Multimed Tools Appl (2011) 51:863–879
DOI 10.1007/s11042-009-0417-2

SCface – surveillance cameras face database

Mislav Grgic · Kresimir Delac · Sonja Grgic

Published online: 30 October 2009

Abstract In this paper we describe a database of static images of human faces. Images were taken in uncontrolled indoor environment using five video surveillance cameras of various qualities. Database contains 4,160 static images (in visible and infrared spectrum) of 130 subjects. Images from different quality cameras should mimic real-world conditions and enable robust face recognition algorithms testing, emphasizing different law enforcement and surveillance use case scenarios. In addition to database description, this paper also elaborates on possible uses of the database and proposes a testing protocol. A baseline Principal Component Analysis (PCA) face recognition algorithm was tested following the proposed protocol. Other researchers can use these test results as a control algorithm performance score when testing their own algorithms on this dataset. Database is available to research community through the procedure described at www.scface.org.

Keywords Video surveillance cameras · Face database · Face recognition

1 Introduction

Interest in face recognition, as a combination of pattern recognition and image analysis is still growing. Many papers are written and many real-world systems are being developed and distributed [1, 12, 19, 22]. As a non-invasive biometric method, face recognition is attractive for national security purposes as well as for smaller scale surveillance systems. However, in order to be able to claim that any face recognition system is efficient, robust and reliable, it must undergo rigorous testing and verification, preferably on real-world datasets. In spite of many face databases currently available to researchers [7–9, 15 www.face-rec.org/databases], we feel they lack the real-world settings part. Images in other currently available databases are usually taken by the same camera e.g. Video database of moving faces and people [14], CMU PIE [18], FERET [16, 17] and are not taken using

M. Grgic (✉) · K. Delac · S. Grgic
University of Zagreb, Unska 3/XII, Zagreb, Croatia
e-mail: mislav.grgic@fer.hr
URL: http://www.vcl.fer.hr/mgrgic/

proper, commercially available, surveillance equipment (e.g. some multimodal databases that include sets for voice recognition—BANCA [3], XM2VTSDB [13]).

Images in those databases are mainly taken under strictly controlled conditions of lighting, pose, etc., and are of high resolution (high quality capturing equipment is used). Although some of the most frequently used databases have a standardized protocol (in order to assure reproducibility of results) all this is still far from real-world conditions. For real-world applications, algorithms should operate and be robust on data captured from video cameras (of various qualities) on a public place. Illumination should be uncontrolled (i.e. coming from a natural source) with varying direction and strength. Head pose should be as natural as possible. All this mimics a really complex imaging environment, such as is expected when building an airport security system or any other national security system. The obvious lack of such image sets is the main reason for a low number of studies on face recognition in such naturalistic conditions, resulting in very high recognition rates suggesting that face recognition is almost a solved problem. This was the main motivation for collecting our database. Our database mimics the real-world conditions as close as possible and we think that using our database in experiments will show that real-world indoor face recognition by using standard video surveillance equipment is far from being a solved problem.

The problem specifically addressed with our database is the law enforcement person identification from low quality surveillance images or any other identification scenario where the subject's cooperation is not expected. Recognizing how important this topic is, researchers are performing more and more experiments using CCTV (Closed Circuit Television) images, exploring performance of both humans [10] and computers [4]. Since up to now there were no publicly available CCTV face images databases, researchers were forced to perform the experiments using their images they captured themselves, which in turn often results in having a low number of subjects and results which have questionable statistical significance. We feel that our SCface database will successfully fill this gap.

For this purpose, we collected images from 130 subjects. The database consists of 4,160 images in total, taken in uncontrolled lighting (infrared night vision images taken in dark) with five different quality surveillance video cameras. Two cameras were able to operate in infrared (IR) night vision mode so we recorded that IR output as well (we will address these images as IR images; for cameras details please see the Appendix). Subjects' images were taken at three distinct distances from the cameras with the outdoor (sun) light as the only source of illumination. All images were collected over a 5 days period.

This database was collected mainly having identification rather then verification in mind, with the additional motive being the fact that identification is a more demanding recognition problem with its one-to-many comparisons. The proposed protocol and initial testing with PCA [20] will also be along these lines of thought. Reported results will accordingly show rank 1 recognition percentages, but Cumulative Match Score (CMS) curves [16] will be omitted, since such a detailed analysis is beyond the scope of this paper.

Here is a short summary of what makes this dataset interesting to face recognition research community:

i) Different quality and resolution cameras were used;
ii) Images were taken under uncontrolled illumination conditions;
iii) Images were taken from various distances;
iv) Head pose in surveillance images is typical for a commercial surveillance system, i.e. the camera is placed slightly above the subject's head, making the recognition even more demanding; Besides, during the surveillance camera recordings the individuals *were not* looking to a fixed point;

v) Database contains nine different poses images suitable for head pose modeling and/or estimation;
vi) Database contains images of 130 subjects, enough to eliminate performance results obtained by pure coincidence (chances of recognition by pure coincidence is less than $1/130 \approx 0.8\%$);
vii) Both identification and verification scenarios are possible, but the main idea is for it to be used in difficult real-world identification experiments.

The rest of this paper is organized as follows: Section 2 gives a detailed database description including the acquisition, naming conventions, database forming stage, image sets details and database demographics; Section 3 discusses potential uses of the database; Section 4 proposes evaluation protocols and results presentation; Section 5 describes an *ad hoc* baseline PCA test on this database in extremely difficult experimental setup where a separate database is used for training; Section 6 concludes the paper.

