

Automatic Feature Extraction and Face Synthesis in Facial Image Coding

Fu-Che Wu, Tzong-Jer Yang, Ming Ouhyoung

Communication & Multimedia Laboratory,
Dept. of Computer Science and Information Engineering,
National Taiwan University, Taipei, 106, Taiwan

Abstract

In this paper, a real-time method that automatically extracts facial feature points is proposed. Based upon this technique, we also present a procedure to encode facial images efficiently. For clarity, the proposed method is divided into analysis part and synthesis part. The analysis part includes face tracking, feature extraction, and pose estimation. All of these procedures are applicable for live video applications. In synthesis part, a new face image is synthesized using one base image and a small amount of feature images. The current performance of feature extraction is about 8 frames per second on a pentium II 233MHz PC with an image size 320x240.

1. Introduction

Extraction of face features is an important issue in facial image related applications, for example, personal identification, very low bit-rate communication for videophone, and human-computer interface. Problems of feature extraction arise from different illumination, head pose, facial expression, as well as occluding objects (e.g., eyeglasses). In recent years, numerous methods have been proposed, for example, contour models[1][2], eigenfaces face[3], template matching, and feature-based matching[4].

In the paper, the features to be extracted are the canthus of the two eyes and a nostril of the nose. These 3 feature points form a rigid body motion. Furthermore, these points are darker than other area of a face, so that it's easier to extract these points. A method that iterative shrink search area is developed. In this method, we first locate the face region, and then find possible locations of features. Finally, a local search is used to locate the exact positions.

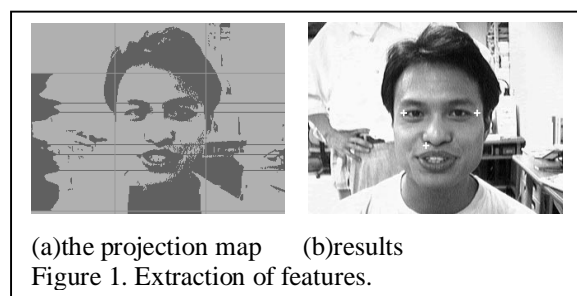
After the three feature points are extracted, an iterative method can be used to estimate the corresponding head motion[5].

In synthesis part, a new face image, instead of using an explicit head model [6]. We use an image based approach that requires only a base image and a small amount of feature images.

2. The Proposed Automatic Feature Extraction

To locate the face region, we first identify the hair region that appears darker. Then the face region can be determined by growing the region with the similar color sampled below the hair region[7].

The original image is then converted to a binary image where pixels with color values larger than a pre-defined threshold are set to 1, and other pixels are set to 0. On this binary image, pixels with value 1 are



accumulated both horizontally and vertically, as shown in Figure 1. Then the following rules are applied to locate feature points.

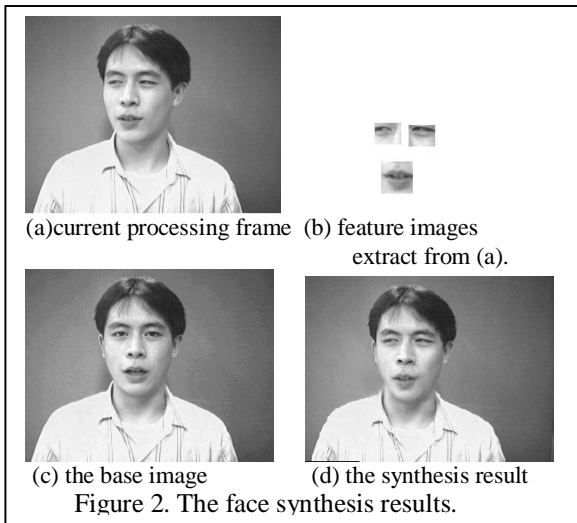
1. Because of the symmetry of a human face, the central peak of the vertical projection represents the line passing through the nose and the mouth.
2. After finding the nose-mouth line, the left and right

peaks are located to find the left and right eyes, respectively.

