Occlusion Detection for ICAO Compliant Facial Photographs

Markus Storer, Martin Urschler and Horst Bischof
Institute for Computer Graphics and Vision, Graz University of Technology
Inffeldgasse 16/II, 8010 Graz, Austria
{storer,urschler,bischof}@icg.tugraz.at

Abstract

Facial image analysis is an important computer vision topic as a first step for biometric applications like face recognition/verification. The ICAO specification defines criteria to assess suitability of facial images for later use in such tasks. This standard prohibits photographs showing occlusions, thus there is the need to detect occluded images automatically. In this work we present a novel algorithm for occlusion detection and evaluate its performance on several databases. First, we use the publicly available AR faces database which contains many occluded face image samples. We show a straight-forward algorithm based on color space techniques which gives a very high performance on this database. We conclude that the AR faces database is too simple to evaluate occlusions and propose our own, more complex database, which includes, e.g., hands or arbitrary objects covering the face. Finally we extend our first algorithm by an Active Shape Model in combination with a PCA reconstruction verification. We show how our novel occlusion detection algorithm outperforms the simple approach on our more complex database.

1. Introduction

Analysis of facial images in a biometric context is an important research topic in computer vision with face recognition being one especially prominent area of application. There is a high interest in biometrics due to a large number of potential commercial and law enforcement applications (e.g., biometric authentication or surveillance) requiring highly accurate recognition of biometric features like the face. Besides accuracy, robustness to occlusions is another very important aspect of facial image analysis and recognition. The topic of face recognition has received significant research attention over the last two decades [25] and has also led to many commercial systems [17].

We are specifically interested in analyzing facial portrait and near-portrait images in the context of the International Civil Aviation Organization (ICAO) standard [9] for machine readable travel documents (MRTDs). The main intention of this standard is to define how images of arbitrary people have to look like in order to perform robust and highly accurate face recognition/verification. Studies show that without proper registration of facial images face recognition performance degrades significantly [18]. Therefore, a part of the ICAO specification describes a standardized coordinate frame based on eye locations for the purpose of geometrical alignment. Starting with this standardized coordinate frame, which is computed in a so-called *canonization* step, one can derive criteria to define images with and without occlusions. In this work we concentrate on occlusions due to extraordinary glasses and objects covering parts of the face (hands, hair, or other objects). See Figure 3 and Figure 5 for some examples.

The problem of occlusions in the context of face recognition has recently been studied in [7], where the authors show different amounts of degradation in recognition performance depending on the location of the facial occlusion. A number of techniques have emerged to make the recognition algorithm itself robust to occlusions. Early work in this direction has been proposed by Leonardis and Bischof [10] who showed how to handle occlusions in an eigenface [21] framework. Their key idea was to extract eigenspace coefficients by a robust hypothesize-and-test paradigm using subsets of image points instead of computing the coefficients by projecting the data onto the eigenimages. Li et al. [11] presented a local non-negative matrix factorization (LNMF) to learn spatially localized part based subspace representations from visual patterns. Their use of localization constraints showed good performance on the AR face database. Extending this work, Oh et al. [16] proposed a selective LNMF technique with a partial occlusion detection step on a number of disjoint image patches. These patches are represented by a PCA to obtain corresponding occlusion-free patches, followed by the LNMF procedure used exclusively on the bases of the occlusion-free image patches. A different direction was pursued by Martinez [14] who described a probabilistic approach that compensates for imprecisely localized, partially occluded faces under different facial ex-

978-1-4244-7030-3/10/$26.00 ©2010 IEEE

pressions. He divides the face into a number of local regions and matches them to a single prototype by a probabilistic scheme. He demonstrates robustness in the presence of occlusion of 1/6 to 1/3 of the facial area at the cost of only a slight decrease in accuracy. All of these presented approaches have in common that their goal is to perform face recognition in the presence of occlusions. However, in our task we explicitly want to detect occlusions to sort out unsuitable images for a subsequent recognition step. This is in accordance with the ICAO specification which prohibits occluded facial images.

A tightly related topic is face hallucination which was made popular in work by Baker and Kanade [1]. Here occluded parts of a face are recovered by using generative face models. Different terms describing this area of research are face recovery and regeneration or face image inpainting. Some examples of recent work are presented in [12, 23]. Our face images are used for machine-readable travel documents, so we do not want to modify given occluded images. Therefore we do not focus further on this research direction.

