Online Modeling and Tracking of Pose-Varying Faces in Video

Xiaoming Liu
Visualization and Computer Vision Lab
General Electric Global Research Center
Schenectady, NY, 12309
Email: liux@research.ge.com

Tsuhan Chen
Advanced Multimedia Processing Lab
Department of Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh, PA, 15213
Email:tsuhan@cmu.edu

DESCRIPTION

Human face recognition is an interesting and popular topic in the vision community. The most important part of a face recognition system is to handle all kinds of variations through *modeling*. There are many different kinds of variations, such as pose, illumination, expression, aging, etc., among which pose variation is the hardest one to model [1].

We proposes a face *mosaicing* approach to model both the facial appearance and geometry from pose-varying videos, and apply it in face tracking and recognition. The basic idea is that by approximating the human head as a 3D ellipsoid, multi-view face images can be back projected onto the surface of the ellipsoid [2], and the surface texture map is decomposed into an array of local patches [3]. Patches from multi-view images are allowed to move locally in order to achieve better corresponding. Also the amount of movement is an indication of how much the actual head geometry deviates from the 3D ellipsoid. Finally the corresponding patches are trained to model facial appearance. And the deviation model obtained from patch movements is used to model the face geometry. We can see that mapping from the image plane and the ellipsoid's surface takes care of the pose variation, and the patch-based appearance model takes care of other variations, such as expression, etc. This approach can be used to build either an individual face model, as shown in the first part of the video, or a universal face model, as shown in Figure 1.

During the online modeling process, the face position and pose in the first frame is assumed to be known for a given video sequence. For each subsequent frame, the algorithm estimates the face position and pose, and generates a texture map, which is further utilized in updating the mosaic model. The top-left part of the video shows the input face video overlaid with the tracked face position, which is displayed as an ellipsoid. The top-right part shows a weighting map indicating the importance of each patch, and as well as the mean of the deviation model displayed as a vector map. The middle-left part shows the texture map generated from the current frame. The middle-right part shows the mean of the appearance model. And the bottom shows the first two eigenvectors of the appearance model.

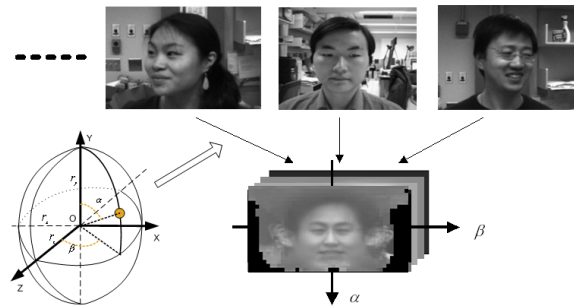In addition to being applied in image-based face recognition

Fig. 1

A UNIVERSAL STATISTICAL FACE MOSAIC MODEL CAN BE GENERATED FROM IMAGES OF MULTIPLE SUBJECTS WITH DIFFERENT POSES.

[3], the face mosaic model can also be used in video-based pose-robust face tracking and recognition. Face tracking is to determine the location, pose of a face in each frame. Since the mapping parameter $\mathbf{x}$ contains all these information, the face tracking is equivalent to estimating $\mathbf{x}$. Given one video frame, we use the condensation method for this estimation. After estimating $\mathbf{x}$ and generating a texture map, we use the distance between the texture map and the individualized mosaic model for the recognition purpose. Using our mosaic model, we have observed satisfying tracking and recognition performance from video sequences with face images. The second part of the video shows the tracking results (both position and poses) of a video sequence containing pose-varying faces. The last part of the video shows a recently collected face video database named Face In Action (FIA) database [4]. By mimicking the *passport checking* scenario, six cameras capture human faces simultaneously from three different poses. We perform the collection in both indoor and outdoor environment, as well as three sessions monthly apart.

REFERENCES

[1] P.J. Phillips, P. Grother, R.J Micheals, D.M. Blackburn, E Tabassi, and J.M. Bone, "FRVT 2002: Evaluation Report", March 2003.
[2] Xiaoming Liu and Tsuhan Chen, "Geometry-assisted Statistical Modeling for Face Mosaicing", In *ICIP 2003*.
[3] Xiaoming Liu and Tsuhan Chen, "Pose-Robust Face Recognition Using Geometry Assisted Probabilistic Modeling," to appear in *CVPR 2005*.
[4] http://amp.ece.cmu.edu/projects/FIADataCollection/