C21 Performance of Matching Algorithms in Non-Standard Expression-Variant Faces

Petra Urbanová, PhD, Masaryk University, Kotlarska 2, Brno, Czech Republic 611 37, CZECH REPUBLIC; and Igor Chalás, Masaryk University, Dept of Computer Graphics and Design, Faculty of Informatics, Brno, South Moravian Region 60300, CZECH REPUBLIC*

After attending this presentation, attendees will understand the basic principles on which mesh-to-mesh and deep learning algorithms for 3D face recognition are grounded. Attendees will also be more familiar with the performance of 3D-to-3D and 2D-to-3D matching algorithms as tested on expression-variant 3D recordings of human faces.

This presentation will impact the forensic science community by providing insight into current state-of-the-art forensic facial identification and by introducing approaches designed or applicable for processing of 3D faces.

The role of images (Identification (ID) photographs or surveillance videos) in personal identification is sometimes downplayed because image-based forensic evidence is, on many occasions, presented for examination in very poor quality, with non-standard framing, viewpoints, or under unspecified conditions. Furthermore, many images are unsuitable for automatic or semi-automatic face recognition systems due to non-standard poses, occlusion, or facial expressions. Development of 3D sensors, which has allowed recording depth information of human faces, has had a great impact on improving the robustness of facial recognition algorithms. Installations of 3D video surveillance systems and 3D capturing devices built for outdoor use have increased chances of securing 3D images at crime scenes and/or in the course of forensic investigations. 3D technology has been shown to compensate for many of the shortcomings traditionally associated with conventional 2D images and to improve performances of matching systems in personal identification if standard faces are processed; however, this technology becomes less successful if dynamic body features, such as hairstyle, facial hair, and, most importantly, facial expressions are incorporated.

Following previous studies, which indicated that matching 3D records of human faces outperform existing systems based on 2D images if tested on controlled expression-invariant faces, the present study sought to extend the objectives by exploring performance of matching algorithms using non-standard expression-variant 3D faces. The tested dataset was composed of an array of 3D faces with nine different facial expressions, including one with a neutral appearance, collected from 150 participants (for a total of 1,350 scans). For each individual, the array was acquired within one scanning session using the VECTRA® XT 3D imaging system.

Two matching algorithms varying in complexity and computation requirements were tested. The first falls under the rubric of 3D-to-3D image matching and was based on extracting 3D mesh geometry descriptors. The second applied 3D-to-2D conversion prior to computation and was based on a learning algorithm featuring convolution networks. For mesh-based matching, algorithms incorporated into FIDENTIS Analyst, a software application developed specifically for processing 3D faces, were tested. The program utilizes the Iterative Closest Point (ICP) algorithm for one-to-one, one-to-all, and one-to-average registration of 3D faces. Dissimilarity of aligned 3D faces is subsequently expressed via signed and absolute closest vertex-to-vertex distances. One scan per individual was selected randomly and tested against the FIDENTIS 3D Face Database (N~2,100 subjects; www.fidentis.cz). This was carried out repeatedly on the entire tested subset. A match/non-match decision was subsequently made while employing a classification model provided by linear discriminant analysis.

For the second approach, a deeper learning algorithm incorporated in CAFFE framework (caffe.berkeleyvision.org) was adapted. Prior to processing, 3D textured meshes were converted automatically into 2D color images (resolution 256 by 256 pixels) accompanied by corresponding depth maps using a simple in-house converter (Face viewer). In order to train the system, the dataset was split into training (75%) and testing (25%) subsets. To classify the images as matching/non-matching, the Berkeley Vision and Learning Center (BVLC) Reference CaffeNet classification model was employed. In both instances, performance of the tested algorithms was assessed by Receiver Operating Characteristics (ROC) curves (cross-validated by repeated random sub-sampling) and expressed in terms of Area Under the Curve (AUC), likelihood ,and odds ratios.

*Presenting Author

The results showed that both tested approaches were challenged by the presence of facial expressions in the tested data and their performance was poorer than when tested on 3D faces with neutral expressions; however, the deep learning approach proved that it could be very efficient if trained on a very large training dataset and, in a more user-friendly version, could be of help for forensic experts. Both tested algorithms were also shown to possess high demands on hardware and computation time. The work which deals with non-standard 3D images and departs from the recent trends of employing advanced computation and learning techniques may improve methods of facial recognition used in the forensic settings. The main purpose of this presentation is to present these prospects to the forensic community.

3D Face Recognition, Face Expression, Matching Algorithms