An obvious choice for an occlusion detection algorithm is the widely used Active Appearance Model (AAM) [5]. Here the strategy is to use the generative AAM model fitting algorithm starting with a suitable model initialization on a face portrait image. By fitting the model to the occluded image one could derive a quality measure (e.g., the final sum-of-squared differences) to make a decision if an occlusion is present or not. Here, the main problem is that the original AAM model formulation is not very robust to occlusions. Some extensions of the AAM model in the presence of occlusions have been presented in the literature [8, 20, 24], however their holistic approach poses a basic difficulty during model fitting, since the quality measure driving the fitting optimization always is influenced by the occluded part to a certain degree and the non-convex optimization is prone to get stuck in local minima. Due to their popularity and widespread availability we will show where this class of algorithms tends to fail in our occlusion detection task.

In this work we propose a novel system to automatically detect occlusions from canonized facial images in Section 2. This is an important pre-processing step for the training of face recognition/verification, but could also be used for the testing step. In Section 2.1 we start with a straight-forward occlusion detection method based on color space techniques and perform occlusion detection experiments on the publicly available AR database [13]. We show why this database is not sufficient to evaluate an algorithm for facial occlusion detection, and we present our own more challenging database which we created specifically for this task. We extend our first algorithm and include an Active Shape Model (ASM) [6] approach followed by a PCA based verification step described in Section 2.2. This second method is able to solve the occlusion detection problem on our own more difficult database. Finally, we discuss and summarize our findings in Section 3.

2. Occlusion Detection

We start with a simple and straight-forward approach for occlusion detection in Section 2.1 which we refer to as *Method 1*. It is based on automatic color correction techniques and on the HSV color space. It turns out that this method is already well suited and sufficient to detect occlusions on the publicly available AR face database [13].

We created our own collection of images which extends the variations exhibited by the AR database in terms of further illumination conditions and types of occlusions. We show that the simple Method 1 does not perform very well on this more challenging database. Hence, we exploit our findings of our color experiments of Method 1 and extend this first method by an Active Shape Model [15] in combination with a projection of facial parts to separate Principal Component (PCA) subspaces. We refer to this extension as *Method 2* explained in detail in Section 2.2.

In our proposed system we make use of input images in the canonized coordinate frame according to the ICAO specification. In order to be able to analyze arbitrary facial images we have to transform them first into this coordinate frame based on eye locations. For this purpose we use a robust face and facial component detection stage followed by a probabilistic voting scheme for the most probable face and eye position [22]. The canonized image is finally derived by warping the input image according to the eye locations.

2.1. Method 1

Our first approach is based on automatic color correction and on the H-channel of the HSV color space. Before transforming the image into the HSV color space an automatic color correction is applied. It reduces the effects of global illumination and could also be referred to as automatic white balancing based on color temperatures.

Our automatic color correction algorithm assumes that the average surface color in a scene is gray. This means that the shift from gray of the measured averages on the three channels corresponds to the color of the illuminant. Three scaling coefficients, one for each color channel, are therefore set to compensate this shift [2, 3]. Every RGB-pixel value is adjusted according to

$$\begin{pmatrix} r_{adj} \\ g_{adj} \\ b_{adj} \end{pmatrix} = \begin{pmatrix} \overline{Y}/\overline{R} & 0 & 0 \\ 0 & \overline{Y}/\overline{G} & 0 \\ 0 & 0 & \overline{Y}/\overline{B} \end{pmatrix} \begin{pmatrix} r \\ g \\ b \end{pmatrix}, \quad (1)$$

where $\overline{Y}$ is the mean value of the luminance channel and $\overline{R}$, $\overline{G}$ and $\overline{B}$ correspond to the mean values of the three planes of an RGB-image.

