# Generating Ground Truth Data for Fingerprint Minutia Conformance Testing

Christoph Busch
2010-04-16

This document describes the purpose for generating ground truth data for fingerprint minutia and explains the requirements for the procedures in this effort.

## 1 Background Information

### 1.1 Conformance Testing

Since the completion of the ISO Minutia Interoperability Standard IS 19794-2 [2] the ISO/IEC JTC1 WG3 has considered conformance testing for this standard. Currently a conformance testing standard is developed as 29109-2 [3], that categorizes three different conformance testing levels:

1. *Level: Basic Data Field Testing*
   All data fields exist properly (e.g.in the correct encoding.)
   ISO Definition: Conformance testing methodology that checks field by field and byte by byte conformance with the specification of the Biometric Data Interchange Record as specified in the base standard, both in terms of fields included and the ranges of the values in those fields.
   NOTE: This type of testing tests syntactic requirements of the base standard.

2. *Level: Internal Consistency Testing*
   All data fields are filled with meaningful values and the fields are internally consistent.
   ISO Definition: Conformance testing methodology that tests the internal consistency of the Biometric Data Interchange Record under test, relating values from one part or field of the Biometric Data Interchange Record to values from other parts or fields of the Biometric Data Interchange Record.
   NOTE: This type of testing tests syntactic requirements of the base standard

3. *Level: Semantic Testing*
   The extracted features (minutia coordinates) are within tolerances bounds to real minutia coordinates.
   ISO Definition: Conformance testing methodology that tests that a Biometric Data Interchange Record produced by an Implementation Under Test is a faithful representation of the Input Biometric Data Record subject to the constraints of the parameters in the metadata records[4].
   NOTE: This type of testing tests semantic requirements of the base standard

### 1.2 Goal of Generating Ground Truth Data

The driving motivation for this activity to generate ground truth data was to conduct Level 3 conformance testing on a reilable database[1][5].

Several additional benefits can be achieved with this ground truth data. E.g. a correlation of human sample quality assessment to NFIQ sample quality assessment could be measured[6]. Furthermore a training system for dactyloscopic experts could be designed that visually displays true minutiae to a trainee - after he has place his measurement.

### 1.3 Expert Crew

The dactyloscopic experts of the German Federal Criminal Police Office (BKA) are kindly volunteering to provide support and place real minutia data in approx. 2000 images (1000 image pairs).
The expert crew will consist of 3 dactyloscopic experts. Those experts are members of the quality assurance team and thus highly experienced. The team will start working on the data in January 2009. As of June 2009 more than 300 images have been processed. Further interests in contributing to the data base has been shown from the Netherlands, the Czech Republic and Australia.

### 1.4 Image Database

The fingerprint images needed to generate ground truth data has kindly be provided by the National Institute of Standards and Technology (NIST). The fingerprint image data was selected from the special databases at NIST namely SD14 (all rolled data and mostly ink with few live scanned images) and SD29 (flat data /plain impression but all ink). The selection was done in June 2008 and targeted avoiding any systematic effects stemming from the way that this data was collected. Fingerprint images selected should equally represent male and female data subjects. Furthermore the images should represent fingerprint image data at different image quality levels according to NFIQ. (There might not get be many NFIQ=5 as we'd like since NFIQ=5 usually constitutes about 1.5 percent of the population)
NIST has provided a 5000 image pair selection. The images are stored in five sets on five CDs with equal characteristics, such that 1000, 2000 or 5000 image pairs could be processed - depending on the personal resources and the effective time consumption.
Furthermore is planned to generate a second database with latent forensic prints and a third database with live scan prints.

### 1.5 Minutia Placement

In order to minimize the impact of the human factor in generating the ground truth data, three dactyloscopic experts will analyze the image data independently. Each expert will specify minutia, core and deltas and correspondant assurance levels. These measurements are stored in a simple ASCII file (*.gtm).

### 1.6 Computing gt-minutia

Later on ground truth minutia coordinate, angles and assurance level (in the remainder of this text this will be referred to as "gt-minutia") will be computed as the mean (or other statistics as nearest neighbour) of all 3 expert placements, where the variance of the coordinates can indicate further trustworthiness of this gt-minutia.

## 2 Process Requirements

In order to minimize systematic effects in the ground truth data the expert crew is kindly asked to respect the following requirements.

1. Anonymity: Each expert should be given an expert ID that is not known to his colleagues. The expert ID is embedded in the filename storing the generated gt-minutia data.

