Synthetic Facial Textures

Taly Gilat, Deirdre O'Brien, Augusto Román
{tgilat, dbobrien, aroman}@stanford.edu

1 Background and Motivation

A full set of 183 Cyberware 3-D scans of adult heads is available for our use. Each head is represented in two images. The first image is a profile image containing the radius of the head in a cylindrical projection from a central, vertical axis. The second image is the texture image containing the associated color (in RGB) at each point of the first image.

It has been shown in previous work that a combined, morphable model created from similar scans can be used to perform useful and interesting image processing tasks such as image approximation[1]. These morphable models are stored as linear combinations of a basis set of profile and texture sets, each of which can take a significant amount of memory. Compressing arbitary 3-d models can be used to reduce the required information for the profile data[5], however the texture data remains very large.

We propose a structure that exploits the information available in the profile data to render a good approximation of the original head using a small number of features from the color image. This would greatly reduce the storage space and transmission bandwidth required.

2 Proposed Approach

The basis of our approach is to represent the color image by a small number of parameters. We plan to use models for eyes, lips, eye-brows and skin, and store or transmit the model parameters in place of the color image. From the profile information it is possible to identify the positions of the facial features and using the models for these features we can 'paint' the features onto the profile data.

2.1 Implementation

The first task will be to align the images from different people with scalings to fit different head sizes and shapes into the same model. Further processing may be needed to compensate for the variation in the images due to side lighting. The next stage will segment the image into skin and non-skin parts.

The skin covers the largest area and we propose representing it as a texture. There are a number of texture analysis and synthesis approaches described in the literature, recent work includes [2, 3, 6, 7, 8, 9]. Many of these build on the algorithm proposed by David Heeger [4], which matches various statistical parameters of the analysis and synthesis features. Much of the recent focus has been to develop methods that incorporate both texture and structural details, such as the regularity associated with brick walls. The skin appears to be a relatively simple unstructured texture and it may be sufficient to use the simpler texture synthesis techniques.

We will begin modeling skin as a single texture throughout. We expect, however, that a small number of different skin textures may be needed, to describe different areas e.g. cheeks, chin stubble and forehead skin types. At the borders between these types, textures will be combined to give a smooth transition.

From an initial inspection of the images it appears that different depths of skin, have different tones. This is particularly notable in the shadows in ears and in wrinkles. We hope to model this tone variation to improve the visual quality of the rendering.

A template will be created for the lips, eyes, eyebrows and perhaps mustaches using the whole dataset. Color parameters will be stored to add to the template during rendering. Given the relatively small area covered by these attributes, we expect that the mean color information will be adequate.

2.2 Performance evaluation

The reduction in bytes required to store the original image and our model parameters will provide a metric for the effectiveness of the algorithm for compression.

We do not seek to produce an exact replica of the color image but rather to produce a 3-D reconstruction that has the same features. For example, the reconstruction for a white person with black stubble may return a very large mean squared error when compared to the original, although the 3-D rendering appears very accurate. Thus mean squared error is not an informative metric here.

The most useful performance evaluation may be a subjective comparison of our 3-D approximation and the original 3-D rendering. Alternatively we could compress the original color image using JPEG to have a size comparable to the size of our parameter file and subjectively compare our 3-D approximation to that using the JPEG image file.

References

[1] V. Blanz and T Vetter. A morphable model for the synthesis of 3d faces. In *SIGGRAPH*, pages 187–194, 99.

[2] A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. In *ICCV*, 1999.

[3] Alexei A. Efros and William T. Freeman. Image quilting for texture synthesis and transfer. *Proceedings of SIGGRAPH 2001*, pages 341–346, August 2001.

[4] D. J. Heeger and J. R. Bergen. Pyramid based texture analysis/synthesis. In *COMPUTER GRAPHICS PROCEEDINGS*, pages 229–238, 1995.

[5] H. Hoppe M. Eck. Automatic reconstruction of b-spline surfaces of arbitrary topological type. In *SIGGRAPH*, pages 325–334, 96.

[6] J Portilla and E P Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1):49–71, October 2000.

[7] Li-Yi Wei and Marc Levoy. Fast texture synthesis using tree-structured vector quantization. In *SIGGRAPH*, 2000.

[8] Li-Yi Wei and Marc Levoy. Texture synthesis over arbitrary manifold surfaces. In *SIGGRAPH*, 2001.

[9] Y. Q. Xu, S. C. Zhu, B. N. Guo, and H. Y. Shum. Asymptotically admissible texture synthesis. In *International workshop. on Statistical and Computational Theories of Vision*, Vancouver, Canada, July 2001.