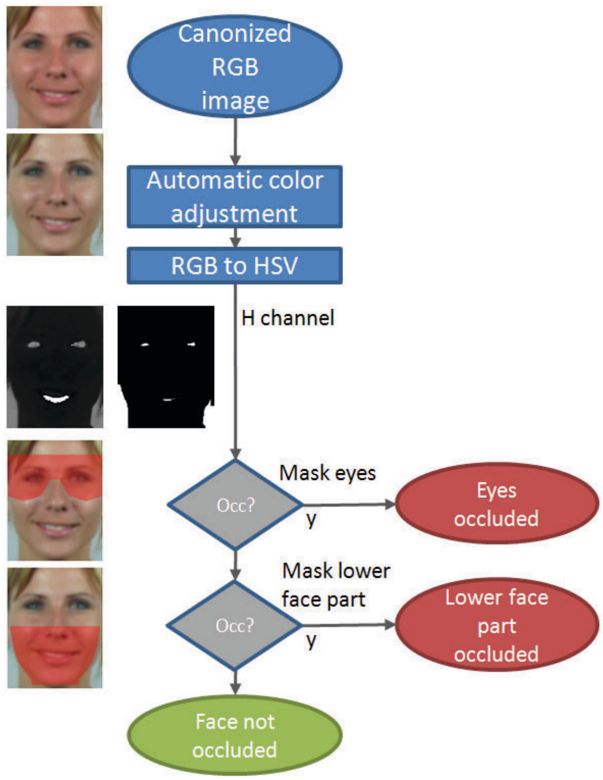


Figure 1. Method 1. The canonized and color adjusted image is transformed to the HSV color space. After binarization of the H-channel of the HSV color space, occlusion masks are used to calculate the level of occlusion on the lower facial part and around the eyes region.

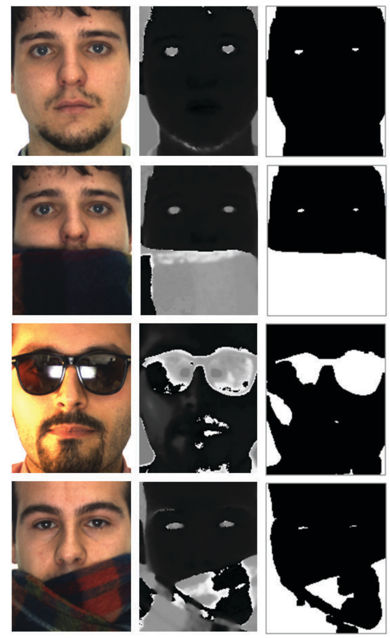


Figure 2. Creation of an occlusion map based on the HSV color space. (first column) original images, (second column) H channel of HSV color space and (third column) the corresponding maps gained after thresholding the H channel image.

Based on several experiments using different color spaces, e.g., RGB, YUV, LAB, XYZ, YCbCr, we found that the H-channel (representing the hue values of the image) of the HSV color space is best suited for occlusion detection on facial images. The H-channel image is binarized and some morphological post-processing is applied to remove small isolated regions. We define masks for the lower facial part and the eyes to obtain a final value for the level of occlusion. This whole chain is illustrated in Figure 1. Figure 2 shows some examples of extracting the H-channel of an image and the final occlusion map after binarization. Note that almost always the beards of male individuals are part of the non-occluded facial region in the binarized H-channel image, thus they do not contribute to an occlusion area.

Experimental results

Using Method 1 we conducted experiments on the publicly available AR face database [13]. The AR face database consists of more than 3000 frontal view facial color images of 135 people showing variations in gender, facial expression, illumination conditions and occlusions (sun glasses and scarves). The size of the images is 768×576 pixels. Those individuals wearing a scarf or sun glasses are labeled as occluded. Some representative examples are presented in Figure 3. With Method 1 we reached an equal error rate (EER) of 4.5%. The corresponding ROC curve is shown in Figure 4. Note that these results are slightly worse than the results presented in Oh et al. [16], however, their method explicitly trains on the occlusions of the AR face database, which is rather unrealistic, since real occlusions occur in a significantly larger variety, while our method works completely unsupervised. Given the very restricted set of possible occlusions present in the AR face database, we conclude that one needs a database with more variation in order to assess the occlusion detection performance realistically.

2.2. Method 2

Method 1 discussed in Section 2.1 is already well suited to detect occlusions on the publicly available AR database. We created a more challenging database of occluded and non-occluded facial images (see Figure 5) where Method 1 does not perform very well. The main reason is that the occlusion detection is exclusively based on color space techniques. In our database we also have skin colored facial occlusions, e.g., hands occluding the face. Examples are depicted in Figure 5a and Figure 5f where Method 1 would fail.

