# VIDEO DATABASE FOR FACE RECOGNITION

*P. Bambuch, T. Malach, J. Malach*

EBIS, spol. s r.o.

**Abstract**

**This paper deals with video sequences database design and assembly for face recognition system working under specific limiting conditions. Features used for recognition have to be insensitive to surroundings influences as scenario[1], scene content and illumination, face dimensions and pose, image resolution etc. Since current face recognition algorithms are partially independent on surroundings and it is not easy to increase substantially face recognition systems performance, alternative approach to performance increase is to limit general operating conditions – surroundings. This could be achieved by using training and testing databases close to expected operating condition of face recognition system.**

**Assembled database is intended for development and testing of face recognition system integrated in access control system. Video data were taken in an indoor environment by video surveillance cameras of various video resolutions. Database contains 5115 images of 275 human subjects.**

**Face detection is the first step in face recognition, thus face detection system influences performance of face recognition. Therefore assembled database has been used for various face detection algorithms evaluation. In this paper test results of Viola-Jones based face detector providing best results are presented. Test results show the influence of scenario and illuminating conditions on the performance of face detection system.**

## 1   Introduction

Essential part of face detection/recognition research is system testing and its methodology. Test credibility is significantly dependent on used database. There are several available face databases for training and testing of face detection/recognition algorithms at the time. However, contemporary face databases have limited range of variability of image data and do not represent the real surroundings sufficiently. The list of current face images databases is available to researchers at http://www.face-rec.org/databases. A lot of these databases were assembled mostly for laboratory testing.

This work is a part of research and development project IVECS[2] focused on video processing systems interconnecting CCTV (Closed Circuit TV) system and ACS (Access Control System). One of a few functions of the system is face recognition. The goal is to put into practice contemporary face recognition algorithm, in such a way to reach reliable operation of the system. Development of face recognition system working under general conditions is still challenging task. Therefore some limiting operating conditions – scenarios, illumination, face dimensions and image resolution have been set to improve performance under real conditions. Training and face database should represent real surroundings taking into account limiting operating conditions. A face database meeting project's needs is not available therefore CCTV system for video sequence capturing was designed and installed and adequate face database has been assembled.

In the second part of the paper the overview of available face databases with brief description is presented. Third part describes IFaViD - IVECS Face Video Database. More details on video sequences acquisition, limiting conditions and database structure are provided. Test results of Viola-Jones based face detection system [1, 2] for different scenarios and illumination types are presented in part four. The Viola-Jones detector was implemented within Matlab environment.

---

[1] Scenario is a time ordered sequence of human actions during video sequence capture.

[2] IVECS (Intelligent Video Modules for Entry Control Systems to Critical Infrastructure Facilities).

## 2 Current available face databases

Difficulties with face detection/recognition occur when dealing with testing methodology and test databases [3]. For employment of face detection/recognition method in real surroundings true performance evaluation is desired. Therefore test database must trustfully represent reality. The current available databases could be grouped for selection purposes by their properties.

First database group consists of artificially captured images where human's behavior is controlled and/or background of images is not complex enough to represent real surroundings. Face Place database (http://www.face-place.org), IMM Face Database (http://www2.imm.dtu.dk/~aam), Yale Face Database (http://cvc.yale.edu/projects/yalefaces/yalefaces.html) and others belong to this group. Performance evaluation based on such databases does not generally lead to usable results.

Second database group attempts to represent real surroundings and natural human's behavior. SCFace database (http://www.scface.org/) is large enough to produce reliable results, but faces captured on images are too small to be applicable for robust face recognition system. Another aspect is that human's behavior seems to be influenced by camera system, which was used for image acquisition. The Chokepoint face database (http://itee.uq.edu.au/~uqywong6/chokepoint.html) seems to meet basic requirements for algorithm's testing and trustworthy results may be expected in real surrounding. The Chokepoint database covers only one scenario and thus it does not meet IVECS project's needs. That's why decision to assemble new database IFaViD has been made.

## 3 IFaViD – database description

This part describes definition of limiting conditions for video sequences capturing. Database specification, image acquisition and ground truth specification (Labeling) are also included.

