Face Segmentation and Parsing in OFIQ

09.11.2023

Johannes Merkle

secunet Security Networks AG

https://de.wikipedia.org/wiki/StyleGAN

OFIQ

- For face recognition, the facial images should meet various requirements
  - Capture related: illumination, exposure, sharpness, background, …
  - Subject-related: Frontal pose, neutral expression, no occlusions, …
- ISO/IEC 29794-5:2024 will define quality requirements and assessment algorithms

  → Talk by Patrick Grother

OFIQ

- Open Source Face Image Quality (OFIQ)
  - Open source implementation of facial image quality assessment algorithms (FIQA)
  - Reference implementation of ISO/IEC 29794-5:2024
  - Founded by the BSI
  - Usable in commercial and non-commercial applications
  - Covers various quality aspects but also outputs a unified quality score

OFIQ

- Many algorithms in OFIQ require segmentation of face image
- Restrict computation to areas of interest
  - Face, eyes, background, etc.
- Check presence of certain objects in image or certain image regions
  - E.g. hair, hands or objects in face region (occlusion)

Approaches for Face Segmentation

- Segmentation of face region and face parts can be done by facial landmarks
  - E.g. eyes, mouth
- Segmentation by CNNs
  - Many publications and free implementations
  - Trained on labelled datasets (manual labels or synthetic images)

Face Segmentation CNNs

- Type of classes distinguished depend on training set
  - E.g. glasses, neck, beard/moustache, head coverings, occlusions
  - Definitions of “occlusion” differ, e.g. for transparent glasses
  - For some datasets, no trained CNNs are available, e.g. mut1ny, FaceSynthetics

CelebAMask-HQ

FaceOcc

mut1ny face/head segmentation dataset

LaPa

FaceSynthetics

Face Segmentation Methods in OFIQ

- Landmarked region segmentation
- Face parsing
- Occlusion-aware face segmentation

Landmarked region segmentation

- Face region without forehead
- Convex hull of face landmarks
- Mouth is also segmented by landmarks
- Landmarks are computed using ADNet
  - https://github.com/huangyangyu/ADNet
  - MIT license
  - Trained on WFLW (98 landmarks)
  - Very accurate, even for challenging images

Face Parsing

- https://github.com/VisionSystemsInc/face-parsing.PyTorch
  - MIT license
- Computed on aligned images of size 400x400
  - Takes 65 ms on Intel Core i9-10900X @ 3.70GHz (CPU)
- Distinguished classes:
  - Face parts: face, eyes, eyebrows, nose, lips, ears,
  - Subject parts: hair, neck, head coverings, clothing,
  - Accessories: glasses, earrings, necklace
  - Background

Occlusion-aware Face Segmentation

- https://github.com/face3d0725/FaceExtraction
  - X. Yin and L. L. Chen: *FaceOcc: A Diverse, High-quality Face Occlusion Dataset for Human Face Extraction*. arXiv abs/2201.08425, 2022.
  - Permission for OFIQ granted by authors
- Computed on aligned images of size 224x224
  - 15 ms on Intel Core i9-10900X @ 3.70GHz (CPU)
- Outputs un-occluded area of face
  - Occlusions include frames of glasses

Usage of Face Segmentation in OFIQ Algorithms

- Background Uniformity
- No Occlusion of the Face / Mouth / Eyes
- Sharpness
- Moment of the Luminance Distribution
- Over-Exposure and Under-Exposure
- Illumination Uniformity, Natural Colors
- No Head Coverings

Background Uniformity

- Get segmentation of background by face parsing
- Erode background mask with kernel of size 4
- Consider only image region above mouth
  - Avoid segmentation errors for clothing

