

The Effect of Image Resolution on the Performance of a Face Recognition System

B.J. Boom, G.M. Beumer, L.J. Spreeuwers, R. N. J. Veldhuis
Faculty of Electrical Engineering, Mathematics and Computer Science
University of Twente
P.O.Box 217 7500 AE Enschede, The Netherlands
Email: {b.j.boom, g.m.beumer, l.j.spreeuwers, r.n.j.veldhuis}@el.utwente.nl

Abstract—In this paper we investigate the effect of image resolution on the error rates of a face verification system. We do not restrict ourselves to the face recognition algorithm only, but we also consider the face registration. In our face recognition system, the face registration is done by finding landmarks in a face image and subsequent alignment based on these landmarks. To investigate the effect of image resolution we performed experiments where we varied the resolution. We investigate the effect of the resolution on the face recognition part, the registration part and the entire system. This research also confirms that accurate registration is of vital importance to the performance of the face recognition algorithm. The results of our face recognition system are optimal on face images with a resolution of 32×32 pixels.

Keywords—Face Recognition, Face Registration, Image Resolution

I. INTRODUCTION

In video surveillance applications it is often difficult to obtain good quality recordings of the faces of the observed individuals. One important factor determining the quality of the recordings is the resolution of the images, which is often much lower than the resolutions typically used in face recognition. Therefore, we decided to investigate the lowest resolution at which a face recognition system still can achieve acceptable performance. In a face recognition system several processing steps are needed to recognize a person's face. Here, we investigate the sensitivity of the recognition part and the registration part to the image resolution.

Currently available face recognition systems usually require face images with more than 50 pixels between the eyes. In literature several papers can be found which use much lower resolutions. Zhao et al [1] use a combination of PCA and LDA for face recognition on a resolution of 24×21 pixels and claim that their approach will even give good results on 19×17 pixels. Kukharev et al [2] report that the images should be larger than 28×23 pixels using PCA and LDA. Wang et al [3] investigate the effects of resolution on face recognition and conclude that results improve until a resolution of 64×48 pixels and remain constant for higher resolutions using both PCA and a combination of PCA and LDA. In [4], Czyz et al mention that results only slightly decrease using face images of 16×16 pixels. Ekenel et al [5] investigate the frequency subbands that perform best. By first low pass filtering and

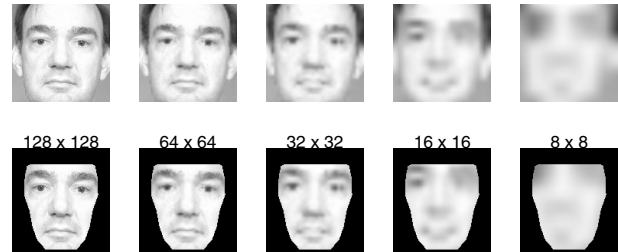


Fig. 1. Different resolutions face images are shown above for the registration and below for recognition

subsequently downsampling, their results of PCA, ICA1 and ICA2 hardly get worse on face images with resolutions down to 16×16 pixels.

Other papers investigate the enhancement of the resolution of images using face hallucinating or super-resolution [6], [7]. This paper might give an indication to which resolution the image must be enhanced for reliable face recognition. In [6] it is suggested that the resolution required to find landmarks should probably be higher than the resolution used for performing face recognition. In [6] it is also noticed that not only face recognition depends on resolution but also the registration, which in many cases relies on finding facial landmarks.

The papers found in literature only mention face recognition as a function of image resolution, while we also investigate the face registration as a function of image resolution. In comparison to other papers we use a larger dataset containing 3699 face image. While in [3] very limited research is done in the choose of the number of PCA/LDA components, we do more elaborate research into this subject.

This paper is organized as follows. In section 2 we explain how we generate face images with different resolutions. In section 3 the methods we use for face registration and recognition are described. In section 4 the experiments and results are given on different resolutions. Finally, in section 5 conclusions are presented.

