ISSN 2229-5518

Face Modeling using Segmentation Technique

M.Priya, Dr.S.Purushothaman

Abstract - This work focuses on 3D facial modeling using three images of a face. Each image is taken at 90^{o}. Each image is segmented to identify skin, to locate eye centres, nose profile, mouth profiles. All the three images are combined to form a 3D facial model. The segmented portions of the images are placed on a standard computer plastic model. Subsequently, the various features of the plastic model are animated corresponding to the various positions of the features in the subsequent images.

Index Terms – Modeling, Contextual segmentation, Facial parameters.

1. Introduction

Modeling virtual human has attracted more and more attention from both the research and industrial community. 3D-modeling of human has wide applications from Virtual Reality application (requires real-time) to medical application (requires high accuracy). With the growing power of computer speed and multimedia ability, people would like to have their virtual counterpart as 3D data on the computer screen and utilize it for various applications as follows:

- human-machine interface
- advanced multimedia
- augmented reality
- immersive Virtual Reality
- simulation of human behavior with virtual human
- medical application
- communication (through network)
- multi-media games

There are two basic types of techniques for obtaining 3D human models, according to the different requirements for the models. The first technique focuses on the accuracy and precision of the obtained object model shapes, such as those used in computer aided design (CAD) systems for industrial purpose or medical application. The second technique concentrates on the visual realism and speed for animation of the reconstructed models, such as those used in virtual reality applications.

Systems using the second type of technique focuses on more practical aspects such as how cheap the hardware is and how easy it is to use. These techniques are usually model-based. There are several approaches to the reconstruction of either a face or a body from photographs. These approaches concern mainly visual realism using a high quality image input.

Some methods take a generic model and then both structural and shape information extracted from photographs is used to modify the generic model while others use silhouette information in several views to reconstruct the shape.

The approaches are simple and efficient, but the shape is not as accurate as the one from a laser scanner. The virtual reality application, usually the fast animation capacity with efficient shape representation with less numbers of points compensate with the accurate shape.

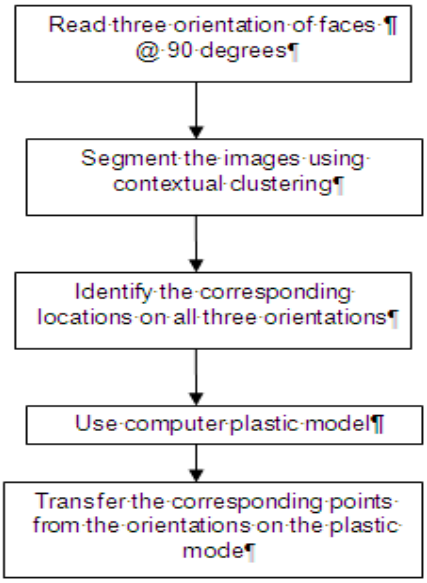
In this work, more on real-time applications since it is no longer fantasy to imagine that one can see herself / himself in a virtual environment moving, talking and interacting with other virtual figures or even with real humans. By advances in algorithms and new developments in the supporting hardware, this fantasy has become a reality. In addition, newly defined moving picture experts group (MPEG-4) supports the standard parameters for communication through network in real-time.

The issues involved in realistic modeling a virtual human for the real-time application purposes are as follows:

- acquisition of human shape data
- realistic high-resolution texture data
- functional information for animation of the human (both face and body)

For the photograph input, use the frontal view and the side view of a face while the frontal view, the side view and a back view are used for a body. Photographs of the whole body cannot provide sufficiently high-resolution for facial information in order to construct a good face model and further facial animation.

2. Schematic Diagram



3. Contextual segmentation

Segmentation refers to the process of partitioning a digital image into multiple regions (sets of pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. The result of image segmentation is a set of regions that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics. Several general-purpose algorithms and techniques have been developed for image segmentation.

Contextual clustering algorithms which segments an image into background (ω_0) and object region (ω_1). The pixel intensities of the background are assumed to be drawn from standard normal distribution.

1. Define decision parameter T_{cc} (positive) and weight of neighborhood information β (positive). Let N_n be the total number of pixels in the neighborhood. Let Z_i be the intensity value of pixel i.

2. Initialization: classify pixel with $z_i > T_\alpha$ to ω_1 and pixels to ω_0. Store the classification to C_0 and C_1.

3. For each pixel I, count the number of pixels u_i, belonging to class ω_1 in the neighborhood of pixel i. Assume that the pixels outside the image area belong to ω_0.

4. Classify pixels with

$$z_i + \frac{\beta}{T_{cc}}\left(u_i - \frac{N_n}{2}\right) > T_\alpha$$

to ω_1 and other pixels to ω_0. Store the classification to variable C_2.

5. If $C_2 \neq C_1$ and $C_2 \neq C_0$, copy C_1 to C_0, C_2 to C_1 and return to step 3, otherwise stop and return to C_2.

4. Facial parameter

The Face and Body animation Ad Hoc Group (FBA) has defined in detail the parameters for both the definition and animation of human faces and bodies. Definition parameters allow a detailed definition of body/face shape, size and texture. Animation parameters allow the definition of facial expressions and body postures. These parameters are designed to cover all natural possible expressions and postures, as well as exaggerated expressions and motions to some extent (e.g. for cartoon characters).

The animation parameters are precisely defined in order to allow an accurate implementation on any facial/body model. Here we will mostly discuss facial definitions and animations based on a set of feature points located at morphological places on the face.