2. Independent operation: The experts conduct the gt-minutia placements independent and are not influenced by their colleagues.

3. Avoid AFIS impact: There should be no impact of any automatic or semi-automatic minutia extraction functionality of any AFIS system (as those minutia extraction algorithms are what Level 3 conformance testing is influencing at most).

It should be documented how examiners mark minutiae, and what rule(s) they follow. Are they trained to mark a ridge ending minutia as skeleton ending or as valley bifurcation?

## 3 Ground Truth Data

### 3.1 Working Environment

The working environment should be familiar to the experts. However any support of automatic AFIS functionality must be avoided. Furthermore it is not possible to export and store data records from the AFIS that would include all the relevant data fields.
Thus it was decided that an independent software should be generated. Expert should first identify and mark each minutia and angle on an image followed assessing the various quality assurance levels of marked minutiae.

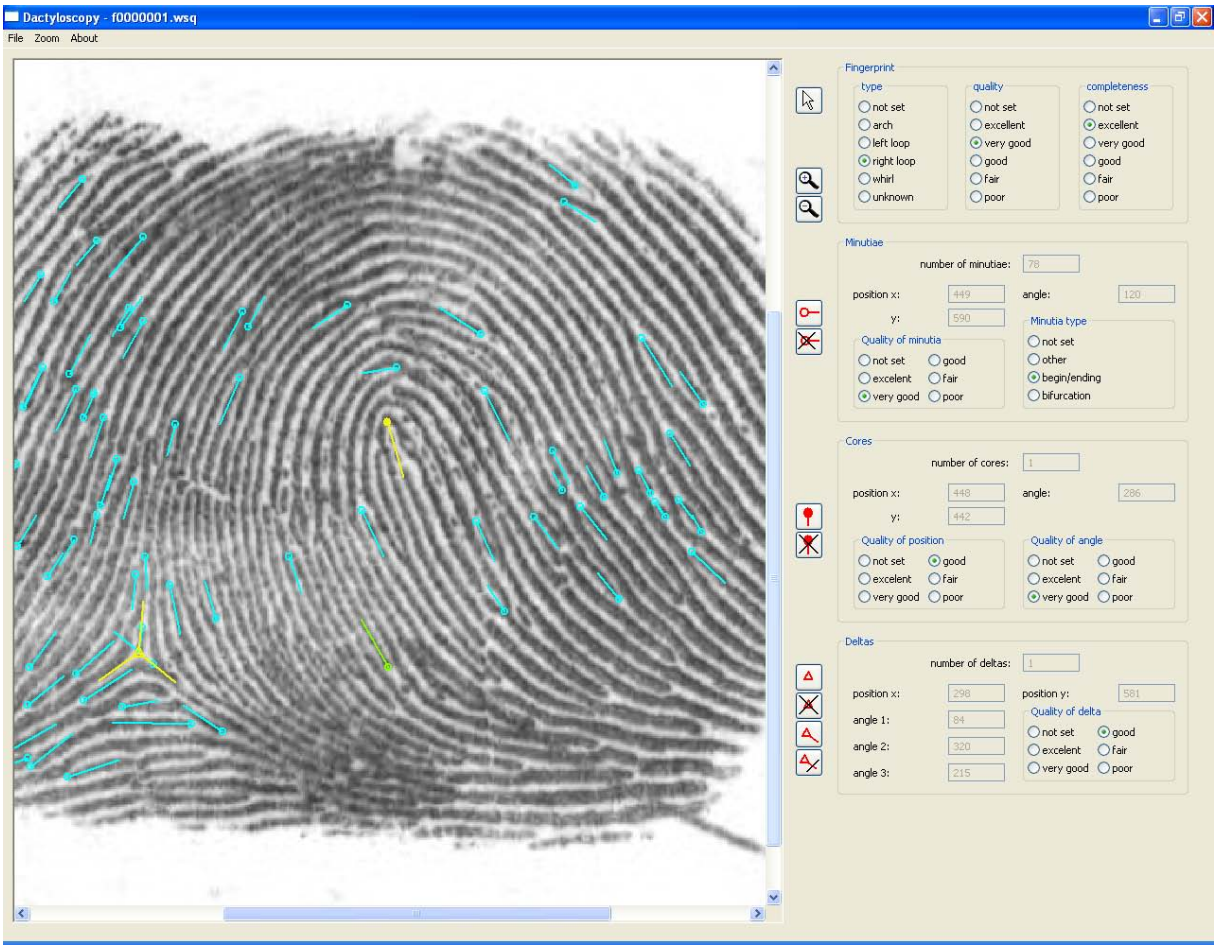


**Figure 1. Graphical User Interface to place real minutiae, delta and core and associated meta data.**

Such software has been developed at GUC. See figure 1 for a sample screenshot. This software runs on WindowsXP . The software including source code will be hosted on the NIST open source server.

