

The technology of FaceVACS

FaceVACS processes each image (8 bit greyscale) as follows:

- Face localization: The image is analyzed to determine the position and size of one or more faces. (In all of the following steps, we assume that only a single face is found.)
- Eye localization: The positions of the eye centers in the face are determined.
- Image Quality Check: The quality of the face image is checked to see whether it is sufficient for the steps that follow.
- Normalization: The face is extracted from the image and is scaled and rotated such that an image of fixed size results, with the eye centers at fixed positions in that image.
- Preprocessing: The normalized image is preprocessed with standard techniques such as histogram equalization, intensity normalization, and others.
- Feature extraction: In the preprocessed image, features are extracted that are relevant for distinguishing one person from another.
- Construction of the reference set: During enrollment the facial features of (usually) several images of a person are extracted and combined into a reference set, also called the "biometric template".
- Comparison: For verification, the set of extracted features is compared with the reference set of the person who the person in the image just processed is claimed to be; for identification, the feature set is compared to all stored reference sets, and the person with the largest comparison value is selected; in both cases the recognition is considered successful if the (largest) comparison value - which is interpreted as a similarity value - exceeds a certain threshold value.

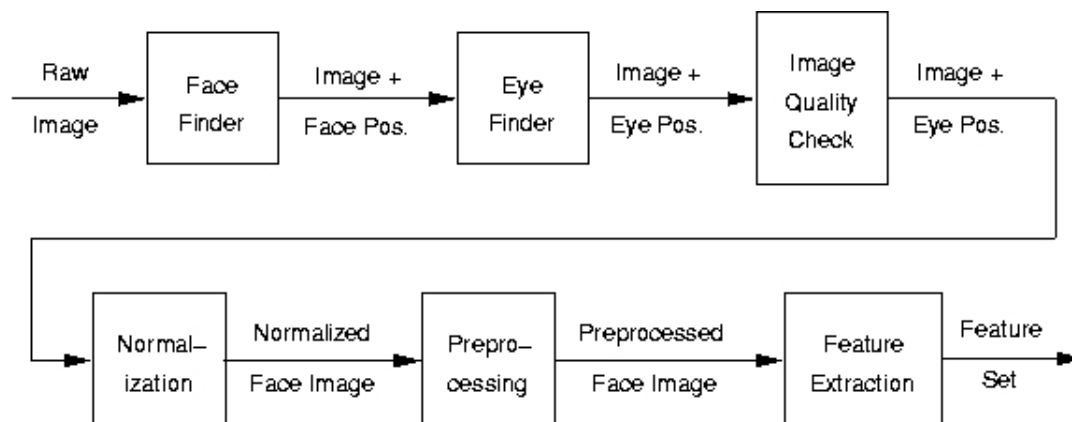


Figure 1. FaceVACS architecture: Feature set creation

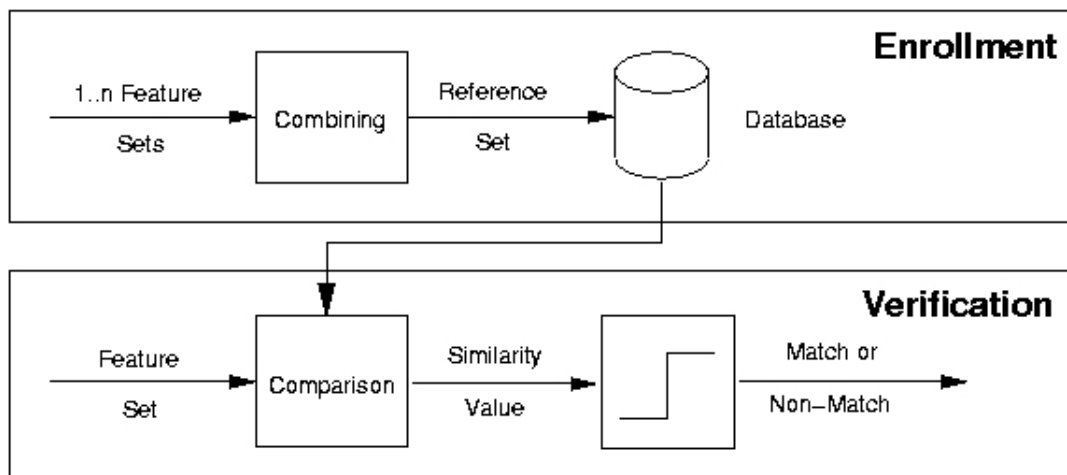


Figure 2. FaceVACS architecture: enrollment and verification

In addition, FaceVACS has a "live check" to ensure that the face in front of the camera is a real one and not just a photograph. To this end, the changes in appearance occurring during movement of the face

(rotations around the vertical axis in particular) are exploited. Due to the special 3D structure of a real face, those changes are quite different for a real face than for a photo. So when the user wants to pass the live check, he or she should briefly rotate his or her head back and forth.

In the following subsections, more details of the individual steps are given. An example image is used to illustrate the effect of each processing stage.



Figure 3. Example image

1. Face and eye localization

To locate the face, a so-called image pyramid is formed from the original image. An image pyramid is a set of copies of the original image at different scales, thus representing a set of different resolutions. A mask is moved pixelwise over each image in the pyramid, and at each position the image section under the mask is passed to a function that assesses the similarity of the image section to a face. If the similarity value is high enough, the presence of a face at that position and resolution is assumed. From that position and resolution, the position and size of the face in the original image can be calculated.

From the position of the face, a first estimate of the eye positions can be derived. In a neighborhood around these estimated positions, a search for the exact eye positions is started. This search is very similar to the search for the face position, the main difference being that the resolution of the images in the pyramid is higher than the resolution at which the face was found before. The positions yielding the highest similarity values are taken as final estimates of the eye positions.