The FAP are encoded for low-bandwidth transmission in broadcast (one-to-many) or dedicated interactive (point-to-point) communications. FAPs manipulate key feature control points on a mesh model of the face to produce animated visemes (visual counterpart of phonemes) for the mouth (lips, tongue, teeth), as well as animation of the head and facial features like the eyes or eyebrows. All the FAP parameters involving translational movement are expressed in terms of Facial Animation Parameter Units (FAPU).

These units are defined in order to allow the interpretation of FAPs on any facial model in a consistent way, producing reasonable results in terms of expression and speech pronunciation. They correspond to fractions of distances between some essential facial features (e.g. eye distance). The fractional units used are chosen to allow enough accuracy.

5. Results and discussion

Figure 1 person 1

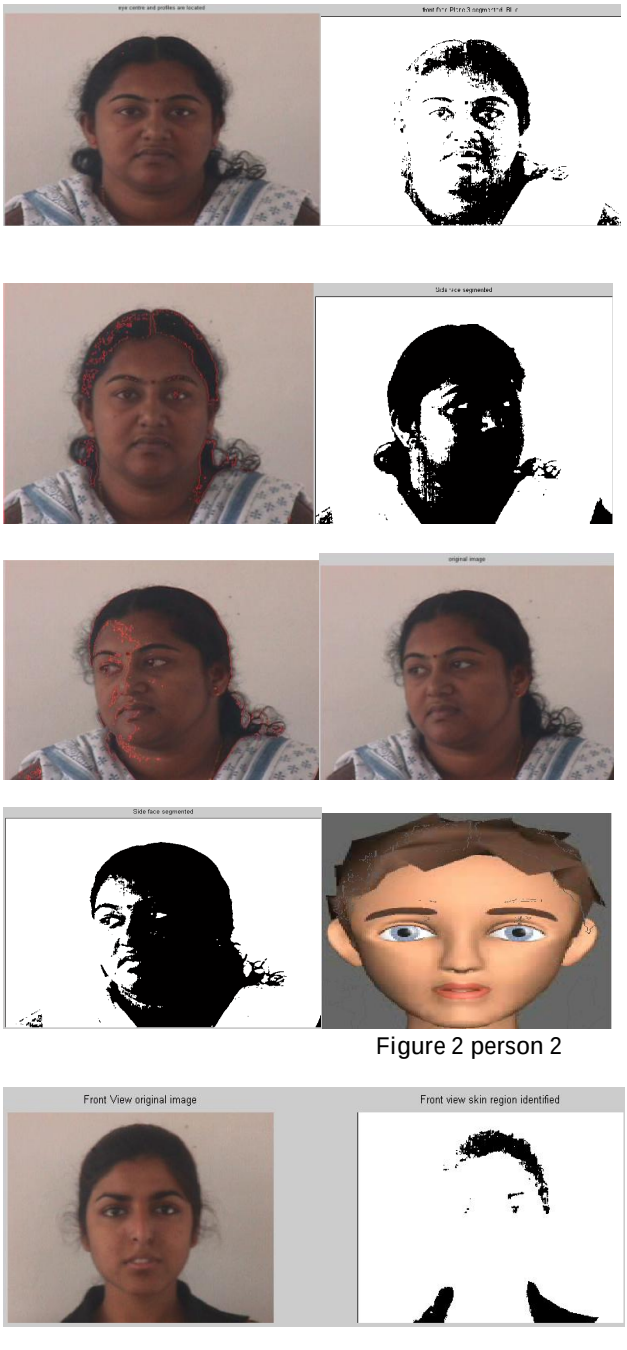


Figure 2 person 2

Figure 3 person 3

6. Conclusion

In this work, 3 persons facial expressions have been considered. For each person, left, front and right postures were considered. This work focuses on 3D facial modeling using three images of a face. Each image is taken at 90^{o}. Each image is segmented to identify skin, to locate eye centres, nose profile, mouth profiles. All the three images are combined to form a 3D facial model. The segmented portions of the images are placed on a standard computer plastic model. Subsequently, the various features of the plastic model are animated corresponding to the various positions of the features in the subsequent images.

References

[1] J. Canny, "A Computational Approach to Edge Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pp. 679-698, June 1986.

[2] I. Craw, H. Ellis, and J. Lishman, "Automatic Extraction of Face Features," Pattern Recognition Letters, vol. 5, pp. 183-187, 1987.

[3] I. Craw, D. Tock, and A. Bennett, "Finding Face Features," Proc. Second European Conf. Computer Vision, pp. 92-96, 1992.

[4] G. Burel and D. Carel, "Detection and Localization of Faces on Digital Images," Pattern Recognition Letters, vol. 15, no. 10, pp. 963- 967, 1994.

[5] M.C. Burl, T.K. Leung, and P. Perona, "Face Localization via Shape Statistics," Proc. First Int'l Workshop Automatic Face and Gesture Recognition, pp. 154-159, 1995.

[6] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," Proc. IEEE, vol. 83, no. 5, pp. 705- 740, 1995.

[7] J. Cai, A. Goshtasby, and C. Yu, "Detecting Human Faces in Color Images," Proc. 1998 Int'l Workshop Multi-Media Database Management Systems, pp. 124-131, 1998.

[8] U. Park, H. Chen, and A.K. Jain, "3d model-assisted face recognition in video," in Proc. Canadian Conference on Computer and Robot Vision, May 2005, pp. 322–329.

[9] K. Bowyer, K. Chang, and P. Flynn. A survey of approaches and challenges in 3d and multi-modal 3d+2d face recognition. *CVIU*, 101(1):1–15, 2006.