### 3.2 Ground Truth Data Format

The data is stored in a simple text file. The filename carries the ID number of the dactyloscopic expert.

### 3.3 Relevant Data Fields

The following information will be stored.

1. *Pattern type* 1st Level classification will be done according to the following Fingerprint Classification Codes:
   A = Arch
   L = Left Loop
   R = Right Loop
   W = Whorl
   U = Unknown

2. *Sample quality level* Level of difficulty to analyze in correspondence to NFIQ ranging from 1 “excellent", 2 “very good", 3 “good", 4 “fair" down to 5 “poor".

3. *Sample completeness level* Level of difficulty to analyze in correspondence to NFIQ ranging from 1 “excellent", 2 “very good", 3 “good", 4 “fair" down to 5 “poor".

4. *Minutia type*: ISO 19794-2 Clause 7.4.2.1
   Undetermined, ridge ending, ridge bifurcation. The bits *00* will represent a minutia of *other* type, *01* will represent a ridge ending and *10* will represent a ridge bifurcation. Experts trained to classify minutiae according to forensic schemes (eg. begin, end, island, etc.) should see chapter 3.4 for a conversion table.

5. *Minutia Position*: ISO 19794-2 Clause 7.4.2.2
   The coordinates of the minutia ( horizontal X and vertical Y).The origin of the coordinate system is, as in ISO/IEC 19794-2, the upper left corner and the X and the Y coordinates shall be in pixel units.

6. *Minutia Angle*: ISO 19794-2 Clause 7.4.2.3
   Absolute angle of the minutia. The angle is measured counter-clockwise from the horizontal axis to the right. The angle of a minutia is in degree. The granularity shall be, as in ISO/IEC 19794-2 of (360/256) degrees. The value shall be a non-negative value between 0 and 255, inclusive.

7. *Minutia Quality*: ISO 19794-2 Clause 7.4.2.4
   The quality (accuracy) figure shall assess both position and angle and range from 100 as a maximum to 1 as a minimum.

8. *Number of Cores*: ISO 19794-2 Clause 7.5.3.1
   The number of core points represented.

9. *Core Position*: ISO 19794-2 Clause 7.5.3.3
   The coordinates of the core ( horizontal X and vertical Y).

10. *Core Position Quality* The quality (accuracy) figure shall range from 100 as a maximum to 1 as a minimum.

11. *Core Angle*: ISO 19794-2 Clause 7.5.3.4
    The angle of the core shall be recorded in one byte in units of 1.40625 (360/256) degrees. The core angle is measured increasing counter-clockwise starting from the horizontal axis to the right. The value shall be a non-negative value between 0 and 255, inclusive. For example, an angle value of 16 represents 22.5 degrees.

12. *Core Angle Quality* The quality (accuracy) figure shall range from 100 as a maximum to 1 as a minimum.

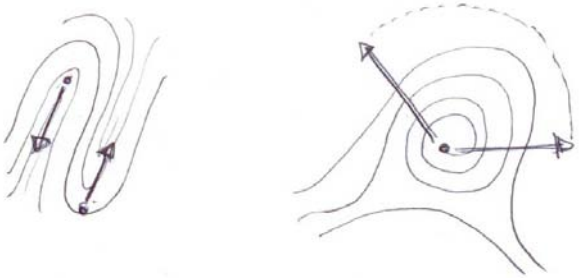

**Figure 2. Left:whorl with two cores, where position and angle have good quality scores. Right:whorl with one core, where position has good quality score but angle has poor quality score.**

13. *Number of Deltas*: ISO 19794-2 Clause 7.5.3.5
    The number of delta points represented.

14. *Delta Position*: ISO 19794-2 Clause 7.5.3.7
    The coordinates of the delta ( horizontal X and vertical Y).