2 Database description

Our database was designed mainly as a means of testing face recognition algorithms in real-world conditions. In such a setup, one can easily imagine a scenario where an individual should be recognized comparing one frontal mug shot image to a low quality video surveillance still image. In order to achieve a realistic setup we decided to use commercially available surveillance cameras of varying quality (see Appendix for details). Since two of our surveillance cameras record both visible spectrum and IR night vision images, we decided to include IR imagery in the database as well.

2.1 Image acquisition: equipment setup and imaging procedure

Capturing face images took place in Video Communications Laboratory at the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia [11]. Capture equipment included: six surveillance cameras, professional digital video surveillance recorder, professional high-quality photo camera and a computer. For mug shot image acquisition we used a high-quality photo camera. For surveillance camera image acquisition we used five different (commercially available) models of surveillance cameras (see Appendix for details) and for IR mug shots a separate surveillance camera was used. Five surveillance cameras were installed in one room at the height of 2.25 m and positioned as illustrated in Figs. 1 and 2. The only source of illumination was the outdoor light, which came through a window on one side. Two out of five surveillance cameras were able to record in IR night vision mode as well. The sixth camera was installed in a separate, darkened room for capturing IR mug shots. It was installed at a fixed position and focused on a chair on which participants were sitting during IR mug shot capturing. The IR part of the database is important since there are research efforts in this direction, e.g. [12] and [5], yet no available database provides such images.

The high-quality photo camera for capturing visible light mug shots was installed the same way as the infrared camera but in a separate room with the standard indoor lighting and it was equipped with adequate flash. Mug shot imaging conditions are exactly the same as would be expected for any law enforcement or national security use (passport images or any other personal identification document). All six cameras (five surveillance and one IR mug shot) were connected to a professional digital video surveillance recorder, which was

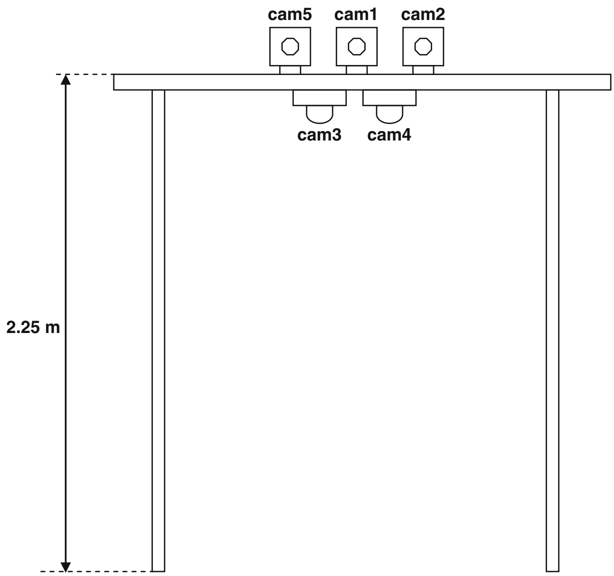


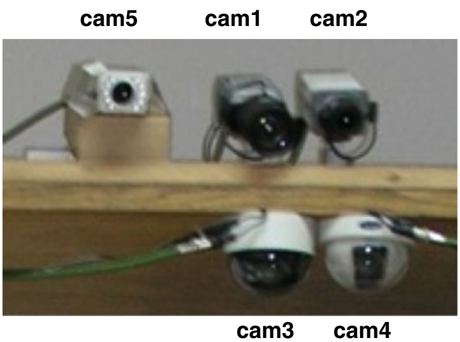
Fig. 1 Cameras set up

recording all six video streams simultaneously all the time on internal hard disk. For storage of images and for controlling surveillance cameras we used Digital Sprite 2 recorder with adequate software (DM Multi Site Viewer) provided from manufacturer for connecting and controlling over personal computer. Settings recommended by the vendor were used for capturing images on Digital Sprite 2, because those settings represent most commonly used parameters set up in real life surveillance systems (see Appendix for details).

The surveillance cameras are named cam1, cam2, cam3, cam4 and cam5. Cam1 and cam5 are also able to work in IR night vision mode. We decided to name the images taken by them in the IR night vision mode as cam6 (actually cam1 in night vision mode) and cam7 (actually cam5 in night vision mode). Camera for taking IR mug shots was named cam8. All cameras (surveillance and photo) were installed and fixed to same positions and were not moved during the whole capturing process.

The capturing was conducted over a 5 days period. All participants in this project have passed through the following procedure. First they had to walk in front of the surveillance cameras in the dark and after that they had to do the same in uncontrolled indoor lighting.

Fig. 2 Cameras positions



During their walk in front of the cameras they had to stop at three previously marked positions (Fig. 3). This way 21 images per subject were taken (cam1–7 at distances of 4.20, 2.60 and 1.00 m).

After that, participants were photographed with digital photographer's camera at close range in controlled conditions (standard indoor lighting, adequate use of flash to avoid shades, high resolution of images). This set of images provides nine discrete views of each face, ranging from left to right profile in equal steps of 22.5 degrees. To assure comparable views for each subject, numbered markers were used as fixation points. As a final result, there are nine images per subject with views from −90 to +90 degrees including another mug shot at 0 degrees (see Fig. 4).

In the end, subjects went into the dark room where high quality IR night vision surveillance camera was installed for capturing IR mug shots at close range. In overall that gives 32 images per subject in the database.