We start by improving the color detection branch as shown in Figure 6. We transform our lowpass filtered can-

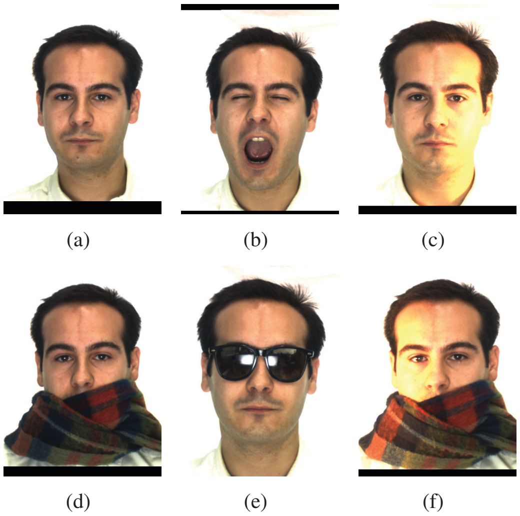


Figure 3. Samples from the AR database. The images feature frontal view faces with (a)-(b) different facial expressions, (c) several illumination conditions, (d) occlusion of the lower facial part by a scarf, (e) occlusion of the eyes by sunglasses and (f) combinations of these variations.

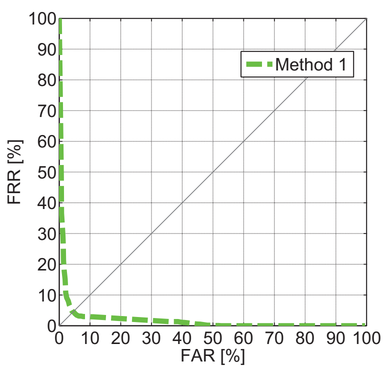


Figure 4. ROC curve of Method 1 evaluated on the AR database.

onized image to the HSV color space and extract the H-channel. After binarization and some morphological operations (as used for Method 1) we obtain the occlusion map. We define an occlusion mask for the forehead and the lower facial part. The eyes region will be considered as the method proceeds. The forehead mask is used to calculate the level of occlusion of the forehead. This is especially important if e.g., somebody wears a cap (Figure 5b).

If the approach at that stage claims an occlusion of the lower facial part, we validate this claim by finding similar colors based on a given color mask, Figure 7. The color mask C is defined based on the position of the canonized image and is marked by the red lines in Figure 7a. We pick up the H-pixel values $h_j \in C$ and form a unimodal Gaussian model $\mathcal{N}(\mu, \sigma^2)$. It turns out that using only the H-values for constructing the Gaussian model is superior to a multivariate Gaussian model constructed from several color queues. Using the Gaussian model we compute the

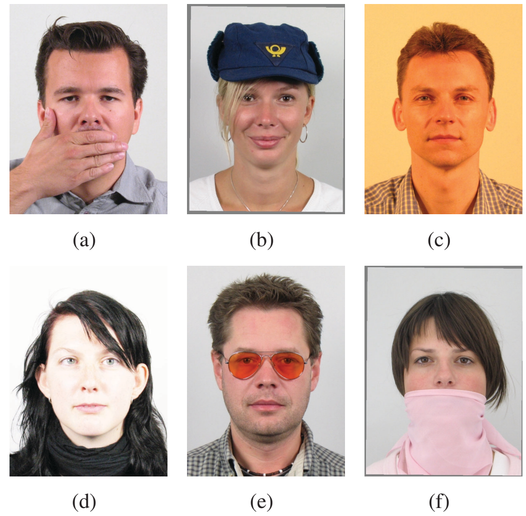


Figure 5. Samples from our own database. In addition to the variations exhibited by the AR database, our own database shows some more variations, e.g. (a) occlusions by skin-similar color of the lower facial part, (b) occlusions of the forehead, (c) variation of the color tone of the overall image, (d) extreme lighting conditions, (e) tinted glasses in several colors and (f) several colored occlusions of the lower facial part (also skin-similar color).

probability map (Figure 7b) in the facial image domain Ω. Therefore, we calculate the probability p_i, $i \in \Omega$, of every pixel ν_i to determine how similar it is to the marked pixels in Figure 7a:

$$p_i = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{\nu_i - \mu}{\sigma}\right)^2\right). \qquad (2)$$

After binarization and some morphological operations we get the final facial map (Figure 7c).

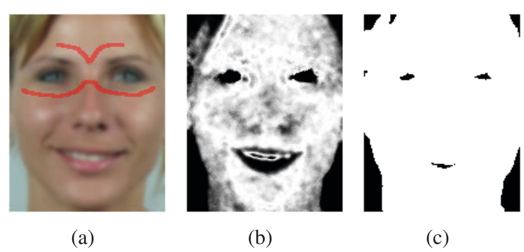


Figure 7. Determining similar colors. (a) The H-channel pixel values marked by the red lines are used to construct a Gaussian model. (b) Probability map of similar colors, (c) probability map after binarization and some morphological operations.