II. FACE IMAGE RESOLUTION

In this paper we investigate the face registration and recognition as a function of the image resolution. Instead of

taking images from faces which have different distances to the camera, we have tried to simulate the effect of lowering the resolution. By using simple downsampling or taking the mean pixel value, the image will still contain high frequency components. To simulate the effect of the camera lens we used a gaussian low pass filter with $\sigma = 0.375$, followed by downsampling the image with a decimation factor of 2 in both dimensions. This approach is similar to the Gaussian pyramid used in [3] and [6]. In our research we create the following resolutions: 128×128 , 96×96 , 64×64 , 48×48 , 32×32 , 24×24 , 16×16 , 12×12 and 8×8 . Since cutting out the region of interest (ROI) and finding the landmarks in the lowest resolution would lead to serious quantisation errors, we decided to upscale all images to the same resolution. After downscaling, we scaled the image up to 256×256 for face registration and 128×128 for face recognition using a bilinear interpolation. The results can be seen in Figure 1.

III. FACE RECOGNITION SYSTEM

In this section the different steps of our face recognition algorithm are discussed. The following steps are performed in our system: face detection, registration, feature extraction and classification, the latter two are taken together as 'recognition'.

A. Face Detection

For face detection we used the OpenCV implementation [8] of the face detection algorithm first proposed by Viola and Jones [9]. The region found by this algorithm is then used in our face registration algorithm. We did not investigate the sensitivity to resolution of the face detection algorithm in this paper.

B. Face Registration and Normalization

The methods for locating (MLLL: Most Likely Landmark Locator) and correcting (BILBO) the landmarks are published in [10], [11] and summarized below. Based on the corrected landmarks the image is aligned, the ROI is determined and the face image is normalized.

1) *MLLL*: This algorithm searches for landmarks, which are typical facial features, easily distinguishable by a human observer, in the region given by the face detection algorithm. In our case we search for 17 landmarks, see Figure 2. The algorithm searches in positions around the mean location of a landmark in the detection region. The landmark is found at the location where the likelihood ratio $L_{u,v}$ is maximum. The likelihood ratio is given by:

$$L_{u,v} = \frac{p(x_{u,v}|L)}{p(x_{u,v}|\bar{L})}, \quad (1)$$

Here the vector $x_{u,v}$ contains the gray-levels of a subimage at the location (u, v) . The parameters of the probability densities $p(x_{u,v}|L)$ and $p(x_{u,v}|\bar{L})$ are respectively learnt from examples of manually labelled landmarks and non-landmarks.

2) *BILBO*: BILBO is used to correct the outliers produced by the MLLL algorithm. The landmarks which have been found by MLLL are put into a vector s , which we call the shape. The shape s is projected onto a subspace of trained correct shapes, resulting in a modified shape s' . The landmark positions which significantly changed in s' will be corrected. The complete method is described in [10]. We use the same parameters as can be found in [10] for our experiments. The results of MLLL (dots) and some corrections by BILBO (squares) are shown in Figure 2.

3) *Face Alignment*: The face images are aligned using a rigid transformation based on the landmarks

4) *Face Normalization*: We remove background and hair by taking an ROI as can be seen at the bottom row of Figure 1 and then we normalize the energy of the image inside the ROI.

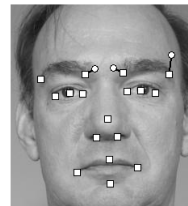


Fig. 2. Landmarks found by MLLL and corrected by BILBO

C. Face Recognition

For face recognition we do feature reduction by subsequently performing PCA [12] and LDA [13]. We use the algorithm proposed in [14] which uses the log-likelihood ratio to classify face images. For each class i the similarity score S is calculated by:

$$S_{y,i} = -(y - \mu_{W,i})^T \Sigma_W^{-1} (y - \mu_{W,i}) + y^T \Sigma_T^{-1} y - \log |\Sigma_W| + \log |\Sigma_T| \quad (2)$$

Here y is a vector which is a representation of the face image after feature reduction, Σ_T is the total covariance matrix, Σ_W is the within class covariance matrix and $\mu_{W,i}$ is the class average.