2.2 Naming conventions and final database forming

After the capturing procedure we had to extract only the subjects' faces from the gathered images. Firstly, all captured images were transferred from Digital Sprite 2 recorder on a PC hard drive and were named in the following manner:

- Surveillance cameras (cam1–7): *subjectID_camNum_distanceLabel.jpg*
- IR frontal mug shot: *subjectID_cam8.jpg*
- Visible light mug shot: *subjectID_frontal.jpg*
- Different pose images: *subjectID_angleLabel.jpg*

This way every image in the database gained a unique name carrying information both about a subject's unique ordinal and at what distance and imaging conditions is the image taken. Distance labels 1, 2 and 3 represent distances of 4.20, 2.60 and 1.00 m, respectively

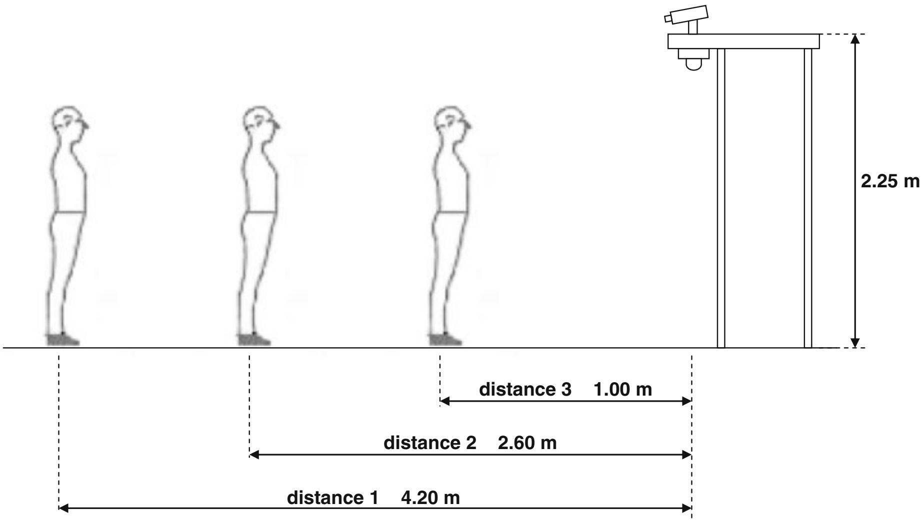


Fig. 3 Surveillance cameras capturing distances

Fig. 4 Example of different pose images

(Fig. 3). As an example, the filename *001_cam1_1.jpg* means that this image shows subject 001 captured with surveillance camera 1 at a distance of 4.20 m. The whole surveillance image set (including two mug shots) for one subject is presented in Fig. 5. A textual database with coordinates of eyes, tip of the nose and center of the mouth is created (Fig. 6), containing the details for all images in the database (except for different pose images). This textual database was created using software developed in JAVA (especially for this project) at the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. The software presents an image to the operator who then clicks on the eyes' center point and the software records the coordinates.

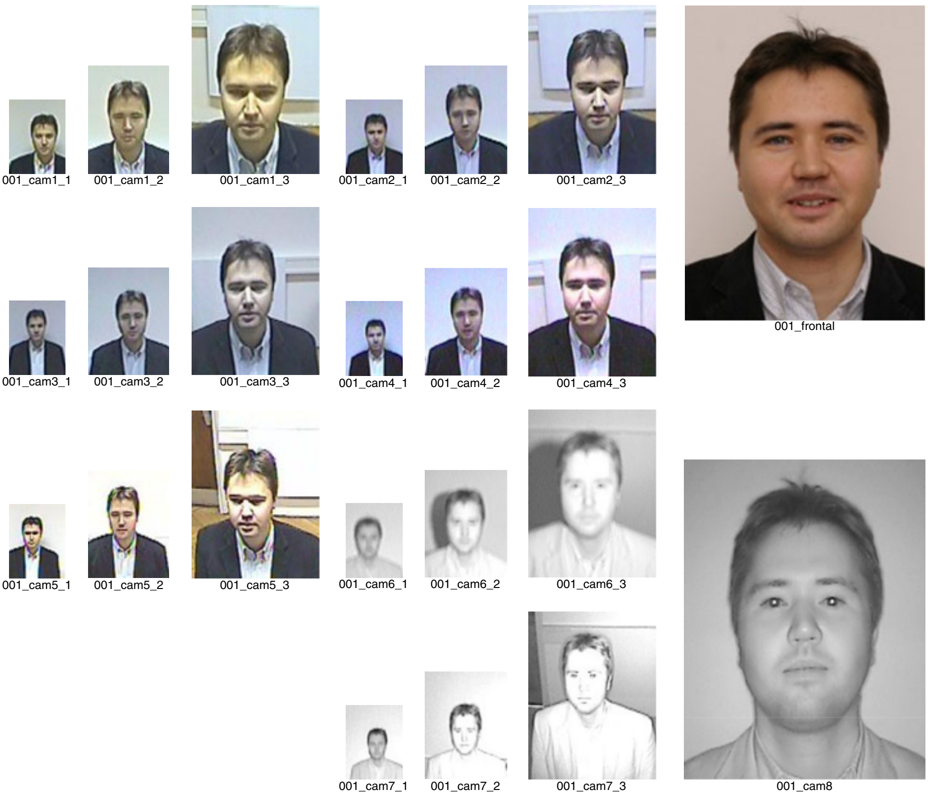


Fig. 5 Example of one image set for one subject

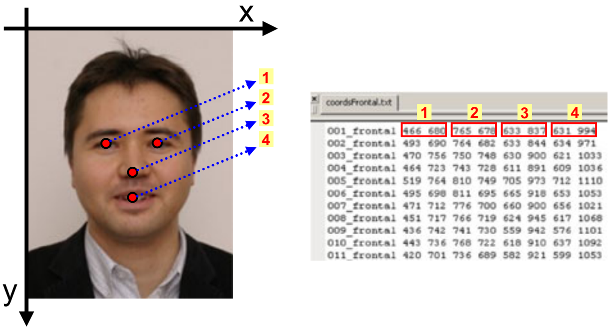


Fig. 6 Coordinates of the centers of the eyes, tip of the nose and mouth, and its corresponding textual database entries

