Biometric System Laboratory
University of Bologna - ITALY

http://biolab.csr.unibo.it

Using synthetic data to evaluate the quality of minutiae extraction algorithms
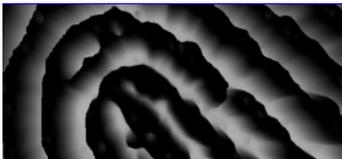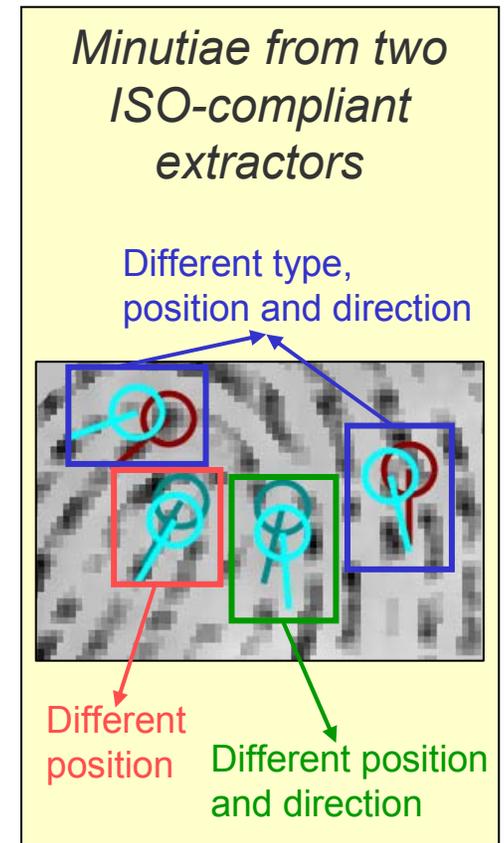
Raffaele Cappelli

Outline

- How to test semantic conformance to minutiae standard formats?

- SFinGe: a synthetic fingerprint generator
  - Generating images with ground truth minutiae

- Experiments
  - The impact of fingerprint quality on feature extraction errors
  - Analysis of minutiae placement errors
  - Correction of minutiae placement to improve interoperability
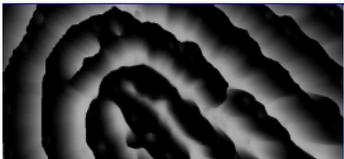
# Semantic conformance to minutiae standard formats?

- Recent interoperability projects (MINEX, MTIT) showed that:
  - Syntactic conformance is not enough for interoperability
  - Some template generators are semantically non-conformant
    - Such non-conformance degrades interoperability
  - Extraction algorithm standardization *should embed testing*
- The INCITS 378 and the ISO/IEC 19794-2 standards contain definitions of minutiae points and rules for their placement
  - Revisions of the two standards are being proposed with *refined guidance on minutia placement*
- How to test *semantic conformance* to data format standards?
  - i.e., is a given record an accurate representation of the true fingerprint minutiae?
  - A lot of data (with ground truth minutiae!) is needed
  - How to ensure the test data is error free (correct ground truth)?
  - How to deal with data protection / privacy issues?
- Does the *fingerprint quality* affect the semantic conformance of feature extraction algorithms?

Minutiae from two ISO-compliant extractors

Different type, position and direction

Different position

Different position and direction

SFinGe – A Synthetic Fingerprint Generator

- SFinGe (the Italian for Sphinx, pron. *sphin-ge*)
  - A software able to synthetically (randomly) generate large databases of realistic fingerprint images
  - Used (and validated!) in FVC2000, FVC2002, FVC2004, and FVC2006
  - Many parameters can be adjusted to control the fingerprint generation and simulate different acquisition devices/conditions:
    - skin conditions, finger placement, skin deformation, noise, scanner background...
  - Several research groups and companies are currently using SFinGe to:
    - compare different fingerprint matching algorithms
    - train pattern recognition techniques that require large learning-sets (e.g. neural network, PCA,...)
    - easily generate a large number of "virtual users" to develop and test medium/large-scale fingerprint-based systems

- References:
  - [Book Chapter] R. Cappelli,"Synthetic fingerprint generation", in D. Maltoni, D. Maio, A.K. Jain and S. Prabhakar, *Handbook of Fingerprint Recognition*, Springer (New York), 2003.
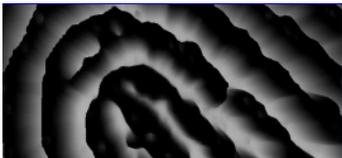
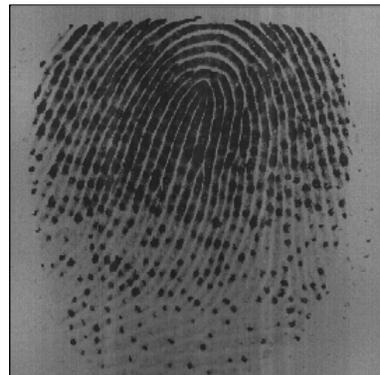Using synthetic data to evaluate the quality of minutiae extraction algorithms