IV. EXPERIMENTS AND RESULTS

A. Experimental Setup

In our experiments we use three datasets, namely the BioID [15], the high-quality FRGC and the low-quality FRGC [16]. The BioID dataset consists of 1521 images of frontal faces of 23 persons, where we use 17 landmarks which are manually labelled in this database. From FRGC version 1, we used 3699 images taken under controlled conditions which is the high-quality FRGC dataset and 1803 images taken under uncontrolled conditions which is the low-quality FRGC dataset, the FRGC version 1 consists of 271 individuals. In our experiments we use the regions found by the face detection algorithm. In the framework of this research we are not interested in the performance of the face detection algorithm. Therefore, respectively 73 and 84 samples are removed from

the high- and low-quality dataset in which the face is not correctly detected.

For training the landmark finder we use the BioID database [15]. To train MLLL, the positive examples are cut out of every image in the dataset. For every positive example 10 negative examples are taken around that facial landmark. The BILBO algorithm is trained on the shapes of all 1521 faces in the BioID database. To verify the results of the registration we use the high-quality FRGC and the low-quality FRGC, where we search for landmarks in the region determined by the face detection algorithm.

The face recognition is performed on the low- and high-quality FRGC datasets. The datasets are randomly split into two subsets, each consisting of approximately half of the images of each person. One subset is used for training and the other for testing. We train and test the face recognition algorithm on the same resolution. The same holds for other parameters, if manual landmarks are used for training then they are also used for testing. The results of the face recognition are measured in Equal Error Rate (EER), at the point of operation where False Accept Rate (FAR) is equal to the False Reject Rate (FRR). To get more accurate results, we repeat the experiments 20 times, randomly splitting the datasets so other subsets are used in the training and test set. The EER in our case is calculated from the total set of matching and non-matching scores of all experiments [17].

B. Experiments

We conducted several experiments, investigating parts of the system and the entire system.

1) *Face Recognition*: In Experiment 1 we use manual landmarks provided with the dataset, to avoid the effects of incorrect registration. By doing this experiment we can determine a minimum resolution which still gives good results and we can also verify if the claims made in literature are valid. We also look at the effect of the components of PCA and LDA under various resolutions. Our hypothesis is that adding more dimensionality will mainly have a positive effect on results of the high resolution images, because then more discriminating details of the face can be used in classification.

2) *Face Registration*: In Experiment 2 we investigate the results of our landmark finder under different resolutions. Our first hypothesis is that the resolution needed for accurate face registration is higher than for face recognition. Our second hypothesis is that at a certain test resolution the landmark finder trained at the same resolution gives the best results. We first compare the automatically found landmarks directly with the manually labelled landmarks. We perform some experiments by training and testing the landmark finder on different resolutions. Finally we investigate the effect of face registration on the face recognition by performing registration on all resolutions, while doing the recognition on the highest resolution.

3) *Face Registration and Recognition*: To study the effect of resolution on the face registration and face recognition we use in Experiment 3 the automatically found landmarks to

register the face images and then perform the recognition. The results of the registration and recognition on the same resolution are calculated to determine the effect of resolution on the overall system.

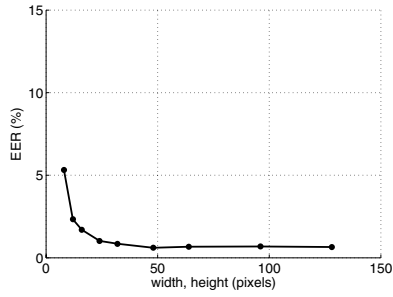
4) *Face Recognition using erroneous landmarks*: Because we only use one landmark finding technique, we decided to investigate what will happen at different resolutions if registration errors are made. Our landmark finding algorithm may perform optimal for a certain resolution, but that doesn't mean other algorithms will. This experiment allows us to predict the EER of face recognition algorithm based on the RMS error made by the registration.

In the next section we present the details and the results of the experiments.