At the end of the color occlusion detection chain, the images showing occlusions with colors similar to skin are still classified as non-occluded. Hence, we create a second occlusion detection step based on fitting an Active Shape Model (ASM) [6], which should also detect non-facial structures. Here we use the recently proposed STASM [15]

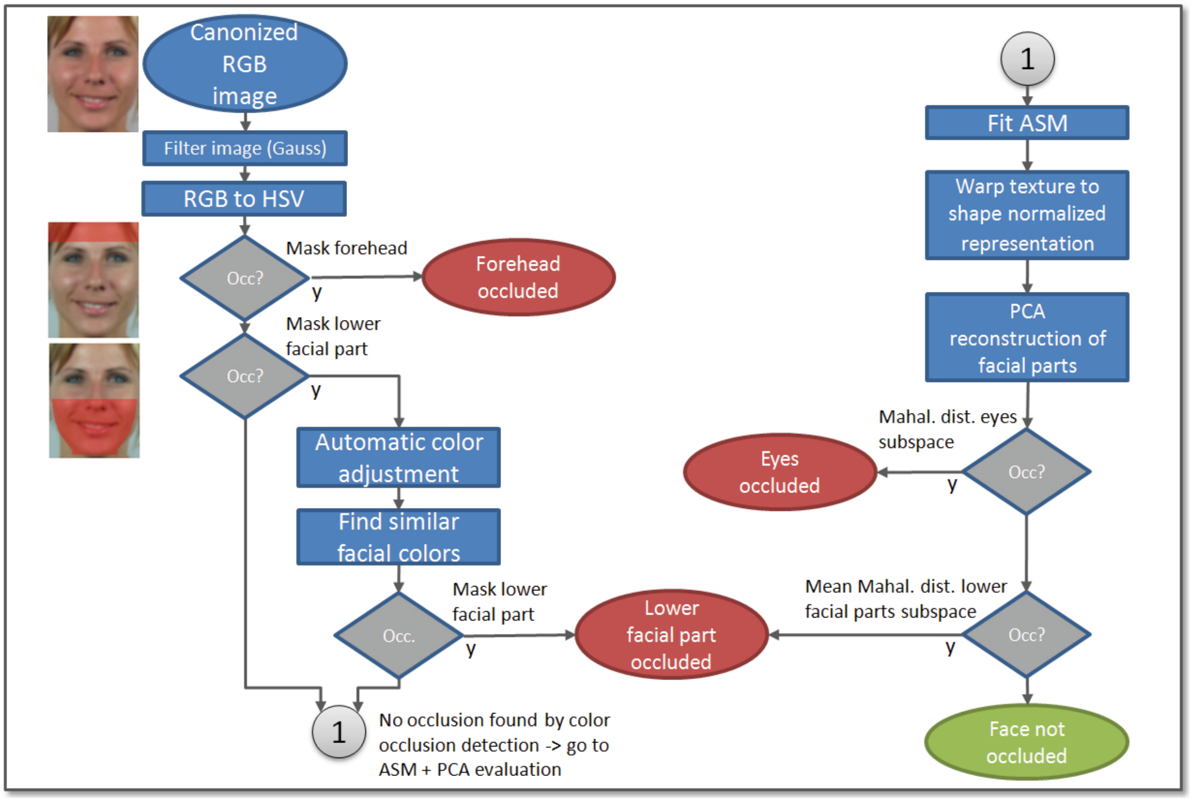


Figure 6. Method 2. The left branch of the whole approach is based on color techniques. If there is no occlusion found by this color occlusion detection, the ASM + PCA approach will be activated.

algorithm which is very robust against varying illumination conditions or partly occluded facial images and showed excellent performance on our data. We also performed experiments with Active Appearance Models [5, 19] but they are by far inferior in terms of fitting accuracy for our task compared to STASM, see Figure 8.

The STASM Algorithm

This publicly available algorithm extends the original Active Shape Model [6] by a number of techniques like two- instead of one-dimensional landmark profiles, extending the set of training landmarks and trimming the covariance matrix by setting a large number of entries to zero. In the following we will describe this algorithm which is an important part of our system.

