Learning Based Automatic Face Annotation for Arbitrary Poses and Expressions from Frontal Images Only

Akshay Asthana, PhD Student (2008-Present), RSISE
Supervisors: Roland Goecke and Tom Gedeon
To be presented at CVPR 2009

Introduction
We propose an approach for automatic annotation of face images at any arbitrary pose and expression (Fig. 1) that only requires annotated frontal images, thus drastically simplifying the model building process. We propose -
- a data-driven approach to learn the correspondence between manually annotated landmark points of the frontal and varying viewpoint images of a face,
- a framework to reconstruct virtual images, of any arbitrary pose and expression, from frontal view images and use these to automatically annotate the unseen images.

Virtual Image Reconstructor

The basic idea here is to learn the correspondence between the landmark points of the frontal and non-frontal images exhibiting arbitrary expressions. We extract 3 vectors: Normalisation Vector, Centroid Vector, and Point Vector from each frontal and non-frontal images (Fig. 2).

- **Normalisation Vector** is a 1D vector containing the location of each of the n landmark points in the normalised frame.

- **Centroid Vector** is a 10 vector containing information about the normalisation distances used to normalise the feature vectors.

- **Point Vector** is a 1D vector containing the location of the centroids of six individual facial features (left and right eyebrows, left and right eyes, nose and mouth) in the normalised frame.

Reconstruction of Virtual Images

Given an annotated frontal face image, a virtual image is reconstructed by predicting the new landmark locations and warping the texture from the frontal image via Piecewise Affine Warping (PAW). The step by step procedure is shown in Fig 3.

Experimental Results

We conducted experiments on CMU PIE, Face Pointing and FERET databases. The regression models, for reconstruction of virtual images, were trained for 6 different poses i.e. 20° Up and 20° Down, 22.5° Left and Right, 45° Left and Right. Given annotated frontal images of an unseen person with arbitrary facial expressions, virtual images were reconstructed (See Fig 5) and were used to automatically annotate unseen images. Fig 4 shows these annotation results.

Future Work

In future, we plan to extend the approach for automatically annotating the frontal images. We also intend to use the reconstructed virtual images to provide a plausible solution to the problem of pose-invariant face recognition from a single image per person.