2.3 Image sets details

2.3.1 Frontal facial mug shots

Facial mug shots are high quality static color images, taken in controlled indoor illumination conditions environment using Canon EOS 10D digital camera. There is one mug shot per subject and those images are labeled with the suffix *frontal* (e.g. 001_frontal.jpg). Images are in lossless 24 bit color JPEG format with the original size of 3,072 × 2,048 pixels, cropped to 1,600 × 1,200 pixels. Cropping was done following the ANSI 385-2004 standard [2] recommendation so that the face occupies approximately 80% of the image. Exact positions of eye centers, tip of the nose and the center of the mouth are manually extracted and stored in a textual file (this was also done for all images in the database except for the different pose set). These mug shot images are what you would expect to find in a law enforcement database or when registering to a security system. There are in total 130 frontal facial mug shot images in the database, one per subject.

2.3.2 Surveillance cameras images (visible light)

Cameras 1–5 are visible light cameras and all images with *camNum* labels cam1–5 represent images taken with surveillance cameras of different qualities. There are three images per subject for each camera, taken at three discrete distances (4.20, 2.60 and 1.00 m). This gives total of 15 images per subject in this set (1950 in total). This set was designed to test the face recognition algorithms in real-world surveillance setup (we again emphasize that we used commercially available cameras of both low and high quality). As can be seen from the Fig. 5, the images differ substantially in quality and resolution. The original images saved from Digital Sprite 2 recorder to PC hard drive all have the same dimensions (680 × 556 pixels, 96 dpi and 24 bit color). They have been cropped in order to remove as much background as possible. Due to different distances at which the images were taken, new (cropped) images are not all the same size. Cropped images have the following resolutions: 100 × 75 pixels for distance 1, 144 × 108 for distance 2 and 224 × 168 for distance 3. All the coordinates mentioned previously were labeled and stored in a textual database for this set of images as well.

2.3.3 IR night vision mug shots

IR night vision mug shots were taken in a separate dark room with a resolution of 426×320 pixels, grayscale. There is one image per subject in this set, yielding a total of 130 of those images in the database.

2.3.4 Surveillance cameras images (IR night vision)

IR night vision surveillance cameras images are completely the same as visible light surveillance images in section 2.3.2 regarding the resolution but are in grayscale. IR surveillance images are labeled as cam6 and cam7, again at three discrete distances. There are six IR night vision surveillance images per subject, thus 780 of them in total.

2.3.5 Different pose images

This set of images was taken with the same high quality photo camera as frontal facial mug shots and under the same conditions. Subjects' pose ranges from −90 (left profile) to +90 degrees (right profile) in nine discrete steps of 22.5 degrees. Image size is 3,072×2,048 pixels and frontal mug shot images of the same size are included for the sake of completeness. There are nine images per subject in this set, which gives 1,170 images in total. The correlation between the angle label in the file name and the actual head pose is given in the Appendix.

2.4 Database demographics

The participants in this project were students, professors or employees at the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. From total of 130 volunteers, 114 were males and 16 females. All participants were Caucasians, between the ages of 20 and 75. The age distribution is displayed in Fig. 7. Besides, the textual file containing each subjects date and year of birth will be distributed with the database. This

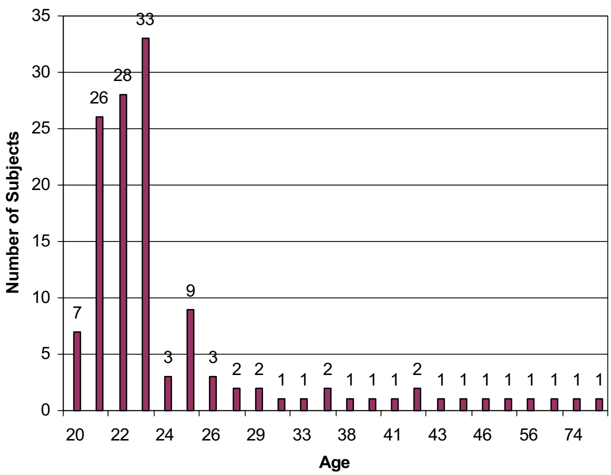


Fig. 7 Database subjects' age distribution

also makes our SCface database unique, because subjects' birthdays are not available in any other database (www.face-rec.org/databases). Moreover, the textual file will contain the following additional information: gender, beard, moustache, glasses.

3 Potential uses of the database

The first and the most important potential use of this database is to test the face recognition algorithm's robustness in a real-world surveillance scenario. In such a setup a face recognition system should recognize a person by comparing an image captured by a surveillance camera to the image stored in a database. If we postulate a law enforcement use as the most potential scenario, this database image, to which the surveillance image is compared to, is a high quality full frontal facial mug shot. We would specifically like to encourage researchers to explore more deeply the small sample size problem (more dimensions than examples). Having one frontal mug shot per subject in our database addresses this issue adequately.