The original ASM makes use of a statistical formulation to combine a set of user-specified landmark points in a training set of annotated images into a generative model of the object of interest. This generative model describes the variation of the object shape from a mean object instance. The ASM relies on a specified ordering of the n landmarks $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ in a training image. Given K suitably aligned training images we can generate K vectors $\mathbf{x}_k$. These vectors

$$\mathbf{x}_k = (x_1, \ldots, x_n, y_1, \ldots, y_n)_k^T$$

form a distribution in a $2n$ dimensional space, and the aim of the ASM is to generatively model this distribution. Therefore, a Principal Component Analysis (PCA) is applied to the training data which results in the mean and the main axes with their corresponding variances of the cloud of points in the high-dimensional space. An approximation of any training instance $\mathbf{x}$ can be calculated from

$$\mathbf{x} \approx \overline{\mathbf{x}} + \mathbf{P}\mathbf{b}, \tag{3}$$

where $\overline{\mathbf{x}}$ is the mean of the distribution, $\mathbf{P}$ is the matrix formed by the t eigenvectors of the covariance matrix of the points and $\mathbf{b}$ is a t-dimensional vector of weights which resembles a set of parameters of a deformable shape model. Modifying this parameter creates different shapes restricted by the information from the training data. In Cootes et al. [6] the fitting of the shape model is performed by an iterative algorithm that finds global pose as well as model parameters $\mathbf{b}$. The fitting procedure is based on the matching of a one-dimensional profile of gray-value and edge information, derived from the training data, to profiles extracted from the current position of the landmarks in the test image. This procedure is very sensitive to the initialization of global pose and model parameters and is prone to get stuck in local minima. The fitting process can be understood as iteratively moving landmark points independently from each other to locations where the profile match is a better one and regularizing the locations of all landmark points by the global PCA shape model.

Figure 8. Comparison of the fitting quality of model based fitting. (first row) AAM, (second row) STASM.

STASM [15] advances this basic model by a number of important extensions. First, they increase the number of necessary landmark points to add redundancy to the model representation and they perturb the landmarks of the training data set by random noise to increase the number of available training data. Second, instead of one-dimensional profiles they use two-dimensional patches at the landmarks. This increases the matching performance at the cost of slightly more computational work. Third, during iterative fitting the global shape model is used with an increasing amount of variation. This means that at lower levels in the image pyramid a small variance around the mean shape is allowed and a more restricted set of eigenvalues is chosen for shape regularization. As fitting proceeds and we reach the original level of the image pyramid eigenvectors are added and the maximal variance is increased in order to loosen the regularization constraints imposed by the shape model. Finally, the patch profile covariance matrix used for matching is optimized by setting components resembling distant points to zero and trimming the resulting approximated covariance in order to be positive definite again. The main purpose of this step is to reduce matching time.

Combining STASM with a PCA sub-component model

We manually annotated n facial images and aligned the obtained facial shapes by applying Procrustes analysis for shape registration. The mean shape is calculated from these aligned shapes. We warp each annotated image to the mean shape representation and split every warped image into three parts, namely the left and right part of the lower face and the eyes region, see Figure 9. We construct a separate color Principal Component Analysis (PCA) subspace $\mathbf{U}_k = \left[\mathbf{u}_{k_1}, \ldots, \mathbf{u}_{k_p}\right]$, $k \in [1, 2, 3]$ for every part. Usually only p, $p < n$, eigenvectors $\mathbf{u}$ are sufficient.

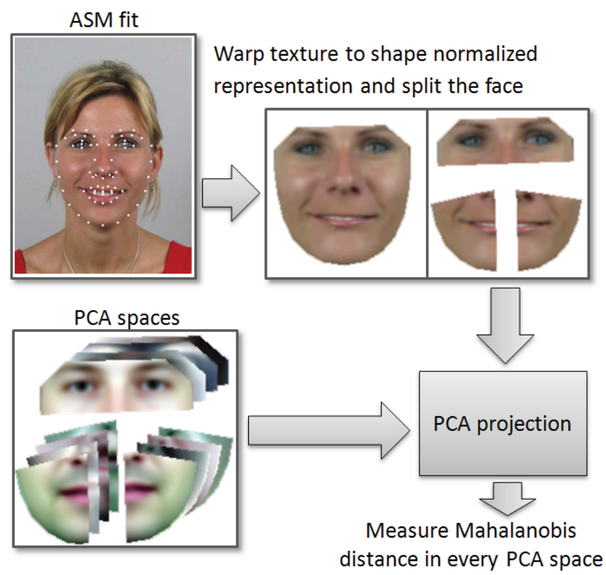


Figure 9. ASM + PCA for occlusion detection.