### 3.1 Structure

The name of project IVECS partly reveals its intended use. System IVECS should be applied in critical infrastructure facilities for authorized staff identification and tracking within facility areas. Thus recognition system has to be able identify only limited (hundreds) number of faces, simultaneously unauthorized (external, strange) faces must not be identified as authorized (internal, known) faces. Database has been assembled to verify this ability of system. Therefore the whole database is divided into two groups: internal persons and external persons. Internal persons are expected to have adequate representing pattern in face recognition system, to be well distinguishable among each other and among external persons. On the other hand, external persons won't have pattern so they should be classified as external, not corresponding to any pattern. To cover this, IFaViD contains 250 external persons. External persons are contained in only several (3 to 8 sequences per person) video sequences while internal persons of total number 25 are captured on many sequences (100 to 150 sequences per person) which represent real external/internal persons sequence occurrence ratio. Such a database should provide representative results for the IVECS project.

### 3.2 Limiting condition

Video processing system, being developed in this project, will be used in industrial environment. Due to this fact, system must be minimally dependent on surroundings and applicable in required scenarios. To maintain reliability of the system the scenario, illumination and face dimensions and resolution requirements have been set.

Scenarios naturally limit human's behavior and actions during video sequence capture, e.g. pose of head is limited to that extent enabling face detection/recognition. Three different scenarios have been chosen according to ACS requirements and are defined as follows:

Scenario A: person walking through door frame or corridor. Person's cooperation with the system is not required, but intentional obstructing is excluded (face coverage, head tilting).

Scenario B: person requesting closed door or gateway access via identification device. In this scenario, person is required to cooperate and interact with identification device in order to get access to requested area. Example of scenario B is in Fig. 1.

Scenario C: person standing/sitting in expected area in front of machine (PC or any other) or any other obstacle (reception desk, counter, check point). Person's cooperation is not required, but intentional obstructing is excluded (face coverage, head tilting).



**Fig. 1 Scenario B.**

The requirement for robust face recognition is sufficient amount of information which is maintained by minimal face dimension and resolution represented by intereye distance[3]. Intereye distance (measured in pixels) in IFaViD is set as a trade-off between standards' requirements, namely Information technology - Biometric data interchange formats - Part 5: Face image data[4] and requirements of leading commercial face recognition systems. Minimal intereye distance in IFaViD is set to 50 pixels. The images with smaller intereye distance are not included in IFaViD database.

Success of face recognition systems substantially depends on face illumination. To avoid difficulties with face recognition task, a face must be uniformly illuminated at appropriate intensity level without shadows and hot spots. Due to this fact the additional lighting sources (standard indoor lighting) have been installed at some less illuminated scenarios areas. Video sequences of scenarios were captured under variable illumination – side daylight combined with additional light source, overhead light, overhead white light combined with camera built-in IR light etc.

## 3.3 Specification

The specifications of IFaViD are defined as follows:

• Internal subjects/persons: 25.

• External subjects/persons: 250.

• Total images: 5115.

• Sex: women and men (1:9).

• Number of nationality: 1.

• Age: 18 to 65.

• Beard and moustache: yes.

• Glasses (transparent): yes.

• Illumination: artificial, daylight, infra-red, combinations.

The image's background is complex and represents real surrounding (e.g. other objects in the scene). Face expression of subjects is naturally various. Some faces with transparent glasses, beards and moustaches are included. Intereye distance for individual subjects spans from 50 up to 150 pixels.

---

[3] The distance between person's eyes.

[4] ČSN ISO/EIC 19794-5:2007.

Illumination is mostly uniform and well-scattered but to avoid lack of complexity some images with imperfect illumination are contained. Typical case is ceiling illuminator instead of ideal-front-face illuminator.

IFaViD examples of faces for defined scenarios are in Fig. 2.



Fig. 2 Example of one image data set per scenario.

## 3.4 Acquisition of video sequences

Database was created by two different types of IP surveillance cameras and one tablet computer camera:

• Axis M1114 used for video capturing in visible spectrum for scenario A and C.

