International Database of Facial Images for Performance and ISO/IEC 19794-5 Conformance Tests

Peter Ebinger and Margarida Castro Neves
Security Technology Dept., Fraunhofer Institute for Computer Graphics Research IGD
Fraunhoferstr. 5, 64283 Darmstadt, Germany
{peter.ebinger|margarida.castro-neves}@igd.fraunhofer.de

René Salamon
Federal Office for Information Security (BSI)
Godesberger Allee 185-189, 53175 Bonn, Germany
rene.salamon@bsi.bund.de

Helmut Seibert
Computer Graphics Center (ZGDV)
Rundeturmstr. 10, 64283 Darmstadt, Germany
helmut.seibert@zgdv.de

Abstract: Digital face images and fingerprint images are included in all newly issued EU passports [Bus06]. Face images stored within electronic passports (ePass) have to conform to ISO/IEC 19794-5 [Inta] which defines a number of quality requirements for facial image data. A software application (QS Software) is used to verify photographic image properties (pose, head/image size, height-to-page ratio, etc.) as well as setup properties (color depth, contrast, brightness, etc.) [Intb]. To ensure reliability and reproducibility of automated facial image processing systems it is necessary to assess the quality of existing software packages. To this end a database of facial image data would be of significant value as a testing base [HSM06].

The goal for project series FIReBIRD (Facial Image REcognition Benchmark Including Realistic Disturbances) is to establish a database of facial images that can be used worldwide for compliance and performance testing and quality verification based on ISO standards. FIReBIRD will be conducted in close cooperation of Fraunhofer IGD and the German Federal Office for Information Security (BSI) and will be carried out in parallel and in international cooperation with the U.S. National Institute of Standards and Technology (NIST) and the British National Physical Laboratory (NPL).

After successful completion of a conceptual framework the international database will be built. This facial image database will consist of distinct data sets used by two separate user groups for two different purposes: public test data that is available for algorithm developers and non-public test data for system evaluation accessible only to authorized organizations such as BSI/Fraunhofer IGD, NPL and NIST.

1 Motivation

Digital face imaging is a core function used by various applications for such services as human examination and automated face recognition, e.g. for access control. A standard data format and digital face image requirements were defined by the International Organization for Standardization (ISO) to allow interoperability among vendors.

Digital face images and fingerprint images are included in all newly issued EU passports [Bus06]. The International Civil Aviation Organization (ICAO) defined requirements for machine readable travel documents in its passport standard 9303 [Gro04, Int06]. These regulations are implemented by all member states of the European Union and by other nations to support border controls by means of biometric systems.

Biometric data stored within the electronic passport (ePass) has to conform to requirements of ISO SC37 data interchange format for biometric data. ISO/IEC 19794-5 [Inta] defines a number of quality requirements for facial image data which include scene constraints (lighting, pose, expression etc), photographic properties (positioning, camera focus etc) and digital image attributes (image resolution, image size etc) [Inta, Intb]. In Germany a similar set of facial image data requirements were defined on a national level [Buna, Bunb] based on the ISO standard.

Before the introduction of the ePass in Germany requirements for facial images used for governmental documents were checked using a sample photo table (Fotomustertafel) based on the specifications in [Pas07]. Submitted facial images were manually inspected by official clerks based on visual comparison of the sample photo table and geometric measurements were checked using a mask/template provided by the Bundesdruckerei. These tools provided only limited means to verify the international binding ISO standard [Inta] as decision results were based on subjective judgment and cannot necessarily be reproduced with consistency.

Meanwhile, quality assessment of submitted facial images is performed using precise geometric measurement based on a new version of the photo mask/template. The new mask/template takes extended geometric requirements of the ISO standards into account while a software application (QS Software) checks photographic image properties (pose, head/image size, height-to-page ratio, etc.) and setup properties (color depth, contrast, brightness, etc.) [Intb].

Verification of whether a specific facial image fulfils international ISO requirements should be automated to ensure process accuracy and image quality. Analysis on whether (and to what extent) products properly examine and evaluate image properties required by the standard and subsequent comparison of the evaluation results of the verification and identification process is needed in order to identify software which most closely conforms to process requirements.

To ensure reliability and reproducibility of automated facial image processing it is necessary to assess the quality of existing systems being used within practical applications. To this end a database of facial image data would be of significant value as a testing base. This image database should contain a collection of valid and invalid facial images (according to ISO standards and thus [Buna, Bunb]).

The database should allow program testing to determine whether generated image data records are a faithful representation of the subject with respect to relevant requirements and parameter constraints. It should be noted that such a collection of image data does currently not exist, at least not with the significant scope and quality required for this degree of testing.