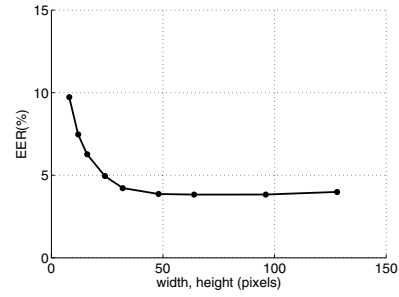
C. Results

1) *Face Recognition (Experiment 1)*: We first perform experiments using the manual labelling given by the FRGC dataset. The EER is calculated for every resolution on the low- and high-quality FRGC database and is shown in Figure 3, where this is done for the low- and high-quality FRGC database. The EERs remain almost constant down to a resolution of 32×32 pixels, below this resolution the EER increases rapidly. This seems to be in accordance with other papers where still good results are mentioned on low resolution face images. Our results on face recognition show that training and testing on low resolutions still give good results. This also means that if the lowest resolution of the face images for a certain system is given, training on this resolution gives good results for all the above resolutions down to 32×32 pixels. For these experiments on the high- and low quality FRGC we use respectively 150 and 90 PCA dimensions and 50 LDA dimensions.

To investigate the influence of the dimensionality at various resolutions, two other experiments are performed. In the first experiment we increase the number of PCA components beginning with 50 PCA components with steps of 20, and we use 50 LDA components (Figure 4). In the second experiment we use 270 PCA components for high quality FRGC and 110 PCA components for low quality FRGC and we increase the number of LDA components beginning with 10 components using steps of 20 (see Figure 5). These experiments are done on both FRGC datasets for the resolutions 128×128 , 64×64 , 32×32 and 16×16 . Figure 4 and 5 show that the results on the resolutions 64×64 and 128×128 are almost the same and that the results for 32×32 are slightly worse. Using more than 90 PCA components on the high quality FRGC will give better results for high resolutions, beyond that the results seem to remain stable on this dataset. On the low quality FRGC dataset using above 110 PCA components will worsen the results. Using more LDA components than the 50 components we already use in the first experiment also seems to worsen the results. In most cases around 30 LDA components seems to be the optimal choice. Figure 4 and 5 also show that the number of components depends more on the database that is used than on the different resolutions.

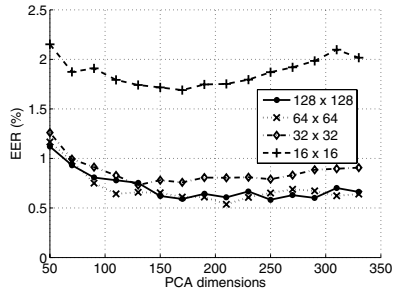


(a) High Quality FRGC

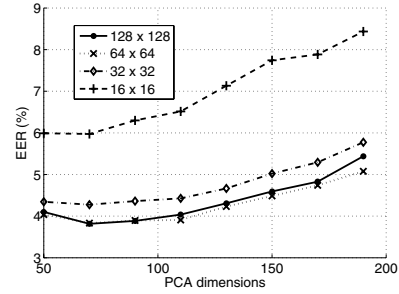


(b) Low Quality FRGC

Fig. 3. Face recognition performed at different resolutions with manual landmarks for registration

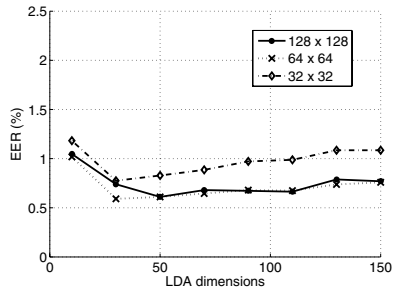


(a) High Quality FRGC

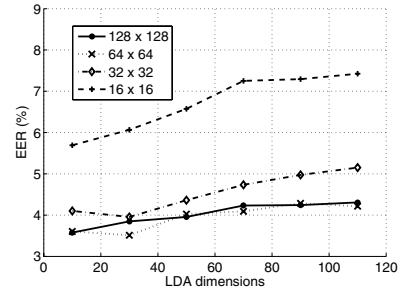


(b) Low Quality FRGC

Fig. 4. EER as a function of the number of PCA components while using 50 LDA components