• Vivotek IP 7361 built-in IR LEDs used for video capturing in infra-red spectrum for scenario A and C.

• Tablet computer camera was used for scenario B.

The surveillance IP cameras were consecutively mounted at five different indoor surroundings. All used surveillance IP cameras have HD (High Definition) resolution. The cameras were set up to various resolutions: 640×480, 800×600, 1024×768 and 1280×960 pixels (for detail see Tab. 1).

Tab. 1 Scenario and camera resolution.

| Scenario | A | | B | C | |
|---|---|---|---|---|---|
| Spectrum | visible | infra-red | visible | visible | infra-red |
| Resolution (px.) | 800×600 | 1280×960 | 640×480 | 1024×768 | 1280×960 |
| | 1024×768 | - | - | - | - |

All cameras were connected to NVR (Network Video Recorder) via Gigabit LAN (Local Area Network). Professional video management system has been used to capture video sequences. Recording started on motion detection event. Used acquisition CCTV system block scheme is in Fig.3.
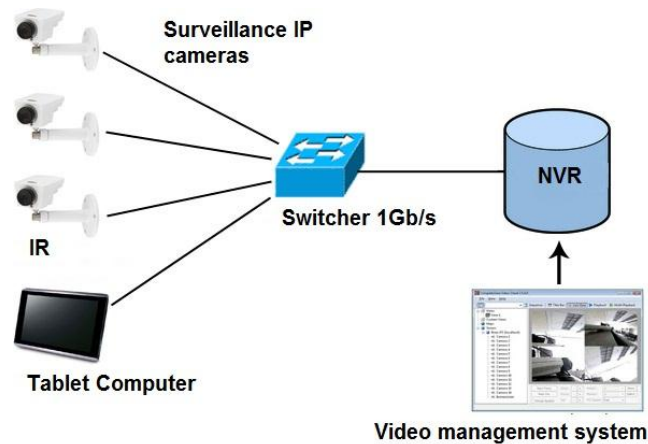


Fig. 3 CCTV block scheme.

## 3.5 Video sequences processing

A Subsequent task was to sort out captured video sequences. Useless video sequences (e.g. not well focused) were deleted, representative video sequences were split into frames - static images and some of them were chosen to trustworthily represent video sequence. Such images were grouped according to characteristic properties (e.g. scenario, illumination, person's ID) into small image sub-databases. These sub-databases were converted back to video sequences. All video sequences were manually labeled. Each video frame has been labeled with ground truth descriptions.

Ground truth (Labeling) contains information (in XML format) about images characteristic properties. Each video frame was manually labeled using freeware tool Event Editor[5]. Example of labeled frame extracted from scenario A video sequence shows Fig. 4. Highlighted objects determine coordinates of subject's face and the eyes in image. Detailed description of all ground truth is in [5].



**Fig.4 Labeled frame.**

## 4 Face detection testing

Very first test conducted using IFaViD was the performance evaluation of Viola-Jones based face detector under defined limiting operating conditions. Tested face detector is written in Matlab programming language. Cascade frontal face classifier was used.

Particular set up of the detector was:

• Minimal size of detection window: $85 \times 85$ pixels.

• Detection window scale factor: 1.2.

• Clustering: 3 neighbors for cluster analysis.

Face detector has been tested on IFaViD involving a total of 3589 images grabbed from 796 video sequences. TPR (True Positive Rate) defined in [4], is the only measure used for face detector performance evaluation for project's needs. There are two reasons: firstly - purpose of test is to unveil influence of real conditions on performance, not to examine ability to distinguish face and background since error rates are very low [4, 6]. Secondly - eventual false phantom faces - FP (false positives) defined in [4] will be eliminated by face recognition task as non faces.

Test results are shown at Fig. 5 which presents influence of defined scenarios and illumination type on TPR.

Results for scenarios A and B show very high TPR (TPR = 0.9982 for scenario A, TPR = 0.9672 for scenario B). Results for scenario C indicates significantly lower TPR (TPR = 0.5864). The reasons for such a TPR drop are high variability of person behavior in front of reception desk and imperfect (overhead) camera and illuminator location – camera and illuminator location is not allowed between receptionist and captured person.

