Object features and face detection performance: Analyses with 3D-rendered synthetic data

Jian Han, Sezer Karaoglu, Hoang-An Le, and Theo Gevers

Computer Vision Lab,University of Amsterdam

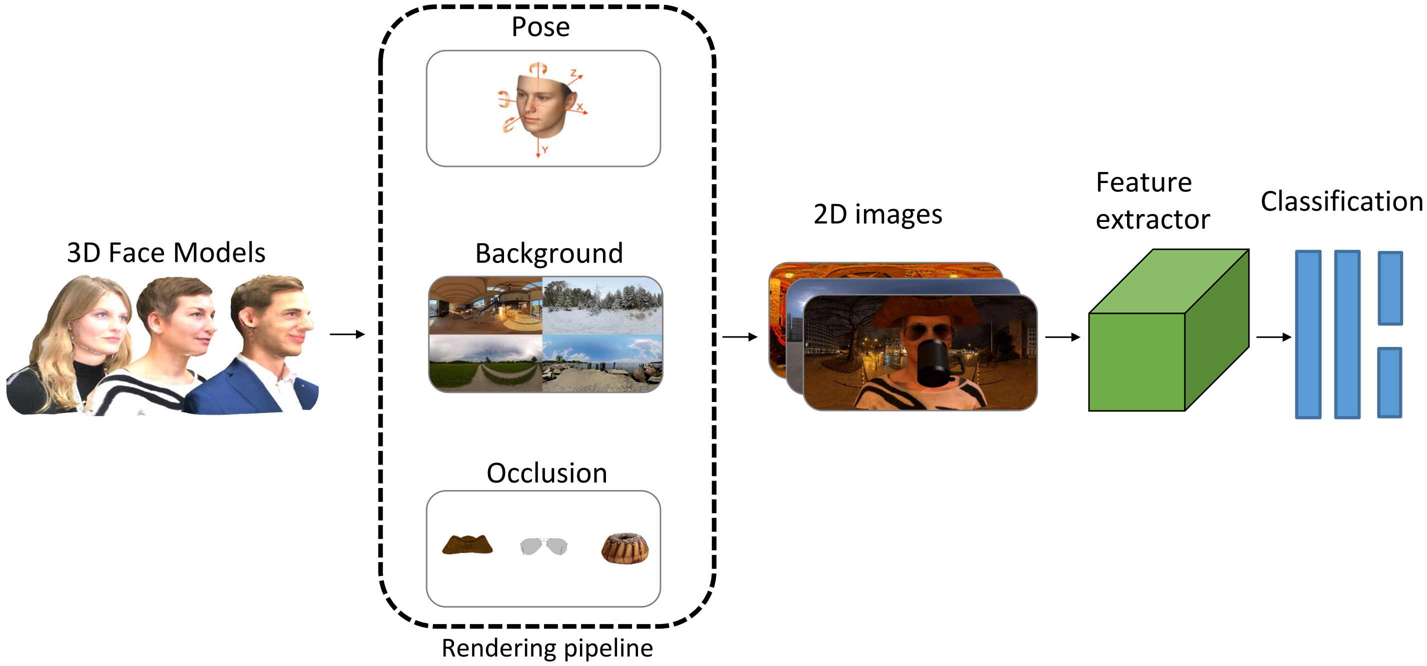
Introduction

- Face detection confronts challenges from features in scale, head pose, expression, facial occlusion and illumination.
- Existing dataset for face detection, the majority of data normally belong to limited range of variations. The faces did not sufficiently to train a robust detector against all potential variations.
- Collecting and annotating real-world datasets with different attributes is unpractical. It is also difficult to fully control the imaging variations in such datasets, or to avoid errors during annotation process. A bias from ground truth may lead to far-reaching impact in deep network.
- We aim to address the issue by using synthetic data, as complementary to real data. We develop a synthetic data generator based on 3D face models. The 2D face synthesis process contains varied viewpoint, scale, illumination, occlusion, and background.

Problem formalization

- Challenging Object features (pose, scale, facial occlusion, blur)
- Face detectors. Faster RCNN is the most representative object detector based on deep network. Single Stage Headless detector is an extremely fast one-step face detector, designed to be scale invariant. Hybrid Resolutions face detector has good performance on tiny face detection by using wide-range contextual information and testing on multiple resolutions.
- Face dataset (MAFA, UFDD, Wider Face). MAFA is a representative dataset of facial occlusion, which is mainly composed of various level of occlusions. UFDD contains faces in different weather conditions and other challenging features concerning lens impediments, motion blur and defocus blur.
- Face synthesis. The rendering pipeline is built on Blender. The parameters of pitch, yaw and roll are selected randomly. for face detection is generated from 3D landmarks. Our rendering pipeline can also be applied on other 3D face models.

Overview



An overview of render pipeline. We manipulate pose, background, occlusion, and illumination on original 3D models, and then render 3D models into 2D images.

Experiments

- Evaluation on object features (scale, head pose, facial occlusion and noise).
- Improve performance through synthetic data. We train on a combination of Wider Face and synthetic data and then test on other dataset (MAFA, UFDD, Wider Face).
- The advantage of synthetic data is that the variations in dataset can be fully controlled in different practical situations.
- False postives have two major sources. The first source is annotation. Different datasets have their own annotation policy. This deviation may turn our reasonable predictions into false positives. The second source of false positives is misleading objects, such as round-shaped objects and human body parts.

Conclusion

- We provide a 2D face synthetic data generator with manipulated features (on pose, scale, background, illumination, and occlusion), which enables specified examinations of face detector performances.
- We conducted detailed analyses between feature and performance, which can be a guide to compare performances of other face detectors.
- Our analyses also reveal some weaknesses of the current face detectors and suggests using synthetic data for future improvement on robustness.