3-D Facial Expression Representation using Statistical Shape Model

Wei Quan, Bogdan J. Matuszewski, Lik-Kwan Shark, Djamel Ait-Boudaoud
Applied Digital Signal and Image Processing Research Centre
University of Central Lancashire
Preston PR1 2HE, UK
{wquan,bmatuszewski1,lshark,dait}@uclan.ac.uk

In the recent years, some research efforts have been made in facial expression representation and recognition for human-computer interaction (HCI) systems, video conferencing and augmented reality, biometrics and medical applications. Most of the current facial expression representation methods are 2-D image based. Inherently these methods have difficulty in handling large head pose variations, changing lighting conditions and subtle variability of facial expressions. By using 3-D/4-D representations instead of 2-D images, the performance of the facial expression recognition can be significantly improved in terms of robustness and accuracy due to the extra information present in the additional dimensions. In our research, a novel approach is proposed for effective 3-D/4-D facial expression representation based on the statistical shape modelling methodology.

The method proposed uses a shape space vector to model surface deformations, and modified iterative closest point (ICP) method to calculate the point correspondence between each surface. The shape space vector is constructed using principal component analysis (PCA) computed for the typical surfaces represented in a training data set. In our work, the calculated shape space vector is used as a feature space for a subsequent facial expression classification. Comprehensive 3-D/4-D face data sets have been used for building the deformation models and for testing. These include 3-D synthetic data generated from FaceGenModeller® software, 3-D facial expression data captured by a static 3-D scanner in the BU-3DFE database and 3-D video sequences captured at the ADSIP research centre using a 3dMD dynamic 3-D scanner.

In order to validate a use of the shape space for the facial expression recognition the landmarks on the 3-D faces from the BU-3DFE database are utilized first. The 83 landmarks are manually marked on the specified area of the face such as eyebrow and lips. The experimental results show that low dimensional shape space vectors exhibit an ability to distinguish between the facial expressions as shown in figure 1. After the validation that the shape space can be used as a feature space for the facial expression recognition the next question which needs to be answered is if the shape space can be effectively build from all available 3-D data instead of the landmarks. In this stage, the 3-D synthetic faces generated from FaceGenModeller® are applied. In order to obtain a correct statistical model, points on each face from the training data set must be in correspondence. This is essential, since an incorrect correspondence can either introduce too much variation or lead to illegal instances of the model. Since the correspondence information is provided in the 3-D synthetic data, a statistical model can be built directly without the need for the correspondence search. The results show that the shape space vector also have ability to distinguish the facial expression based on the whole face figure 2. In the third stage,
the BU-3DFE database is used, without the landmarks, and as there is no correspondence information offered in the BU-3DFE database, this requires an additional method for correspondence search during the model building. Figure 3 shows an example of deformable registration achieved for a real 3-D static data. In the last experiment, the results obtained for the static 3-D faces are extended for the use with the dynamic 3-D face sequences figure 4.

Figure 3: Result for real measurement data: (a) texture of real data, (b) surface of real data, (c) mean shape, (d) registered mean shape, (e) closest distance ≤ 2mm, (f) ≤ 4mm, (g) ≤ 6mm, (h) ≤ 8mm, (i) ≤ 10mm

Figure 4: Faces sampled from the dynamic 3-D sequence