In the occlusion detection with this ASM + PCA approach we first fit the ASM to the input image. We warp the texture, enclosed in the found landmarks, to the mean shape representation. This texture is split as in the training stage of the PCA. The advantage of the split is the increased robustness to bad illumination conditions compared to the whole face. Every obtained facial part $\mathbf{t}_k$ is projected into the corresponding subspace obtaining the PCA coefficients $\mathbf{c}_k$ which correspond to distances from the mean on the axes spanned by the subspace:

$$\mathbf{c}_k = \mathbf{U}_k^T \left(\mathbf{t}_k - \overline{\mathbf{t}_k}\right). \tag{4}$$

We measure the Mahalanobis distance d_k in every partial subspace and thus determine which part of the face is occluded:

$$d_k = \sqrt{\mathbf{c}_k^T \mathbf{\Sigma}_k^{-1} \mathbf{c}_k}, \tag{5}$$

where $\mathbf{\Sigma}_k = \mathrm{diag}\left(\lambda_k\right)$ is the diagonal covariance matrix consisting of the eigenvalues λ_k obtained from the construction of the PCAs.

A further nice property of this ASM + PCA approach is that we can measure the fitting quality of the ASM, which is an unresolved question in the literature. If the ASM fit is poor, the warped and projected texture will lead to a large Mahalanobis distance in the subspace, because this texture is not represented in the facial subspace. On the other hand a good ASM fit will result in a good reconstruction of the facial parts. Hence, that combined approach is a good indicator for the fitting quality of the ASM.

Experimental results

We performed experiments on our own database which consists of 4930 color facial images and is more challenging compared to the AR database used in our first experiments. In addition to the variations of the AR database, our database exhibits further illumination conditions and more types of occlusions, see Figure 5. The size of the images is 480×640 pixels.

For our ASM + PCA approach, we manually annotated 427 facial images taken from the Caltech face database [4] and our own collection (disjoint from our test database). Taking also the mirrored versions of those images doubles the amount of data. For the PCA model we keep 98% of the eigenvalue energy spectrum for each of the three subspaces.

Method 2 gives a significant increase in performance compared to Method 1 on our own database. The EER is decreased from 30.9% to 6.6%. The corresponding ROC curves are depicted in Figure 10. In Figure 11 some typical failure cases of our approach are shown. The algorithm is very fast, mostly depending on the runtime of the STASM. The average runtime[1] to analyze a facial image is 0.3s.

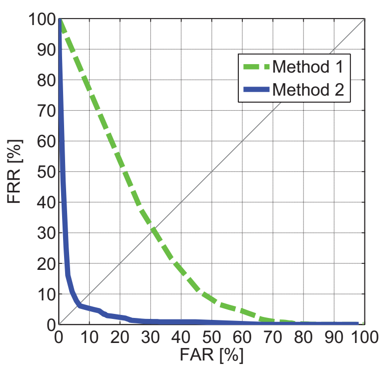
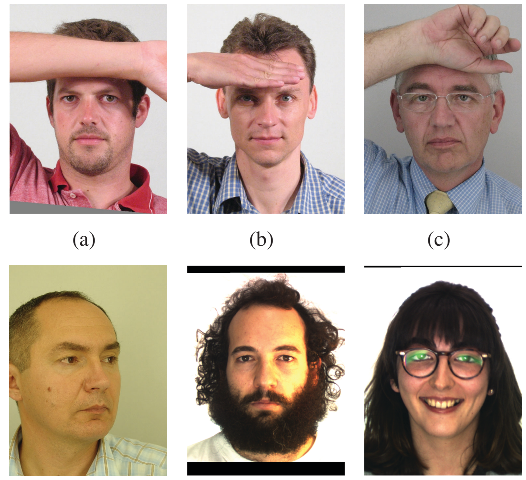


Figure 10. ROC curve of Method 1 and Method 2 evaluated on our own dataset.

[1]The runtime is measured using an Intel Core 2 Duo processor running at 2.4GHz.



(a) (b) (c)

(d) (e) (f)

Figure 11. Typical failure cases resulting from (a)-(c) skin-similar occlusion of the forehead, (d) larger deviations from frontal pose, (e) extreme facial hair and (f) slight specularities exhibited on the glasses.