No Occlusion of the Face

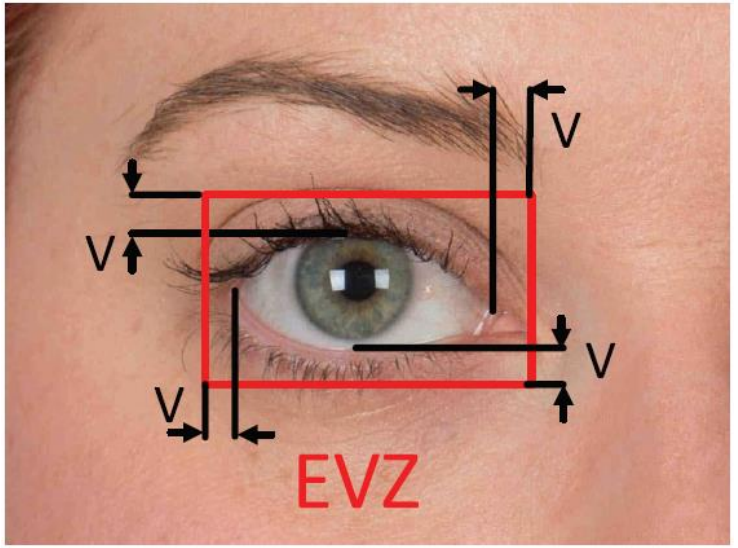
- Get landmarked region segmentation F
- Get occlusion-aware segmentation map O
- Compute degree of occlusion from F and O

No Occlusion of the Mouth

- Segment mouth region M as convex hull of lips landmarks
- Get occlusion-aware segmentation map O
- Compute degree of occlusion from M and O

Eyes Visible

- Determine Eyes Visibility Zone (EVZ) from landmarks
  - Defined in ISO/IEC 39794-5
- Get occlusion-aware segmentation map O
- Compute degree of occlusion from EVZ and O



Sharpness and Moment of the Luminance Distribution

- Landmarked region segmentation

Over-Exposure and Under-Exposure

- Landmarked region segmentation
- Disregard nostrils and beards
  - Intersection with circle centred at eye midpoint with radius to nose tip
- Avoid dark hair and frames of eyeglasses
  - Intersection with occlusion-aware face segmentation

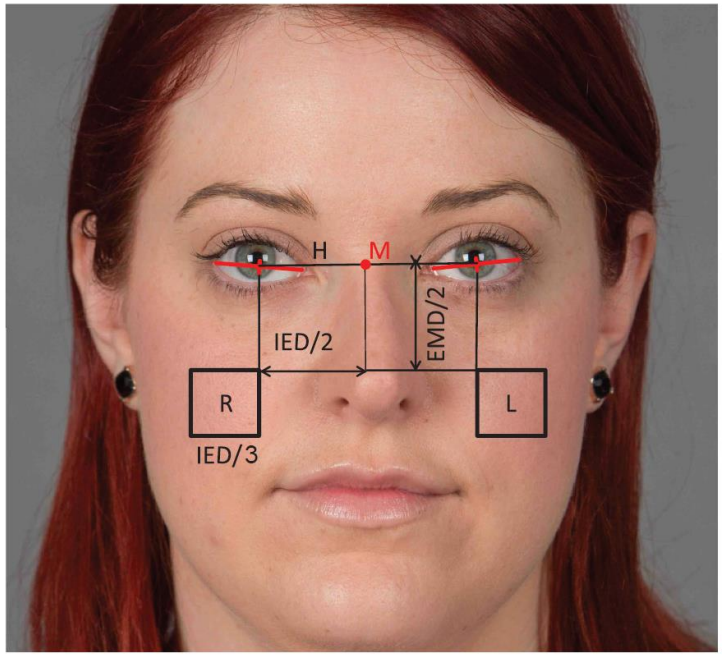
Only necessary for under-exposure



Parts of eyes are included
→ dark irises could be an issue

Illumination Uniformity and Natural Colors

- Segment measurement zones L and M
  - Defined in ISO/IEC 39794-5



No Head Coverings

- Get segmentation map by face parsing
- Restrict map to region above eyes
- Consider pixels with label “head covering” or “clothing”

  → hoods are sometimes labelled as clothes

Summary and Outlook

- Segmentation algorithms work well in most cases
- Two CNNs are deployed
- Short processing times

- Future improvements:
  - Improvement of segmentation for Under- and Over-Exposure
  - Single segmentation CNN for face parsing **and** occlusion-aware segmentation
  - Segmentation of beards/mustaches and nostrils

secunet