Our surveillance cameras were of various quality and resolution (see Appendix) and this issue is the strong point of this database. It remains to be seen how will face recognition algorithms perform in such difficult conditions and how does the quality of capturing equipment and subject's distance from camera influence the results. There is also a potential to test various image preprocessing algorithms (enhancing image quality by filtering techniques), as some of these surveillance images are of extremely low quality and resolution.

By including different pose images of subjects, we made it possible to use this database in face modeling and 3D face recognition.

Other potential uses of this database include but are not restricted to: evaluation of head pose estimation algorithms, evaluation of face recognition algorithms' robustness to different poses, evaluation of natural illumination normalization algorithms, indoor face recognition (in uncontrolled environment), low resolution images influence, etc.

4 Proposed evaluation protocol

4.1 DayTime tests

This should be a straightforward test comparing the mug shot image (visible light, i.e. *subjectID_frontal.jpg* images) to cam1–5 images. Frontal mug shots represent the gallery of known images and cam1–5 images and three distances are probe sets (the main idea is similar to FRVT tests as in [16]). This scenario gives 15 possible different probe sets, varying both in distances from camera and in camera qualities. Comparing the probe image to one gallery image is the most logical real-world (law enforcement) scenario.

4.2 NightTime tests

NightTime tests have two possible galleries that can be used. Cam6-7 images (IR night vision) can be compared to mug shot images (visible light, i.e. *subjectID_frontal.jpg* images) or they can be compared to cam8 (IR night vision mug shot) images. This scenario gives two galleries and six possible different probe sets, again varying both in distances from camera and in camera qualities.

4.3 Performance metrics

Rank 1 results and Cumulative Match Score (CMS) curves are the usual way to report results in identification scenario (for details see [16]), while ROI plots are the usual way of reporting results in verification scenario. Besides, we suggest the use of some form of statistical significance test like McNemar's hypothesis test [6, 21]. Error bars and standard deviations of performance results are also something that would help establish the true robustness and performance improvement of any algorithm. Control algorithm performance scores must accompany performance results of novel algorithms. This control algorithm should be something that is easily implemented and readily available to all, like PCA or correlation. An initial baseline PCA tests will be provided here, but the researchers are encouraged to expand and improve them to make solid grounds for control algorithms comparisons in the future.

4.4 Training scenario suggestion

In order for any tests to be close to expected real-world (law enforcement) applications and for the results to be meaningful, we strongly suggest the use of a separate set of face images for training, especially in the identification scenario. In other words, any experiment should be designed so that images used in algorithm training stage have no overlap with the query images (gallery and probe). We believe that this is the only way to make fair comparisons and to present fair performance results. Other databases (like FERET or CMU PIE) can be used for training, similar to the setup from [4]. All algorithms will probably experience a serious performance drop with this experimental setup, but it will actually show their true strength.

5 Baseline PCA performance evaluations

In order to show exactly how demanding this dataset is and to provide a potential benchmark performance results for further algorithms testing and comparisons, we performed an *ad hoc* experiment using baseline PCA as face recognition algorithm and followed the evaluation protocol proposed in Section 4.

5.1 Image preprocessing

We normalized all images from our SCface database following the standard procedure often described in many face recognition papers. Firstly, all color images were turned to grayscale. By using eye coordinates information, all images are scaled and rotated so that the distance between the eyes is always the defined number of pixels (32 in our case) and the eyes lie on a straight line (head tilt is corrected). Then the image is cropped to 64×64 pixels and masked with an elliptical mask. Two examples of preprocessed images can be seen in Fig. 8. The eyes are at positions (16,16)—left eye, and (16,48)—right eye. Standard histogram equalization (Matlab) is done prior to masking, normalizing the pixel values to [0, 255].

5.2 Experimental design

The experiment performed is set up in a most difficult manner imaginable, but it mimics what can be expected in real-world face recognition applications: *the training set is different*

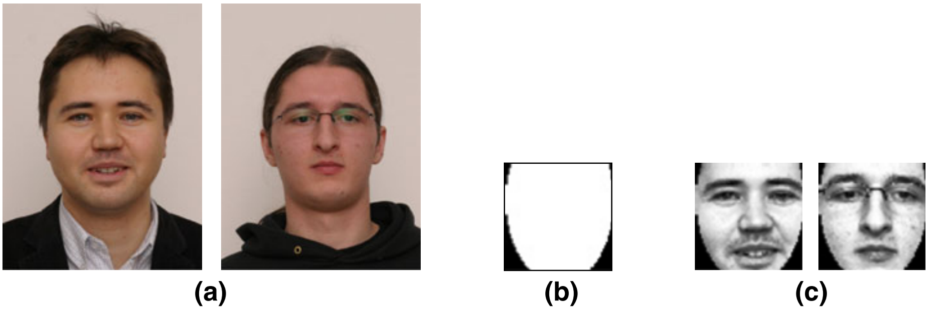


Fig. 8 **a** Original frontal mug shot images; **b** Standard mask; **c** Images after preprocessing

than the test set. In other words, training and test sets completely independent (images used in training will not be used in gallery or probe sets). We will use a totally different database for training to insure just that. This is why the recognition rates achieved will seem so low. This kind of setup, although often discussed, is rarely used in research papers, but we feel it is time to push this research area into a new level.

We will preprocess 501 randomly chosen FERET database images in the same manner as we preprocessed our database images and we will use those images to train baseline PCA. PCA algorithm details are well known and are beyond the scope of this paper. They will not be discussed here and an interested reader is referred to [20]. Gallery and probe images will be from our SCface database but with no overlap between the two sets (no images from the gallery are in the probe set and vice versa). By training the PCA with one image per person (the frontal mug shot) we will also address the small sample size problem (more dimensions than examples).