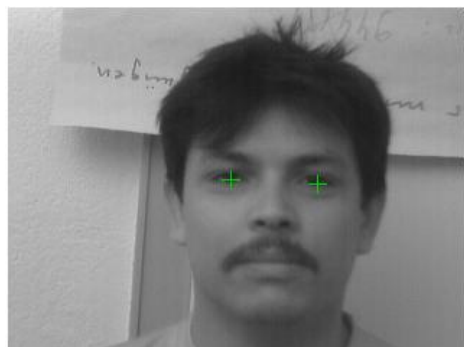


Figure 4. Eye locations found by the algorithm

2. Image Quality Check

To be usable for the subsequent steps, the part of the image occupied by the face has to meet certain quality requirements; e.g., it should not be too noisy or blurred. The quality is measured by means of a set of functions that are applied to the image. If the quality is considered too low, the image is rejected.

3. Normalization and preprocessing

In the normalization step, the face is extracted, rotated and scaled such that the centers of the eyes lie at predefined positions. More precisely, they are positioned to lie on the same horizontal pixel row such that the midpoint of this row is aligned with the midpoint between the centers of the eyes.

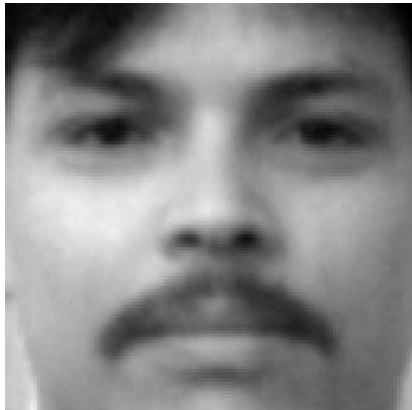


Figure 5. After normalization

The preprocessing step comprises, among other transformations, the elimination of very high and very low spatial frequencies and the normalization of contrast.

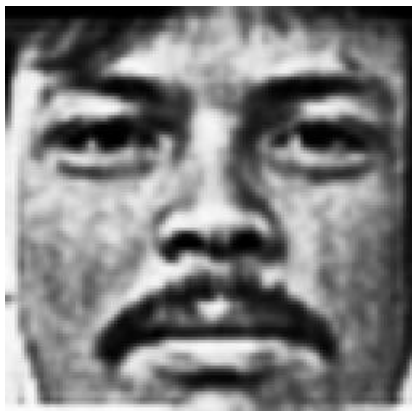


Figure 6. After preprocessing

4. Feature extraction and reference set creation and comparison

Feature extraction starts with local image transforms that are applied at fixed image locations. These transforms capture local information relevant for distinguishing people, e.g. the amplitudes at certain spatial frequencies in a local area. The results are collected in a vector.

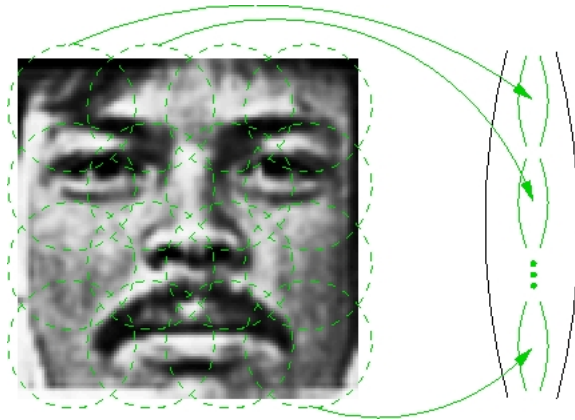


Figure 7. Extracting local features

A global transform is then applied to this vector. Using a large face-image database, the parameters of this transform were chosen to maximize the ratio of the inter-person variance to the intra-person variance in the space of the transformed vectors; i.e., the distances between vectors corresponding to images of different persons should be large compared to distances between vectors corresponding to images of the same person. The result of this transform is another vector that represents the feature set of the processed face image.

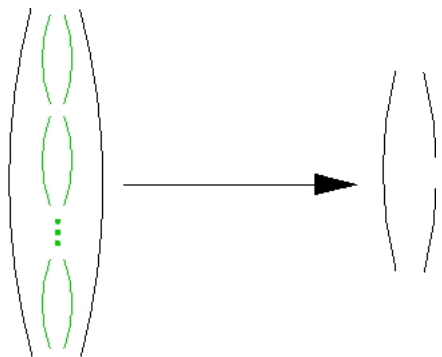


Figure 8. Global transform, yielding the feature set of the face image

For the creation of the reference set, several images are usually taken of each person during enrollment in order to better cover the range of possible appearances of that person's face. The reference set generated for a person consists of up to five feature sets, which are the centers of clusters obtained through a clustering process on the feature sets created from those images.

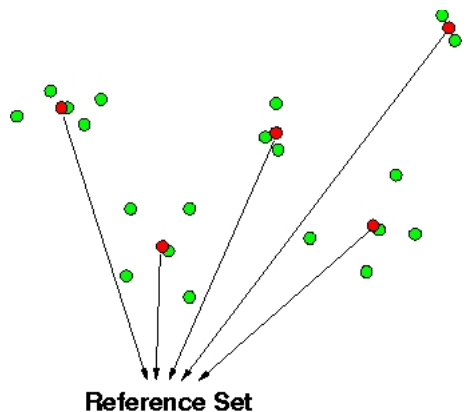


Figure 9. Combining cluster centers (red) into a reference set. (Green dots are feature sets created from images.)

The function that is used to compare a feature set with a reference set is simple and can be computed very fast. It makes identification a matter of seconds, even if a million reference sets have to be compared.