Abstract

Facial image analysis is a major branch of human-computer interaction. Among the techniques, facial landmark fitting is one of the fundamental prerequisites for the further analysis. The landmark fitting task is to address the problem of deforming a group of predefined 2D landmarks into the optimal positions of a given facial image. Many canonical methods succeeded to achieve good performance, e.g. Active Shape Model, Active Appearance Model, Constrained Local Model, etc. However, the holistic constraint methods, i.e. Active Shape Model, may underestimate the subtle local shape variance caused by face diversity. Moreover, in unconstrained environments, the large head pose variation and partial occlusion make the problem even more challenging.

In our work, we firstly introduce the problem of facial landmark localization and its relevant canonical and state-of-the-art techniques. Among the existed methods, when facilitating to the facial images under unconstrained environments, they may encounter problems from the large pose variation, partial occlusion, unpredictable illumination, etc. A pose-robust optimized part mixture and cascaded deformable shape model combined localization method is proposed to alleviate the arbitrary head pose problem. Then a consensus of occlusion-specific regressors method is proposed to overcome the partial occlusion inside facial images. Further, we aim at building a unified framework to jointly deal with the pose and occlusion problems. A pose-conditioned hierarchical part based regression method is designed to condition the pose into several pre-defined subspaces and localize the key positions in a hierarchical way, in which the occlusion is detected by the part regressors and further propagated through the hierarchical structure. The proposed facial landmark localization methods have shown more promising performance than those state-of-the-arts in both accuracy and efficiency, compared on
both lab-environmental databases and multiple challenging faces-in-the-wild databases.

With good key points localization, many higher-level analysis can be conducted, such as facial expression recognition and human interaction understanding. We present a user-defined facial expression recognition task by introducing a fused deep neural network structure, which boosts the generalization ability of the extracted features. We further show a video-based automatic deception detection framework utilizing the landmark localization, expression recognition modules. The improved results from the applications further validate the advantages of our landmark localization methods under types of uncontrolled conditions.