2 FIReBIRD Objectives

The goal for project series FIReBIRD (Facial Image REcognition Benchmark Including Realistic Disturbances) is to establish a database of facial images that can be used worldwide for compliance testing and quality verification based on the ISO standards. The database will be provided and maintained by the following partners:

- German Federal Office for Information Security (BSI)/Fraunhofer IGD, Germany,
- U.S. National Institute of Standards and Technology (NIST), USA, and
- British National Physical Laboratory (NPL), UK.

FIReBIRD will be conducted in close cooperation of Fraunhofer IGD and BSI. It will be carried out in parallel and in international cooperation with the efforts of NIST and NPL.

The facial image database should have distinct data sets used by two separate user groups for two different purposes:

- Public test data that is available for algorithm developers. These data sets can be used for performance and conformance testing for the development of new algorithms.
- Non-public test data only accessible by authorized organizations such as BSI, Fraunhofer IGD, NPL and NIST. These data sets are used for the evaluation of developed algorithms and the implementation of facial image data processing software.

Procedures for building each database should be identical for all partners such that equivalent data quality can be achieved, providing a solid foundation of data for testing of algorithms and related implementations of facial image data quality assessment. The team of organizations will first agree on concept development for required image variations, data structures to be used and defined prototypical setups for the production of facial images. The goal is to complete the development of this structure by the end of 2008.

After successful completion of a conceptual framework it is planned to build the international image database in 2009. This data collection then should be made available (copied as required) to partner nations. If subsets of individual image databases are made available to the public (e.g. to manufacturers of face recognition systems) they will also be provided to partner organizations.

3 Preparatory and Fundamental/Basic Works

3.1 (National) (Privacy-related) Legal Basis/Requirement

In order to legally capture facial images, privacy-related legal requirements for data collection have to be verified in cooperation with the Federal Commissioner for Data Protection and Freedom of Information (BfDI). In particular it must be clarified whether and to what extent (e.g. for the duration of the data collection) additional personal data (such as first and last name, date of birth, place of residence/address, etc.) can/must be collected and stored (e.g. to exclude multiple captures of a specific person). In addition a formal declaration of consent which outlines image storage, usage and third party disclosure will be developed in cooperation with the BfDI to be signed by all volunteers during image acquisition.

3.2 Basic Considerations for the Acquisition of Facial Image Data

Based on the project objectives it will be determined to what extent a demographically representative composition of the facial image database is necessary or desirable and what personal data (such as gender, age, ...) should be stored in the database.

In addition, the basic concept for how the collection of the facial image data will be constructed in Germany will be specified. Currently there are three possible scenarios to be discussed:

1. "Central" collection of facial images data in one location by a single photographer
2. Distributed data collection in several locations in Germany by a single photographer ("mobile acquisition center")
3. "Distributed" capturing of image data at several locations by various local photographers.

3.3 Face Property Analysis

It is known that performance of human and in particular automatic face recognition is significantly impacted by non-biometric characteristics (such as "wearing glasses", "having a beard", "hair hanging down over the forehead," etc.). It is also expected that morphological characteristics such as certain facial or nasal forms, mouth line characteristics, various eye shapes, etc. also have an impact on the performance of detection systems. These properties will be added to the database as metadata for each facial image and identity respectively to prevent the uncontrolled influence of such facial characteristics on the result of evaluations that are performed based on the database that is to be created.

Based on the combination of the current state of anthropology knowledge and the results of the BioFace project series facial characteristics have to be identified and specified that allow to objectively and clearly distinguish different human face types.

4 Analysis of ISO Requirements

4.1 Determination of Relevant and Measurable Image Properties

Review of photographic image and scene properties defined in [Inta] is required with in respect to:

A how relevant image properties are for automatic face recognition (AFR) by 2D and/or 3D camera systems [BNZ+07, BNZ+08], i.e. how various images influence the quality of verification results (such that a false rejection or false acceptance occurs) or the performance of the identification process;

B how relevant image properties are for human face recognition (HFR), i.e. the recognizability of a face degraded that a false rejection or false acceptance occurs and whether property "a" and/or "b" are true;

C if image properties can be measured and/or computed in an automated fashion;

D additional image properties should be identified which are not currently in use for automatic nor human face recognition but may so in future (subject to the availability of more powerful face recognition systems, e.g. advanced 3D face recognition systems).

This study will also consider the experience of Fraunhofer IGD with facial recognition system (FRS) and the findings of the BioFace series of BSI. Thus the list of image properties defined in [Inta] and [Bunb] needs to be checked for completeness.

Tabular 1 shows the structure of the decision matrix. Image properties that will be considered as features include for example pose (roll, pitch and yaw), facial expression or exposure.

