Model-Based De-Identification of Facial Images

**Ralph Gross[1], PhD, Latanya Sweeney[1], PhD, Jeffrey Cohn[2], PhD,
Fernando de la Torre[3], PhD, and Simon Baker[4], PhD**

[1] Data Privacy Lab, School of Computer Science, Carnegie Mellon University
[2] Department of Psychology, University of Pittsburgh
[3] Robotics Institute, Carnegie Mellon University
[4] Microsoft Research, Microsoft Corporation

Abstract

Recent advances in both camera technology as well as supporting computing hardware have made it significantly easier to acquire, transmit, process and store large amounts of image data. As a consequence a number of image databases, specifically face databases, have been collected, often with the expressed goal of sharing the data with others. Due to concerns about the privacy of the individuals visible in the scene, data dissemination is particularly difficult for medical face databases depicting patients.

For most data usage scenarios however, knowledge of the identity of people in the image is not required. This makes the case for image de-identification, the removal of identifying information from images, prior to sharing of the data. The implicit goal of these methods is to protect privacy and preserve data utility, for example the ability to recognize gender or facial expressions from de-identified images. While privacy protection methods are well established for field-structured data, work on images is still limited.

The appearance of a face is influenced by a number of factors including (but not limited to) identity, pose, facial expression, illumination, and gender. Algorithms have been developed to factorize facial appearance into these underlying components using linear and bilinear as well as tensor models. In this paper we describe a novel framework for the de-identification of facial images. We use Active Appearance Models (AAMs) for initial parameterization, before we apply a multi-factor model which unifies linear, bilinear, and quadratic models to factorize the face parameters into identity and non-identity components. The resulting representation is then de-identified according to one of multiple privacy protection models.

We evaluate the proposed face de-identification system on a large-scale still image face database, the CMU Multi-PIE database. We measure face recognition performance on de-identified images as well as data utility (quantified as accuracy of facial expression recognition) using a support vector machine classifier trained on the original data. In comparison to naive blur de-identification as well as to the previously proposed formal k-Same algorithm our method preserves data utility much better while providing equivalent privacy protection.Our framework directly extends from single images to image sequences. We demonstrate this by de-identifying videos of subjects displaying pain expressions.