3. Conclusion

Occlusion detection is an important part of the ICAO specification for assessing suitability of facial images for machine readable travel documents. We presented two approaches which detect occlusions on facial portrait images. The first approach is straightforward and is based on color techniques using the H-channel of the HSV color space. It turns out that this first method is already well suited to sufficiently detect occlusions on the publicly available AR database. We created a more challenging database where the first method showed significant shortcomings. Hence, we improved our first method with a combination of an Active Shape Model (ASM) and a component based PCA subspace reconstruction. This algorithm proved to be very successful on our more difficult database. Furthermore, we can use the combination of ASM and PCA reconstruction as a measure of ASM fitting quality.

Acknowledgements

This work has been funded by the Biometrics Center of Siemens IT Solutions and Services, Siemens Austria.

References

[1] S. Baker and T. Kanade. Hallucinating faces. In *Proc. International Conference on Automatic Face and Gesture Recognition (FG)*, March 2000.

[2] S. Bianco, F. Gasparini, and R. Schettini. Combining strategies for white balance. In *Proc. Digital Photography III, IS&T/SPIE Symposium on Electronic Imaging*, 2007.

[3] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1):1–26, July 1980.

[4] Caltech. Caltech Face Database. http://www.vision.caltech.edu/html-files/archive.html, 1999.

[5] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.

[6] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, 61:38–59, January 1995.

[7] H. K. Ekenel and R. Stiefelhagen. Why is facial occlusion a challenging problem? In *Proc. International Conference on Advances in Biometrics (ICB)*, 2009.

[8] R. Gross, I. Matthews, and S. Baker. Active appearance models with occlusion. *Image and Vision Computing*, 24(6):593–604, 2006.

[9] ISO International Standard ISO/IEC JTC 1/SC37 N506. Biometric data interchange formats - part 5: Face image data, 2004.

[10] A. Leonardis and H. Bischof. Robust recognition using eigenimages. *Computer Vision and Image Understanding*, 78(1):99–118, 2000.

[11] S. Z. Li, X. W. Hou, H. J. Zhang, and Q. S. Cheng. Learning spatially localized, part-based representation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 207–212, 2001.

[12] D. Lin and X. Tang. Quality-driven face occlusion detection and recovery. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.

[13] A. Martinez and R. Benavente. The AR face database. Technical Report 24, CVC, June 1998.

[14] A. M. Martinez. Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample image per class. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(6):748–763, 2002.

[15] S. Milborrow and F. Nicolls. Locating facial features with an extended active shape model. In *Proc. 10th European Conference on Computer Vision (ECCV)*, 2008.

[16] H. J. Oh, K. M. Lee, and S. U. Lee. Occlusion invariant face recognition using selective local non-negative matrix factorization basis images. *Image and Vision Computing*, 26(11):1515–1523, 2008.

[17] P. J. Phillips, W. T. Scruggs, A. J. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott, and M. Sharpe. FRVT 2006 and ICE 2006 large-scale results. Technical report, National Institute of Standards and Technology, Gaithersburg, MD 20899, March 2007.

[18] E. Rentzeperis, A. Stergiou, A. Pnevmatikakis, and L. Polymenakos. Impact of face registration errors on recognition. In *Proc. Artificial Intelligence Applications and Innovations (AIAI)*, pages 187–194, 2006.

[19] M. B. Stegmann. Active appearance models: Theory, extensions and cases. Master's thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, August 2000.

[20] M. Storer, P. M. Roth, M. Urschler, H. Bischof, and J. A. Birchbauer. Active appearance model fitting under occlusion using fast-robust PCA. In *Proc. International Conference on Computer Vision Theory and Applications (VISAPP)*, volume 1, pages 130–137, February 2009.

[21] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[22] M. Urschler, M. Storer, H. Bischof, and J. A. Birchbauer. Robust facial component detection for face alignment applications. In *Proc. 33rd Workshop of the Austrian Association for Pattern Recognition (AAPR/OAGM)*, pages 61–72, May 2009.

[23] D. Yu and T. Sim. Using targeted statistics for face regeneration. In *Proc. International Conference on Automatic Face and Gesture Recognition (FG)*, 2008.

[24] X. Yu, J. Tian, and J. Liu. Active appearance models fitting with occlusion. In *Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*, pages 137–144, 2007.

[25] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, 2003.