(a) High Quality FRGC



(b) Low Quality FRGC

Fig. 5. EER as a function of the number of LDA components while using for the high- and low-quality FRGC respectively 270 and 110 PCA components

2) *Face Registration (Experiment 2)*: In this experiment we search for the location of the landmarks in the different resolution face images. To measure the performance of the landmark finding we use the RMS error which compares manually labelled landmarks with the automatically found landmarks after the alignment. The exact calculation of the RMS error we use, is given below, because a straightforward comparison cannot be used due to the difference in scale in the face images.

- 1) Translate, scale and rotate the groundtruth data so that the eye landmarks are on a horizontal line at a 100-pixels distance from each other.
- 2) Align the shape found to the corresponding groundtruth shape.
- 3) Calculate the Euclidian distance between each landmark

and its groundtruth equivalent.

- 4) Remove the bias caused by the different labelling policies in the databases, i.e. tip of the nose (BioID) versus a point between the nostrils (FRGC).
- 5) Calculate the RMS value of the remaining difference between the found shape and the groundtruth shape, which is now given as a percentage of the inter-eye distance.

In the FRGC database the center of the mouth is labelled, while our methods label the mouth corners. Therefore, prior to calculating the error an estimate of the center of the mouth was obtained by computing the midpoint of the mouth corners. During our experiment the training of MLLL is done using the highest resolution images of the BIOID and for testing we used the high-quality FRGC database. The results are shown in

Table I. Because the resolution of 32×32 gives the best results, we have studied the effect of training on other resolutions.

TABLE I

RMS ERROR ON HIGH QUALITY FRGC FOR DIFFERENT FACIAL FEATURES

Resolution	right eye	left eye	nose	mouth
128×128	3.0	3.2	4.4	2.8
64×64	2.6	2.7	3.7	2.5
32×32	2.1	2.2	3.2	2.2
16×16	3.1	3.2	4.6	2.9
8×8	6.4	6.8	8.2	5.6

The results of these experiments are shown in Table II and III on both the high- and low-quality FRGC, the values in the table are the mean of the RMS errors of the eyes, nose and mouth. Tables II and III show that training on the highest resolution still gives overall better results than training on lower resolutions. Tables II and III also show that the results of training on 128×128 pixels and the highest resolution are almost the same. The reason is that the resolution of the face images of this dataset is around 150×150 pixels and thus close to 128×128 pixels. It seems that MLLL performs best at the resolution of 32×32 pixels, nearly for almost all training resolutions.

We investigate the effects of this registration method on

TABLE II

RMS ERROR ON HIGH QUALITY FRGC TRAINED ON DIFFERENT RESOLUTIONS (COLUMNS)

Test Resolution	Highest	128×128	64×64	32×32
128×128	3.4	3.6	5.1	7.7
64×64	2.9	3.0	4.0	6.9
32×32	2.4	2.5	2.7	4.2
16×16	3.5	3.5	3.4	3.4
8×8	6.7	6.5	6.8	5.9

TABLE III

RMS ERROR ON LOW QUALITY FRGC TRAINED ON DIFFERENT RESOLUTIONS (COLUMNS)

Test Resolution	Highest	128×128	64×64	32×32
128×128	4.3	4.4	4.7	6.9
64×64	4.1	4.2	4.4	6.0
32×32	4.0	4.2	4.0	4.5
16×16	4.3	4.4	4.2	4.3
8×8	5.7	5.6	5.5	5.2

the face recognition, after performing face alignment based on the landmarks obtained by the landmark finder trained on the highest resolution. Because here we are only interested in the performance of the registration under different resolutions,

we performed landmark finding under the different resolution, while doing face recognition under the highest resolution of 128×128 . The results are shown in Figure 6 by the dotted line. Optimal results in EER are also reached at 32×32 . It appears that the RMS error gives a good indication of the EER, because the RMS error and EER follow the same trend.