5 Specification of Test Conditions

5.1 Definition of the Ground Truth

For mandated image property a distinct and "general" reference point or benchmark will be specified and described. The generic term "general" is used within this scope to refer to a term that is independent of a specific implementation or software tool (e.g. information on compression strength is only allowed if it is not based on a specific parameter and/or

	A Relevant for Automatic Face Recognition (AFR)	B Relevant for Human Face Recognition (HFR)	C Measurable / Computable	D Relevant for future use	E Mentioned in ISO Standard [Inta]	F Mentioned in TR-PDÜ [Buna]
Feature 1	Y	N	Y	—	Y	Y
Feature 2	Y	N	N	—	Y	Y
Feature 3	N	Y	Y	—	N	Y
Feature 4	N	N	Y	Y	Y	Y
Feature 5	N	N	N	N	Y	Y
Feature 6	Y	N	Y	—	Y	N
...	...	...	...	...	...	...

Table 1: Decision Matrix for Relevant and Measurable Features of Facial Images

value range of a (software) tool that is used). If a general reference point can not be specified calibration images must be created which can be referenced by users of other (software) tools to map the value range of their (software) tool. As per restrictions on measurable and/or computable image properties, it should be assured that reference points are defined in a non-trivial way. If [Inta] and/or [Bunb] demand image properties that are not measurable or computable, this will be documented as notable restriction within the subsequent database description.

5.2 Specification of the Face Image Properties to be Considered and the Qualitative Extent of the Face Image Database

A list of image properties will be developed for consideration of use in the database based on the internationally approved decision matrix (cf. section 4.1). The final compilation will take into account which image properties can actually be generated and delivered. Based on this list image properties will be defined that will be included within the database to meet [Inta] and [Bunb] requirements. For each property variants for permissible ("good case", "valid image") and unauthorized ("bad case", "invalid image") images have to be specified and put in relation to the ground truth for the respective image property as defined in section 5.1. If no suitable images can be obtained for specific image properties appropriate recommendations should be developed to eliminate shortcomings in the mapping of ISO standards to facial images in the database.

5.3 Specification of Data and Communication Structures

Compilation of national image databases can achieve a large enough sample size of identities and ensure the widest possible spread of morphological characteristics within collected identities. In order to make this possible a common binding storage and interchange format for the facial image data is required for aggregation. For image types specified in section

5.2 an appropriate storage and exchange format for data and metadata will be developed. In addition identification and protection methods for the recorded images will be used that ensure the authenticity of recorded data for later verification (e.g. by electronic signatures or watermarking).

5.4 Development of Image Capturing Setup

An experimental arrangement will be setup which allows the production of all image types that were specified in section 5.2 by photographic means and/or with use of (software) tools. This setup should be reusable for the production of large amounts of data (a few thousand images) and implemented using suitable technical and financial resources, taking into account readily available technical equipment in mainstream photo studios. It must also be ensured that products used for information processing (software, additional hardware, etc.) are available for little or no cost (including all purchase, lease or licensing costs).

The experimental setup includes means for recording 3D scans [BNZ+07, BNZ+08] of each test individual with a 3D camera system showing the head in a neutral pose (cf. section 5.1). This 3D data set is then mapped within the database to the 2D image data of the same test individual. For 3D images an appropriate 3D camera system will be chosen that provides both the 3D model as well as texture information in sufficient quality. Information regarding camera recording accuracy (e.g. resolution) will be added to the image database as metadata.

6 Proof of Concept

This section outlines the proof of concept of the developed image capturing setup. Reassessment of the practical application of developed concepts may require a revision of concepts according to the ISO standards and may result some additional practical tests for quality improvement.

6.1 Setup: Operational and Functional Capabilities

An instance of the specified image capturing setup will be implemented and put into operation. The operational capability of the setup should prove that the facial images of 10 different people can be created according to section 5.2. The usability and robustness of the setup in operation will be investigated and assessed.

6.2 Setup: Completeness

For all image properties and types of section 5.2 it will be demonstrated that the implemented setup is capable of recording all facial data required for the FIReBIRD database. In particular it will be shown that the specified ground truth can be met or at a minimum that variations are within acceptable tolerances. The setup will also be analyzed and evaluated with respect to measurement inaccuracies and errors.

6.3 Test Conditions: Feasibility Analysis / Revision

The database concept and setup will be adjusted based on observations and test results from sections 6.1 and 6.2 in order to exclude measurement errors and inaccuracies that can be largely excluded (thus ensuring a smooth start to practical use).

7 Adjustment of Regulation Documents

Regulation documents [Inta] and [Bunb] should not require image properties that do not allow automatic measurement and/or computation otherwise automation of quality assurance and error-free processing of facial image data can not be guaranteed.

