Using Dynamic 3D Facial Data to Create Appearance Models

Lanthao Benedikt, David Marshall, Paul L. Rosin
School of Computer Science
Cardiff University, U.K

Darren Cosker
School of Computer Science
University of Swansee, UK

Introduction

In this poster, we describe a technique to construct an Appearance Model using dynamic 3D data of facial expressions. Our method extends the work of T.Cootes et al on Active Appearance Model [1], and achieves a high realism and accuracy including fine details such as muscle folds and skin deformations.

A 4D technology-based colour camera is used to capture simultaneously geometry and texture data at high frame rates (48fps). The training process requires the capture of facial expressions based on Action Units which have been validated by researchers from our Psychology Department.

After the model is built, we identify the principal control parameters of the model based on peak Action Units intensities. New facial expressions can be synthesized by altering these parameters.

Modelling and Animation Method

In order to learn the facial dynamics, a small number of landmarks (50 per video frame) are placed at the key features as shown in Figure 1. We first compute a 2D texture map of the 3D head. The landmarks are semi-automatically placed on the 2D texture map and later, the corresponding points on the 3D mesh are computed.

The main challenge is to synthesize fine details such as muscle folds without using a large number of landmarks. Hence, we train our Appearance Model with shape information (e.g. xyz coordinates and surface normals) of all vertices and texture information (e.g. RGB values) of all pixels in the texture map.

We then perform a Principal Component Analysis on the training data matrix and obtain a set of Eigen vectors which represent the different Modes of Variations of the facial dynamics. We usually reorder the Eigen vectors by intensity. Considering that 98% of the variations are usually contained in the first 10 modes, we can reduce the high dimensionality of our model and improve at the same time the processing cost.

Figure 1 – Landmarks are semi-automatically placed at the key facial features e.g. lip boundaries, eye corners....

Using our model, we can analyse a given facial performance by computing the appearance parameter values corresonding to each frame in the video sequence. This allows us to observe the temporal evolution of the appearance parameters across the video sequence [2].

Results

Figure 2 depicts three face poses generated using an Appearance Model which has been built from the Action Unit “disgust”. The appearance parameter corresponding to the first Mode of Variation varies from -2 to 2 by step of 2. We notice that the cheek muscles are accurately rendered without the need to place any landmarks on the cheek surface.

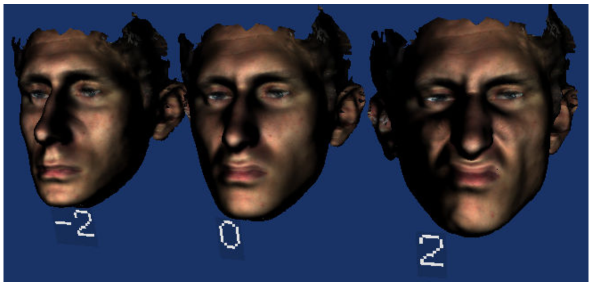

Figure 2 – Synthesized face expressions obtained by altering the value of the first Mode of Variation of an Appearance Model trained from the Action Unit “disgust”.

Given a video footage of an actor performing the Action Unit “disgust”, we can compute the values of the appearance parameters and observe their temporal evolutions, as shown in Figure 3.

Conclusion and Future Works

We dispose at the moment of highly accurate realistic looking 3D face models with parameter control to generate different facial poses. We are able to render fine details e.g. skin deformations and muscle folds despite using only a small number of landmarks.

The next step in the research will be to animate facial dynamics from speech and non-verbal expressions.

References

[1].T. Cootes, G. Edwards, and C. Taylor. Active appearance models. IEEE Trans. PAMI, 23(6):681 – 684, 2001.

[2]. Recent Mirage paper. But which one?