3) *Face Recognition and Registration (Experiment 3)*: In this experiment we investigate the effects of the entire system under different resolutions. We first perform face registration and then the face recognition on the same resolution. The results of face registration and recognition under various resolutions are shown in EER in Figure 6 by the dashed line. The other lines in Figure 6 are the results of only face recognition using the manual landmarks (solid line), and the results of only face registration doing face recognition on resolution of 128×128 pixels (dotted line). Figure 6 shows that our landmark finding algorithms works best at 32×32 and also for the whole face recognition system the performance is best for this resolution. For resolutions above 32×32 pixels the influence of the registration on the EER is significant, while for lower resolutions the EER is dominated by the poor performance of the face recognition. The difference obtained in EER between manual and automatic landmarks is rather large, so there is much to gain by improving the face registration.

4) *Face Recognition by using erroneous landmarks (Experiment 4)*: In this experiment we added Gaussian noise to the manually labelled landmarks and, based upon these landmarks, the face registration and recognition is performed. For the noisy landmarks we calculated the RMS error of the landmarks and the EER of the face recognition algorithm. This experiment is performed on both FRGC datasets and the results are shown in Figure 7. It shows that the EERs of resolutions 32×32 up to 128×128 remain almost the same for all RMS errors. Results in RMS error and corresponding EER of our face registration and recognition for the resolutions 16×16 , 32×32 , 64×64 and 128×128 are shown as the large symbols (respectively, plus, diamonds, cross and dot) in the Figure 7. These results correspond well with the erroneous landmark results, which indicates that with this graph we can roughly predict the results on face recognition if we know the RMS error of the registration.

V. CONCLUSION

In this paper we investigate the effect of image resolution on the result of a face recognition system. The results confirm that face recognition algorithms using a PCA/LDA-based system are not very sensitive to resolution and still give good results at resolutions as low as 32×32 pixels. This also means training on 32×32 pixels will give good results if the face recognition system has to deal with various resolutions above 32×32 pixels. We also show that the optimal choice for the number of PCA and LDA components depends on the resolution, but much more on the dataset that is used. Increasing the amount of PCA and LDA components can help in cases where the resolution and quality of the images is high, however at the risk of overtraining.

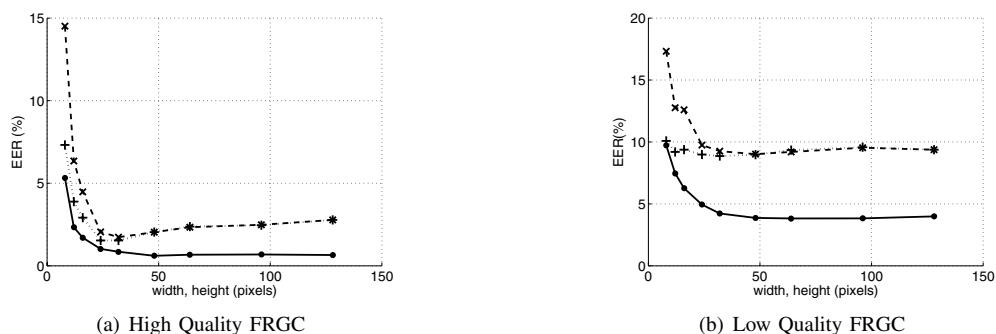


Fig. 6. Face recognition performed at different resolutions with 1. manual landmarks (solid), 2. landmark finding on all resolution and recognition at 128×128 (dotted), 3. landmark finding and recognition on all resolutions (dashed)

