation (middle). Asymmetric intensity variations (right).

1. Asymmetric Appearance Models

In this section, the authors introduce the *Appearance models* (AM). originally introduced by Cootes*et al.* [1], is widely used for object modeling, especially in the context of facial processing. Many research efforts have solved this as an optimization problem to find improved fitting algorithms and reduce matching errors.

1.1. Implementation

AM combines a shape model and a texture model. To define the face AM, the author needs to train changes in face shape (as a shape model) and face intensities (as a texture model).

Considering 64 point-landmarks, as illustrated in Fig. 1, right, and using the MUCT face dataset [2], the authors create an annotated face dataset in order to train a generic face-shape model. Following the standard AM approach [1], and applying a uniform coordinate system, annotated faces are represented by a vector $f = [x_0, y_0, ..., x_i, y_i]^\top$. A face shape model is defined as Eq.1:

$$f = \bar{f} + P_s b_{si} \tag{1}$$

where $\bar{f}$ is the mean face shape applying *principal component analysis* (PCA) on the available face data, P_s is an orthogonal matrix of face-shape variations, and b_s is a vector of face-shape parameters (given in distance units). By applying a translation (t_x, t_y), and a rotation and scaling $(s_x = s \cdot \cos\theta - 1, s_y = s \cdot \sin\theta)$, each sample face is warped into the mean shape model, thus creating a new face F. Let $F = S_t(f)$ be this warped image, where S_t is the warping function, and $t = [s_x, s_y, t_x, t_y]^\top$ is the pose parameter vector. Fig. 2 illustrates the steps for creating the

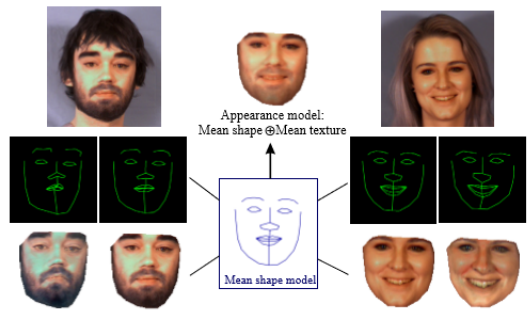

Figure 2. Conversion of face shape and face texture models of two sample faces into a mean appearance model.

appearance face model based on only two sample faces. The second row of Fig. 2 shows examples of shape variations with different deformation parameters applied to each sample face.

To create a face texture (intensity) model, first a symmetric Delaunay triangulation is applied to shape-feature points, for each sample face (Fig. 1, middle). Considering g as a texture vector of a sample face image, similar to the shape-warping stage, the authors have a mapping $g \to g^*$, where g^* is generated after scaling and adding an offset to current intensity g. This way they create a shape-free "intensity patch" for each sample face given in the training dataset. This is done by raster scanning of the texture vector g, and a linear normalization of g for every half of the face as Eq.2:

$$g_L^* = \frac{g_L - \mu_L \cdot \mathbf{1}}{\sigma_L}, g_R^* = \frac{g_R - \mu_R \cdot \mathbf{1}}{\sigma_L} \tag{2}$$

where μ_L, μ_R and σ_L^2, σ_L^2 are means and variances for the left and right part of the face-intensity patch, g_L, g_R are the left and right half of the g vector, g_L^*, g_R^* are normalized data, and $\mathbf{1}$ is a vector of ones.

Similarly,by applyinga PCA to the normalized intensity data, a face intensity-model is estimated as Eq.3:

$$g_L = \bar{g_L^*} + P_{gL} b_{gL}, g_R = \bar{g_R^*} + P_{gR} b_{gR} \tag{3}$$

where $\bar{g}^*$is the mean vector of normalized gray-level or intensity data, P_g is an orthogonal matrix of texture-modes of variations,and b_g is a vector of intensity parameters in gray level units (Fig. 2, third row). The authors apply this as individual processes for each half of the face.

References

[1] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE TPAMI*, 2001. 1

[2] S. Milborrow, J. Morkel, and F. Nicolls. The muct landmarked face database. *PRASA*, 2010. 1