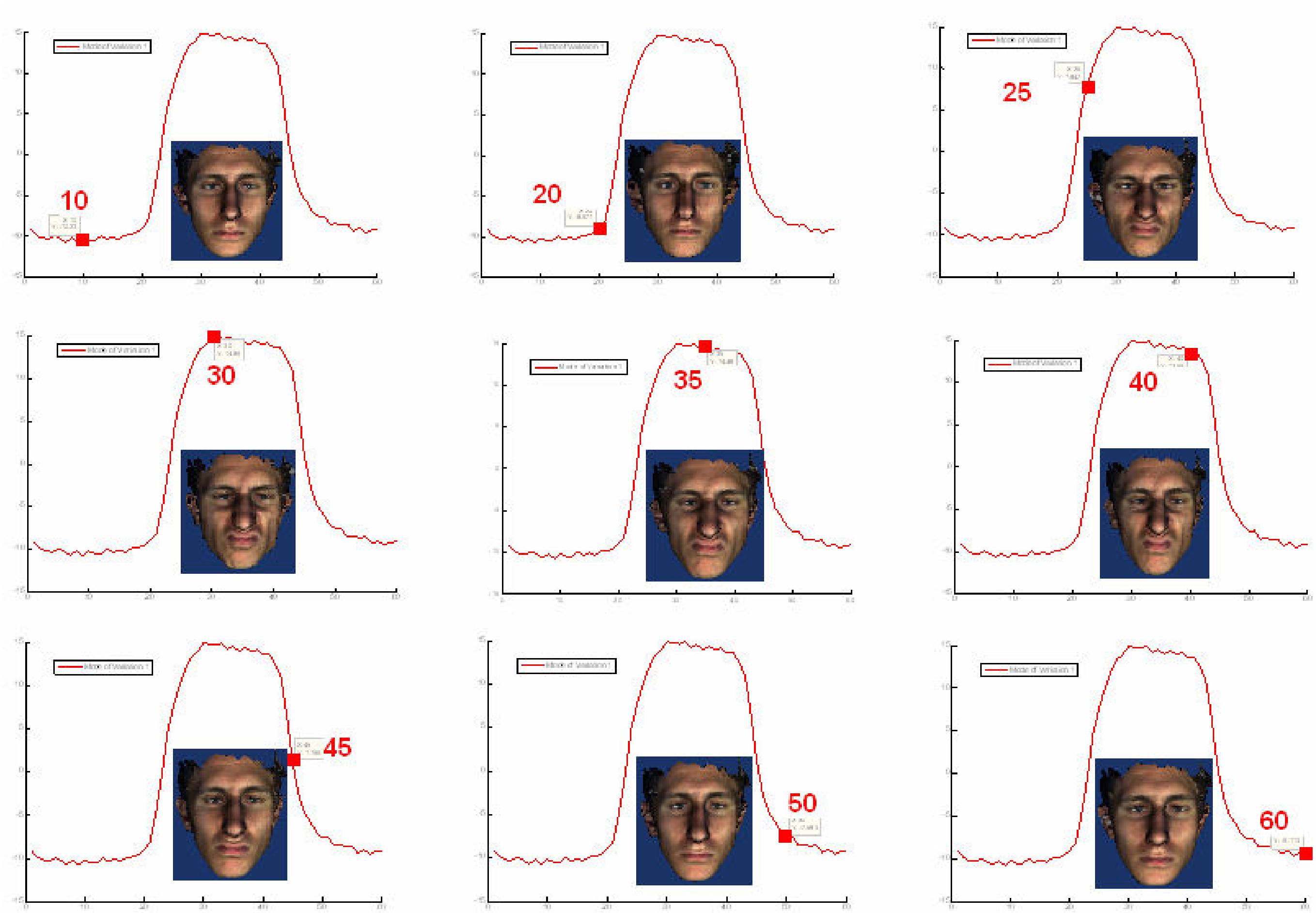

Figure 3 – Video sequence of an actor performing the Action Unit “disgust” and the temporal evolution of the appearance parameter corresponding to the highest Mode of Variation. Graph legend: X axis: frame number. Y axis: value of the appearance parameter.