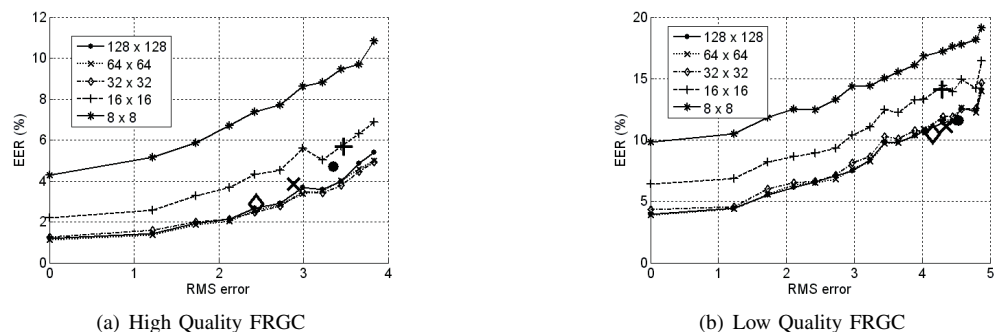


Fig. 7. EER under different RMS errors using erroneous landmarks, the large symbols are the locations using automatically obtained landmarks

We have also investigated the relation between face registration and resolution. Our registration algorithm performs best on the upscaled image with a resolution of 32×32 pixels. The landmark finding was not improved by training on the same resolution as used for testing. Other registration methods may behave differently under various resolutions. We also show that there is much to be gained by accurate registration. This can be seen in the difference in results of the face recognition between manually and automatically found landmarks. This confirms that accurate registration is of vital importance for face recognition. Our entire face recognition system works best at the resolution of 32×32 pixels.

REFERENCES

- [1] W. Zhao, R. Chellappa, and P. Phillips, "Subspace linear discriminant analysis for face recognition," Center for Automation Research, University of Maryland, College Park, Tech. Rep. CAR-TR-914, 1999.
- [2] G. Kukharev and P. Forczmanski, "Data dimensionality reduction for face recognition," *Machine Graphics & Vision*, vol. 13, pp. 99–121, 2004.
- [3] J. Wang, C. Zhang, and H. Shum, "Face image resolution versus face recognition performance based on two global methods," *Proceedings of Asia Conference on Computer Vision*, 2004. [Online]. Available: <http://learn.tsinghua.edu.cn:8080/2001216110/ACCV04.pdf>
- [4] J. Czyz and L. Vandendorpe, "Evaluation of lda-based face verification with respect to available computational resources," *PRIS 2002*, pp. 59–66, 2002.
- [5] H. K. Ekenel and B. Sankur, "Multiresolution face recognition," *Image and Vision Computing*, vol. 23, pp. 469–477, 2005.
- [6] S. Baker and T. Kanade, "Hallucinating faces," in *Fourth International Conference on Automatic Face and Gesture Recognition*, March 2000, pp. 83–89.
- [7] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1167 – 1183, September 2002.
- [8] Intel, "Open computer vision library," <http://sourceforge.net/projects/opencvlibrary/>.
- [9] P. A. Viola and M. J. Jones, "Rapid object detection using a boosted cascade of simple features," in *CVPR (1)*, 2001, pp. 511–518.
- [10] G. Beumer, Q. Tao, A. Bazen, and R. Veldhuis, "A landmark paper in face recognition," in *Automatic Face and Gesture Recognition, FG2006*. IEEE, November 2006, pp. 73–78.
- [11] A. Bazen, R. Veldhuis, and G. Croonen, "Likelihood ratio-based detection of facial features," in *Proc. ProRISC 2003, 14th Annual Workshop on Circuits, Systems and Signal Processing*, Veldhoven, The Netherlands, nov 2003, pp. 323–329.
- [12] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, pp. 71–86, 1991.
- [13] P. N. Belhumeur, J. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," in *ECCV 2*, 1996, pp. 45–58.
- [14] R. Veldhuis, A. Bazen, W. Booi, and A. Hendrikse, "Hand-geometry recognition based on contour parameters," in *Proceedings of SPIE Biometric Technology for Human Identification II*, Orlando, FL, USA, March 2005, pp. 344–353.
- [15] HumanScan, "Bioid face db," <http://www.humanscan.de/>.
- [16] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 947–954.
- [17] G. Beumer, A.M.Bazen, and R. Veldhuis, "On the accuracy of EERs in face recognition and the importance of reliable registration," in *SPS 2005*. IEEE Benelux/DSP Valley, April 2005. [Online]. Available: <http://acivs.org/sps2005/>