3. From the horizontal projection, if the first peak occurs at the beginning of face region, it is the hair region's bottom line, so that next peak would be the brow line. Otherwise, the first peak represents the brow line, the next peak is the eye line, and the next 2 is the nose line.
4. The cross sections of the vertical and the horizontal projection are therefore the feature points. Positions of the feature points are further refined. For examples, at the location of left eye we search for the left most corner.

3. Synthesis of Facial Image

In synthesis part, we have a base image, a small amount of feature images, and some parameters that describe the affine transformation of head pose and locations of features. First, the head area is mapped onto a cylinder, and the cylinder is transformed according to affine transformation parameters. Feature images are then overlapped onto the transformed head images. The pixels around the pasted regions are blurred, and thus make it look more smoothly. The result is shown in Figure 2.



5. Experimental results

Our experiments are performed on six students. Under a well-controlled lighting environment, the proposed method can achieve an accurate rate of over 90% with a frame rate of 8 frames/sec, as shown in Figure 3. The method can be further improved by introducing a better face finding algorithm.

References

- [1] M. Kass, A. Witkin, D. Terzopoulous, "Snakes, Active

Contour Models," Proc. IEEE International Conference on Computer Vision, pp.259-268, 1987.

- [2] Andreas Lanitis, Chris J. Taylor, Timothy F. Cootes, "Automatic Interpretation and Coding of Face Images Using Flexible Models" IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 743-756, Jul. 1997.
- [3] Moghaddam. B., Pentland. A., "Probabilistic Visual Learning for Object representation" IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 696-710, Jul. 1997.
- [4] Roberto Brunelli, and Tomaso Poggio, "Face Recognition: Features versus Templates," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 15, No. 10, pp. 1042-1052, Oct. 1993.
- [5] Tzong-Jer Yang, Fu-Che Wu, and Ming Ouhyoung, "Real-time 3D motion estimation in facial image coding" to appear in Proc. of International Conference on Multimedia Modeling 1998, Lausanne, Switzerland, 1998.
- [6] Demetri Terzopoulos, Keith Waters, "Analysis and Synthesis of Facial Image Sequences Using Physical and Anatomical Models" IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 15, No. 6, pp. 569-579, Jun. 1993.
- [7] Tony S. Jebara, Alex Plentland, "Parametrized Structure from Motion for 3D Adaptive Feedback Tracking of Faces" Proceeding of the IEEE on Computer Vision and Pattern Recognition, pp. 144-150, 1997.

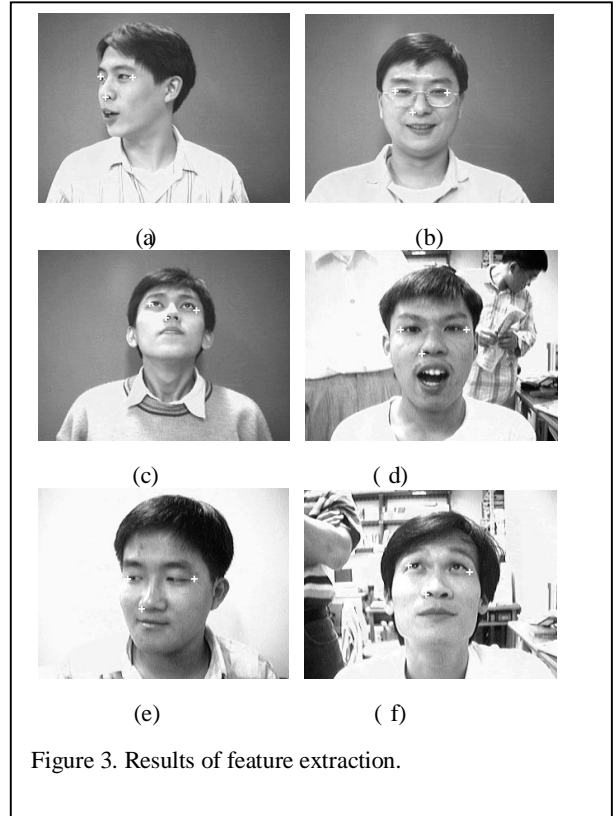


Figure 3. Results of feature extraction.