15. *Delta Angle*: ISO 19794-2 Clause 7.5.3.8
    The three angle attributes of the delta shall each be recorded. The delta angle is measured increasing counter-clockwise starting from the horizontal axis to the right. The value shall be a non-negative value between 0 and 255, inclusive. For example, an angle value of 16 represents 22.5 degrees. If not all three angles can be extracted from the image because of noise or image cropping, the angle fields affected shall be filled by repeating any of the other angle(s) for the same delta.

16. *Delta Quality* The quality (accuracy) figure shall assess both position and angle and range from 100 as a maximum to 1 as a minimum.

### 3.4 Conversion Table

The ground truth data can only encode three different minutia types that have been defined by ISO namely ridge ending, ridge bifurcation and other. Therefor various minutia types that are common in the dactyloscopic literature have to be translated to ISO coding. The following sections will provide examples.

#### 3.4.1 Rdige Ending

Ridge endings should be labeled with the radio button "begin/ending" in the GUI. The following figures all represent ridg endings.

**Figure 3. Left: begin Right: ending**

**Figure 4. Left: embedded line Right: evasive endings**

**Figure 5. Left: hook Right: two endings**

### 3.4.2 Rdige Bifurcation

Ridge bifurcations should be labeled with the radio button "bifurcation" in the GUI. The following figures all represent bifurcations.

**Figure 6. Left: bifurcation Right: hook**

**Figure 7. Left: eye Right: island**

**Figure 8. ridge ramification**

### 3.4.3 Other

All other minutia that are neither endings nor bifurcations (and neither core or delta) should be labeled with the radio button "other" in the GUI. The following figure represent the minutia type other.

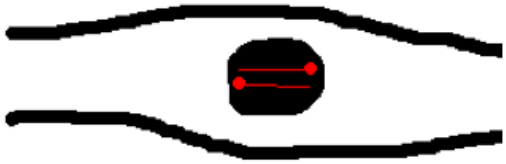

**Figure 9. singular point**

A short line is considered a singular point if the length is less than twice width of the neighbor ridge. Thus this short line should be marked as other. The angle should be set according to the ridge orientation (in the vicinity). In the centre of a delta there should no singular point be marked (use the delta in the GUI instead).

## 3.5 Use of Ground Truth Data

The ground truth data will be be used primarily for level 3 conformance testing. The overall database will be split in two separated fractions. It is intended that the mix between live scan (but roll) and plain (but inked) should be approx. the same in the public as in the sequester fraction.
The *public* fraction should contain approx. 30 percent of the total data and will be available in the public domain in order to be used by any interested body of the biometric community. Thus this fraction will be available to the industrial sector.
The *sequester* fraction should contain approx. 70 percent of the total data and will be provided under strong restriction to testing institution only, which are involved in conformance testing - be it in the academic sector or in the governmental sector. It is essential that the sequester fraction is never made available to stakeholders of the industrial sector in order to avoid that commercial algorithms are tuned to the data. For the long-term it is expected that ISO/IEC JTC1 SC37 WG3 will establish a procedure to autorize testing institution to receive a copy of the sequester database. For the short-term Patrick Grother and Christoph Busch will jointly decide, which institution should get a license. Candidates are NIST(US), NPL(UK), Fraunhofer-IGD(DE), GUC (NO).

## References


[1] C. Busch, D. Lodrova, E. Tabassi, and W. Krodel. Semantic conformance testing for finger minutiae data. In *Proceedings of IEEE IWSCN 2009*, pages pp. 17–23. IEEE, May 2009.

[2] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 19794-2:2005 Information Technology - Biometric Data Interchange Formats - Part 2: Finger Minutiae Data - Technical Corrigendum 1*. International Organization for Standardization, Jan. 2008.

[3] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC FCD 29109-2 Information Technology - Conformance Testing Methodology for Biometric Interchange Formats defined in ISO/IEC 19794 – Part 2: Finger Minutiae Data*. International Organization for Standardization, Feb. 2009.

[4] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC FDIS 29109-1 Information Technology - Conformance Testing Methodology for Biometric Interchange Formats defined in ISO/IEC 19794 – Part 1: Generalized Conformance Testing Methodology*. International Organization for Standardization, Feb. 2009.

[5] D. Lodrova, C. Busch, E. Tabassi, W. Krodel, and M. Drahansky. Semantic conformance testing methodology for finger minutiae data. In *Proceedings of BIOSIG2009*, pages pp. 31–42. GI LNI, September 2009.

[6] NIST. Nist biometric image software (nbis).