After the training, we kept 40% of the eigenvectors with highest eigenvalues, thus forming 200 dimensional face (sub)space in which the recognition will be done. Using the fact that once the PCA is trained the distance between any pair of projected images is constant, we were able to perform virtual experiments. We projected all images from our database onto the derived 200 dimensional subspace and calculated distance matrix (distance matrix being the size n x n, where n is the number of images in our database, and containing the distances between all images). All experiments were then performed on those matrices by changing the list items in probe and gallery sets and without running the algorithms again. Distance metric used was the cosine angle (*COS*).

5.3 Performance results

5.3.1 DayTime experiments

As can be seen from Table 1, the results never exceed 10% recognition rate at rank 1. We will restrict ourselves to rank 1 results as detailed performance analysis is beyond the scope of this paper. The results are quite low, ranging from below 1% to about 8%. We can see that distance from the camera has some influence on performance, but it is not a straightforward connection. The results at distance 3 are sometimes lower than distance 2 results. This can be explained by the fact that images taken at distance 3 are more of the top part of the head, thus covering the part of the face normally seen in frontal images. Since we used mug shot frontal images as gallery, the low results for distance 3 are expected.

Table 1 Rank 1 performance results for DayTime experiments. Gallery: mug shot frontal images. Distance metric: cosine angle

Camera/Distance	Rank 1 recognition rate [%]
cam1_1	2.3
cam1_2	7.7
cam1_3	5.4
cam2_1	3.1
cam2_2	7.7
cam2_3	3.9
cam3_1	1.5
cam3_2	3.9
cam3_3	7.7
cam4_1	0.7
cam4_2	3.9
cam4_3	8.5
cam5_1	1.5
cam5_2	7.7
cam5_3	5.4

5.3.2 NightTime experiments

NightTime experiments can be seen in Table 2. When frontal images are used as gallery the results are persistent with the DayTime experiments. However, when night vision mug shots (cam8) were used as gallery we can see some improvements, although not very impressive ones. All other conclusions from DayTime experiments still hold.

5.4 Discussion

A question one must ask after seeing such poor results is not how good baseline PCA algorithm is, but how representative of the entire population is the subset of images used for training. Since we used a totally different database for training, the results are low as expected. This is the most difficult experimental setup for any algorithm. Training with totally different database is just part of the problem. As can be seen at sample images in Fig. 5, images are quite difficult for a simple straightforward recognition algorithm such as PCA. The images vary in resolution, pose, illumination, expression, etc. All this makes this

Table 2 Rank 1 performance results for NightTime experiments. Gallery: mug shot frontal / night vision (cam8) images. Distance metric: cosine angle

Camera/Distance	Rank 1 recognition rate [%]	
	Gallery: frontal	Gallery: cam8
cam6_1	1.5	1.5
cam6_2	3.1	10.0
cam6_3	3.9	10.0
cam7_1	0.7	1.5
cam7_2	5.4	2.3
cam7_3	4.6	7.0

dataset an extremely challenging performance testing tool for any algorithm. DayTime and NightTime (IR night vision imaging) experiment show no significant difference in performance. Again, this can be expected, as the images from cam6 and cam7 are quite similar to others (Fig. 5).

As commented before, distance has a large influence on results due to the different viewing angle at different distances. Besides the angle, image resolution is an important issue to address as well. Images taken at distance 1 (4.20 m) are 100×75 pixels and only a portion of that is the actual face. Although, there were some discussions in the literature before that 19×19 pixels face images are large enough for recognition, this conclusion does not stand here, probably due to extremely difficult imaging conditions.

Camera quality showed no particular influence since we used cameras with various qualities for database acquisition, but the results are consistently low across all probe sets.

6 Conclusion

In this paper we described a database of 4,160 static images of 130 human subjects and the corresponding proposed testing protocol. Commercially available video surveillance equipment of varying quality was used for database acquisition. All cameras were placed as they would be placed in any security system and illumination conditions were uncontrolled. All this should mimic the real-world conditions as close as possible and provide researchers a solid (and very difficult) experimental data. Face images in this database vary in resolution, pose, illumination, expression, etc. This database is not suitable for exploring each individual problem of face recognition field (such as different illumination or expression) but is engineered so that all the usual problems come together in one difficult dataset. By performing a simple baseline PCA on this dataset (using a separate database for training) we showed that recognition performance is quite low compared to the usual PCA performance reported for other databases. By making this database available to all the research community, we hope to encourage the exploration of yet unsolved face recognition issues, including but not restricted to the uncontrolled indoor imaging conditions, non-overlapping training and query datasets and small sample size problem, and recognition from real-life surveillance images. Other algorithms, superior to simple PCA, are yet to be tested using this database and the proposed protocol, thus proving their robustness and efficiency. Hopefully, this challenging database will push the state-of-the-art face recognition to the next level and prompt the researchers to explore a more difficult set of problems in the future. As can be seen from the PCA results on this dataset, face recognition (particularly in indoor, real-world conditions) is far from being a solved issue.