As software based quality assurance of facial images used in sovereign European and in international electronic government documents is the declared goal, in all cases where an unmeasurable or incomputable property is included in the standard additional measures must be taken (see table in section 4.1). This is the case for all entries where column C (Measurable/Computable) of the decision matrix contains an "N" (no) and least one of columns A (Relevant for AFR) and B (Relevant for HFR) contains a "Y" (yes). It must be ensured that the Information or requirements in [Inta] and [Bunb] are modified such that the required image properties are measurable or computable, or additional guidelines and instructions are inserted into the standards that must be adhered to if an image property should be assessed that can not be measured.

If no "Y" was entered in columns A - C or an unquestionable reformulation of relevant paragraphs in regulation documents is not possible, the requirements should be labeled as optional or even deleted.

The goal is to propose reasonable and practical corrections to the regulation documents [Inta] and [Bunb] or subsequent documents on the basis of experience with biometry based verification of public documents.

8 Conclusion

Verification of whether a specific facial image fulfils international ISO requirements [Inta] should be automated to ensure process accuracy and image quality. Therefore a database of facial image data is needed as a testing base. The image database should contain a collection of valid and invalid facial images (according to ISO standards and thus [Buna, Bunb]) and allow program testing to determine whether generated image data records are a faithful representation of the subject. Such a collection of image data does currently not exist, at least not with the significant scope and quality required for this degree of testing.

Therefore the goal for project series FIReBIRD (Facial Image REcognition Benchmark Including Realistic Disturbances) is to establish a database of facial images that can be used worldwide for compliance testing and quality verification based on ISO standards. FIReBIRD will be carried out in parallel and in cooperation with NIST and NPL.

In cooperation with the Federal Commissioner for Data Protection and Freedom of Information (BfDI) privacy-related legal requirements for data collection will be verified. The photographic image and scene properties defined in [Inta] will be reviewed and the relevant and measurable image properties will be determined. It will be evaluated how relevant image properties are for automatic and human face recognition and whether specific image properties can be measured and/or computed in an automated fashion.

Performance of automatic face recognition systems is significantly impacted by non-biometric characteristics and it is also expected that morphological characteristics also have an impact on the performance of detection systems. Therefore these properties will be added to the database as metadata for each facial image and identity respectively.

The facial image database will contain public test data that is available for algorithm developers as well as non-public test data only accessible by authorized organizations that will be used for the evaluation of developed algorithms and the implementation of facial image data processing software.

References

[BNZ+07] Christoph Busch, Alexander Nouak, Xuebing Zhou, Farzin Deravi, Michiel van der Veen, and Jean-Marc Suchier. Towards Unattended and Privacy Protected Border Control. In *Proceedings of Biometrics Consortium Conference*, 2007.

[BNZ+08] Christoph Busch, Alexander Nouak, Xuebing Zhou, Jean-Marc Suchier, Emile Kelkboom, and Tom Kevenaar. 3D Face Recognition for Unattended Border Control. In *Proceedings Net-ID 2008 Conference*, 2008.

[Buna] Bundesamt für Sicherheit in der Informationstechnik. BSI TR-03104 – Technische Richtlinie zur Produktionsdatenerfassung, -qualitätsprüfung und -übermittlung für Pässe, Version 2.1.1.

[Bunb] Bundesamt für Sicherheit in der Informationstechnik. BSI TR-03104 Annex 1 (QS-Gesicht) – Technische Richtlinie zur Produktionsdatenerfassung, -qualitätsprüfung und

-übermittlung für Pässe, Version 2.1.1 – Qualitätsanforderungen bei der Erfassung und Übertragung der Lichtbilder als biometrische Merkmale für elektronische Pässe.

[Bus06] Christoph Busch. Facing the future of biometrics. *EMBO Reports*, vol. 7:23–25, 2006.

[Gro04] International Civil Aviation Organization Technical Advisory Group 15 Machine Readable Travel Documents/New Technologies Working Group. Biometrics Deployment of Machine Readable Travel Documents, Version 2.0, May 2004.

[HSM06] Rein-Lien Vincent Hsu, Jidnya Shah, and Brian Martin. Quality Assessment of Facial Images. In *Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pages 1–6, Baltimore, MD, Sept 2006.

[Inta] International Organization for Standardization. ISO/IEC 19794-5:2005 Information Technology – Biometric data interchange formats – Part 5: Face Image Data.

[Intb] International Organization for Standardization. ISO/IEC IS19794-5:2005, Information Technology – Biometric data interchange formats – Part 5: Face Image Data – Amendment 1 – Conditions for Taking Photographs for Face Image Data.

[Int06] International Organization for Standardization. ISO/IEC TC JTC1 SC17, Supplement to Doc9303-part 1-sixth edition, June 2006.

[Pas07] Verordnung zur Durchführung des Passgesetzes. Bundesgesetzblatt, 2007. Teil I Nr. 52.