---

[5] http://www.fit.vutbr.cz/research/grants/m4/editor/index.htm.cs.iso-8859-2.

Video sequences captured by camera with built-in IR illuminator for scenarios A and C show approximately the same drop of TPR, by circa 10%. This may suggest that capturing images using commercial camera build-in IR illuminator has negative effect on TPR.
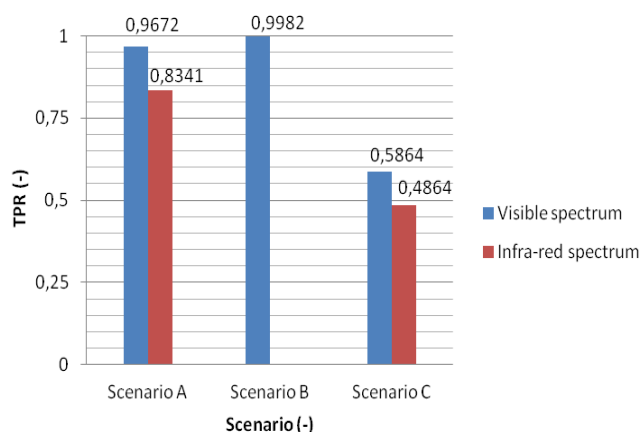


**Fig.5 Test Results showing influence of scenarios and illumination.**

## 5 Conclusion

This paper deals with face detection/recognition system testing databases. This paper tries to point out practical aspects of face detection/recognition system testing. Reaching of trustworthy results is the main target of this work.

Contemporary available databases description and suitability assessment for project's needs testing are presented. The assessment resulted in the need of a new database assembly that would complexly test face detection/recognition algorithms. IFaViD has been assembled from real surroundings video sequences. IFaViD contains 5115 images of 275 human subjects. Video data were taken in indoor surroundings by commercially available video surveillance cameras and one tablet computer camera.

IFaViD contains video sequences captured under limiting conditions – scenarios, illumination and face dimension and resolution conditions, applicable in ACS.

Test results for Viola-Jones based detector under limiting conditions using IFaViD are presented. Results proved significant influence of the way of scenario video sequence capturing on TPR. Results for video sequences captured by commercial camera with build-in IR illuminator show the decrease of TPR by circa 10% in comparison to daylight and white artificial illuminators.

IFaViD is to be used in next project's steps for face recognition algorithm development and testing.

## Acknowledgment

## References

[1] P. Viola, M. Jones. *Robust Real-Time Face Detection*. International Journal of Computer Vision, vol. 52, no. 2, pp. 137-154, 2004.
[2] T. Malach. *Detekce obličeje v obraze*. Bachelor's thesis, VUT, Brno University of Technology, Brno, 2011.
[3] N. Degtyarev, O. Seredin. *Comparative testing of face detection algorithms*. In Proceedings of the 4th international conference on Image and signal processing, Trois-Rivières, QC, Canada, 2010.

[4] M. Yang, D. Kriegman, N. Ahuja. *Detecting Faces in Images: A Survey*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp. 34-58, 2002.

[5] P. Bambuch. *Popis vytvořené obrazové databáze modulu FACE*. EBIS, spol. s r.o., archive no. 0021412, 2012.

[6] T. Malach, P. Bambuch, J. Malach. *Detekce obličeje v obraze s využitím prostředí MATLAB®*. In Proceedings of the 19th Annual Conference Technical Computing Prague 2011, ISBN 978-80-7080-794-1, pp. 78, 2011.

Petr Bambuch
EBIS, spol. s r. o.
Křižíkova 2962/70a
612 00 Brno
pbambuch@ebis.cz

Tobiáš Malach
EBIS, spol. s r. o.
Křižíkova 2962/70a
612 00 Brno
tmalach@ebis.cz

Jindřich Malach
EBIS, spol. s r. o.
Křižíkova 2962/70a
612 00 Brno
jmalach@ebis.cz