Acknowledgement The authors would like to thank Bozidar Klimpak for his technical support during database acquistion and indexing, Kresimir Marusic and Tehnozavod Marusic Ltd. for providing surveillance system equipment, Boris Krzic for providing professional photo equipment and his help with mug shots capturing, Darko Poljak for developing and providing JAVA software used for semiautomatic determination of eyes, nose and mouth coordinates, and to all participants in this project. Portions of the research in this paper use the FERET database of facial images collected under the FERET program. The authors would like to thank the FERET Technical Agent, the U.S. National Institute of Standards and Technology (NIST) for providing the FERET database. The work described in this paper was conducted under the research project "Intelligent Image Features Extraction in Knowledge Discovery Systems" (036-0982560-1643), supported by the Ministry of Science, Education and Sports of the Republic of Croatia.

Appendix

Table 3 Different pose image labels details

Image label	Head pose
SubjectID_F	F - frontal
SubjectID_L1	L1 - left, 22.5
SubjectID_L2	L2 - left, 45.0
SubjectID_L3	L3 - left, 67.5
SubjectID_L4	L4 - left, 90.0
SubjectID_R1	R1 - right, 22.5
SubjectID_R2	R2 - right, 45.0
SubjectID_R3	R3 - right, 67.5
SubjectID_R4	R4 - right, 90.0

For creating SCface database six surveillance cameras were used, a digital photographer's camera, digital video surveillance recorder, personal computer (PC) and other accessories used for connecting and installing cameras, recorder and a PC.

Surveillance cameras

cam1, cam6–Bosch LTC0495/51
cam2–Shany WTC-8342
cam3–J&S JCC-915D
cam4–Alarmcom VFD400- 12B
cam5, cam7, cam8–Shany MTC-L1438

Digital video surveillance recorder was Digital Sprite 2 from Dedicated Micros. The Digital Sprite 2 has 16 input video channels, two monitor outputs and LAN connector. It has 600 GB internal hard disk for storage of video streams from cameras and CD writer for writing data directly on CDs. It was set up for capturing 2 pictures per second from each camera, what gives a total of 12 pictures per second. Each picture size was 40 kB.

Table 4 Surveillance cameras specifications

Type	cam1 Bosch LTC0495/51	cam2 Shany WTC-8342	cam3 J&S JCC-915D	cam4 Alarmcom VFD400-12B	cam5 Shany MTC-L1438
CCD Type	1/3″ IT	1/3″ Sony Super HAD	1/3″ Color	1/3″ IT	1/3″ Sony Super HAD
Active pixels	752×582	795×596	597×537	752×582	795×596
Resolution	540 TVL	480 TVL	350 TVL	460 TVL	480 TVL
Minimum illumination	0,24/0,038 lux (IR)	0,15 lux	0,3 lux	1,5 lux	0 lux (IR LED on at 4 lux)
SNR	> 50 dB	> 50 dB	> 48 dB	46 dB	> 50 dB
Video output	1 Vpp, 75 Ω	1 Vpp, 75 Ω	1 Vpp, 75 Ω	1 Vpp, 75 Ω	1 Vpp, 75 Ω
Comment	IR night vision		dome camera	dome camera	IR night vision

Mug shot camera

Photographer's camera was Canon EOS 10D model with 22.7×15.1 mm CMOS sensor, with 6.3 mega pixels, equipped with Sigma 18–50 mm F3.5–5.6 DC lenses and Sigma EF 500 DG Super flash.

References

1. Abatea AF, Nappi M, Riccioa D, Sabatino G (2007) 2D and 3D face recognition: a survey. Pattern Recogn Lett 28(14):1885–1906
2. ANSI INCITS 385-2004 (2004) Face Recognition Format for Data Interchange. May
3. Bailly-Bailleire E, Bengio S, Bimbot F, Hamouz M, Kittler J, Mariethoz J, Matas J, Messer K, Popovici V, Poree F, Ruiz B, Thiran J (2003) The BANCA database and evaluation protocol. Lect Notes Comput Sci 2688:625–638
4. Chen S, Berglund E, Bigdeli A, Sanderson C, Lovell BC (2008) Experimental analysis of face recognition on still and CCTV images. Proceedings of the 2008 IEEE fifth international conference on advanced video and signal based surveillance, pp 317–324.
5. Chen X, Flynn PJ, Bowyer KW (2005) IR and visible light face recognition. Comput Vis Image Underst 99(3):332–358
6. Delac K, Grgic M, Grgic S (2005) Independent comparative study of PCA, ICA, and LDA on the FERET data set. Int J Imaging Syst Technol 15(5):252–260
7. Face recognition homepage, databases section: www.face-rec.org/databases
8. Gao W, Cao B, Shan S, Zhou D, Zhang X, Zhao D (2004) The CAS-PEAL large-scale chinese face database and evaluation protocols. *Technical Report No. JDL_TR_04_FR_001*, Joint Research & Development Laboratory. CAS
9. Gross R (2005) Face databases, in Li S and Jain A (eds.) Handbook of face recognition, Springer
10. Keval HU, Sasse MA (2008) Can we ID from CCTV? Image quality in digital CCTV and face identification performance. *Proc. SPIE 6982*, paper ID 69820K
11. Klimpak B, Grgic M, Delac K (2006) Acquisition of a face database for video surveillance research. Proceedings of the 48th international symposium ELMAR-2006 focused on multimedia signal processing and communications, Zadar, Croatia, pp 111–114
12. Kong SG, Heo J, Abidi BR, Paik J, Abidi MA (2005) Recent advances in visual and infrared face recognition—a review. Comput Vis Image Underst 97(1):103–135
13. Messer K, Matas J, Kittler J, Luettin J, Maitre G (1999) XM2VTSDB: The Extended M2VTS Database. Second int. conf. audio and vide-based biometric person authentication (AVBPA'99), Washington D.C., USA, pp 72–77
14. O'Toole AJ, Harms J, Snow SL, Hurst DR, Pappas MR, Ayyad JH, Abdi H (2005) A video database of moving faces and people. IEEE Trans. Pattern Analysis and Machine Intelligence 27 (5):812–816
15. Phillips PJ, Flynn PJ, Scruggs T, Bowyer KW, Chang J, Hoffman K, Marques J, Min J, Worek W (2005) Overview of the face recognition grand challenge. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) 1:947–954
16. Phillips PJ, Moon H, Rizvi SA, Rauss P (2000) The FERET evaluation methodology for face-recognition algorithms. IEEE Trans. Pattern Analysis and Machine Intelligence 22(10):1090–1104
17. Phillips PJ, Wechsler H, Huang J, Rauss P (1998) The FERET database and evaluation procedure for face recognition algorithms. Image Vis Comput 16(5):295–306
18. Sim T, Baker S, Bsat M (2003) The CMU pose, illumination, and expression database. IEEE Trans. Pattern Analysis and Machine Intelligence 25(12):1615–1618
19. Tan X, Chen S, Zhou Z, Zhang F (2006) Face recognition from a single image per person: a survey. Pattern Recogn 39(9):1725–1745
20. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Neurosci 3(1):71–86
21. Yambor W, Draper B, Beveridge JR (2002) Analyzing PCA-based face recognition algorithms: eigenvector selection and distance measures. In: Christensen H, Phillips J (eds) Empirical evaluation methods in computer vision. World Scientific, Singapore
22. Zhao W, Chellappa R, Rosenfeld A, Phillips PJ (2003) Face recognition: a literature review. ACM Comput Surv 35(4):399–458

Mislav Grgic received B.Sc., M.Sc. and Ph.D. degrees in electrical engineering from the University of Zagreb, Faculty of Electrical Engineering and Computing (FER), Zagreb, Croatia, in 1997, 1998 and 2000, respectively. From July 1997 he is working at the Department of Wireless Communications, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. He was a visiting researcher at the University of Essex, Colchester, United Kingdom (1999/2000). In April 2007 he was promoted to Associate Professor. He participated in 5 scientific projects financed by the Ministry of Science, Education and Sports of the Republic of Croatia and 3 international COST projects of the European Commission. Currently he is a project leader of the research project: "Intelligent Image Features Extraction in Knowledge Discovery Systems" financed by the Ministry of Science, Education and Sports of the Republic of Croatia and a member of the Management Committee of the European project COST Action IC0604. He published more than 110 papers in books, journals and conference proceedings in the area of image and video compression, content-based image retrieval, face recognition and digital mammography (computer-aided detection and diagnosis of breast cancer). Prof. Grgic is a senior member of IEEE. He is a Vice Chair of the IEEE Croatia Section and a collaborating member of the Croatian Academy of Engineering (HATZ). Prof. Grgic participates in more than 40 conference international programs committees, he is member of 6 journal editorial boards and he serves as a technical reviewer for various scientific journals. Prof. Grgic received bronze and silver medals "Josip Loncar" from FER for outstanding B.Sc. and M.Sc. theses, and "Vera Johanides" award from the Croatian Academy of Engineering for scientific achievements in the area of multimedia communications, in 1997, 1999 and 2005, respectively. Prof. Grgic is a member of the Working Group for the preparation of negotiations for the accession of the Republic of Croatia to the European Union, responsible for the Chapter 10 of the acquis communautaire: "Information Society and Media" and an appointed member of the Committee on Transportation, Communications and Maritime Affairs of the Croatian Parlament.

Kresimir Delac received his B.Sc.E.E degree in 2003, and Ph.D. degree in 2007 from University of Zagreb, Faculty of Electrical Engineering and Computing. He is a recipient of the silver medal "Josip Loncar" from the Faculty of Electrical Engineering and Computing in Zagreb for an outstanding Ph.D. thesis work. His research interests are image analysis and pattern recognition in biometric and medical applications. He is working as a consultant in a privately held company and is also affiliated to University of Zagreb, Faculty of Electrical Engineering and Computing, as a researcher on the project “Intelligent Image Features Extraction in Knowledge Discovery Systems” financed by the Ministry of Science, Education and Sports of the Republic of Croatia and as a guest lecturer. He is also a member of the Management Committee of the European project COST Action IC0604.

Sonja Grgic received the B.Sc., M.Sc. and Ph.D. degrees in electrical engineering from University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia, in 1989, 1992 and 1996, respectively. She is currently full Professor at the Department of Wireless Communications, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. Her research interests include television signal transmission and distribution, picture quality assessment, wavelet image compression and broadband network architecture for digital television. She was a member of the international program and organizing committees of many international workshops and conferences. She was a visiting researcher at the Department of Telecommunications, University of Mining and Metallurgy, Krakow, Poland. She is the recipient of the silver medal “Josip Loncar” from the Faculty of Electrical Engineering and Computing in Zagreb for an outstanding Ph.D. thesis work and annual award “Rikard Podhorsky” from Croatian Academy of Engineering (HATZ) for year 2006. She has participated in 10 domestic and international scientific projects. Currently, she is principal investigator of the project “Picture Quality Management in Digital Video Broadcasting” financed by the Ministry of Science, Education and Sports of the Republic of Croatia. She is author or co-author of 16 papers published in scientific journals, more than 120 papers published in conference proceedings of international scientific conferences as well as of 15 reviewed studies and expert works. She was editor of 6 international conference proceedings. She is a member of IEEE and SMPTE.