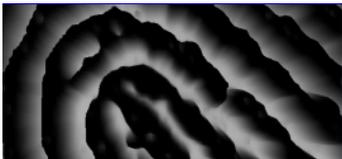# SFinGe - Examples

Real   Synthetic
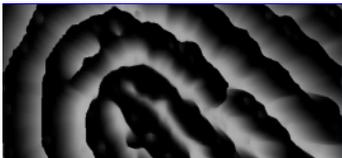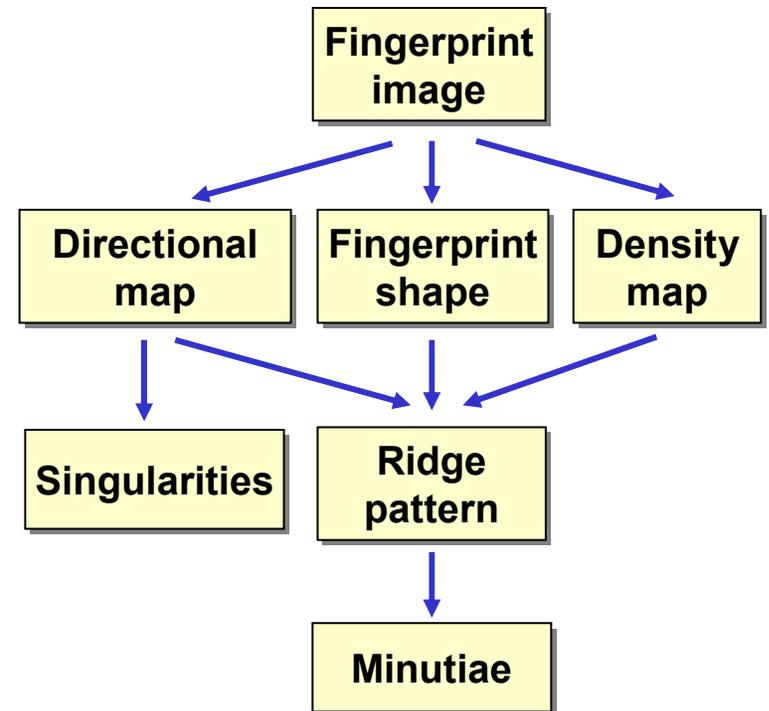
Real   Synthetic

Real   Synthetic

Typical feature extraction steps (from a real fingerprint)
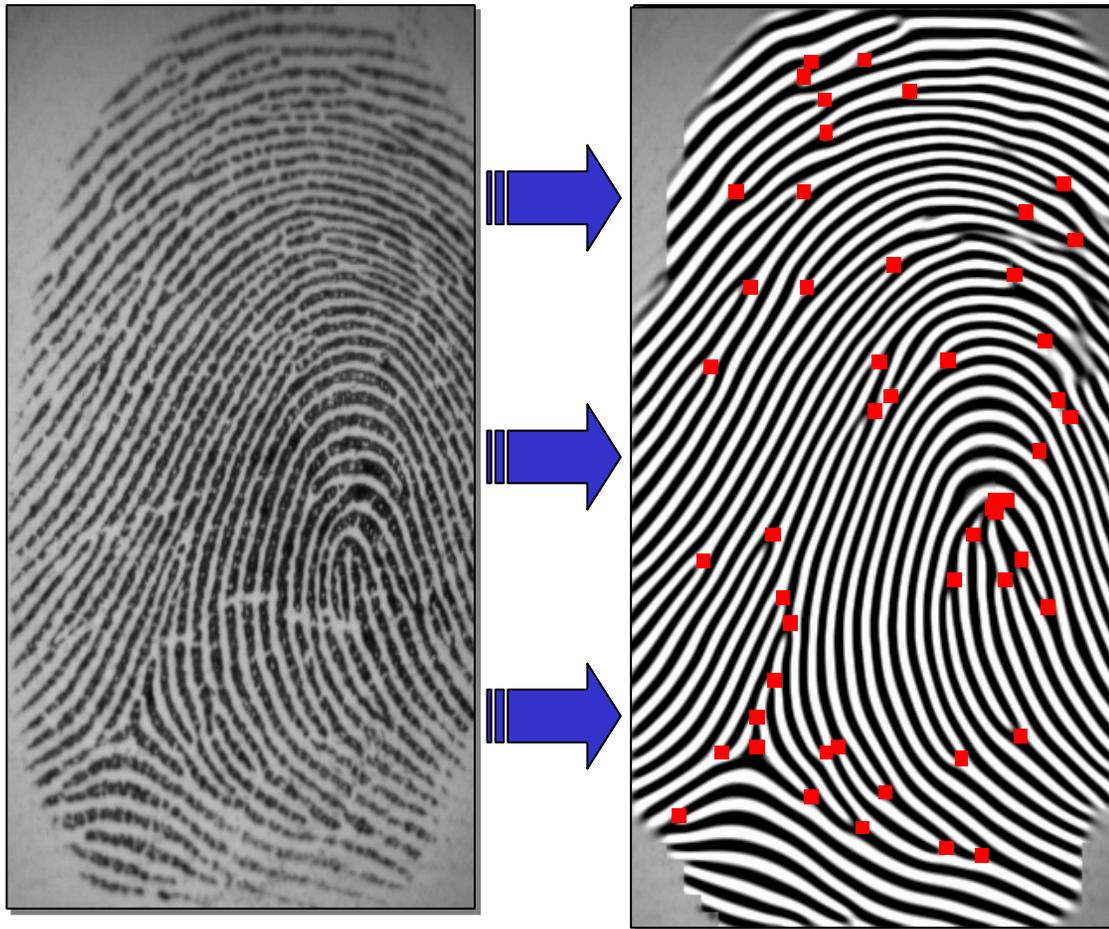
How SFinGe works (1)


```
Singularities
     │
     ▼
Directional     Fingerprint     Density
   map            shape           map
     │              │              │
     └──────────────┼──────────────┘
                    ▼
                 Ridge
                pattern
                    │
                    ▼
              Fingerprint
                 image
```
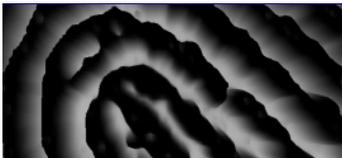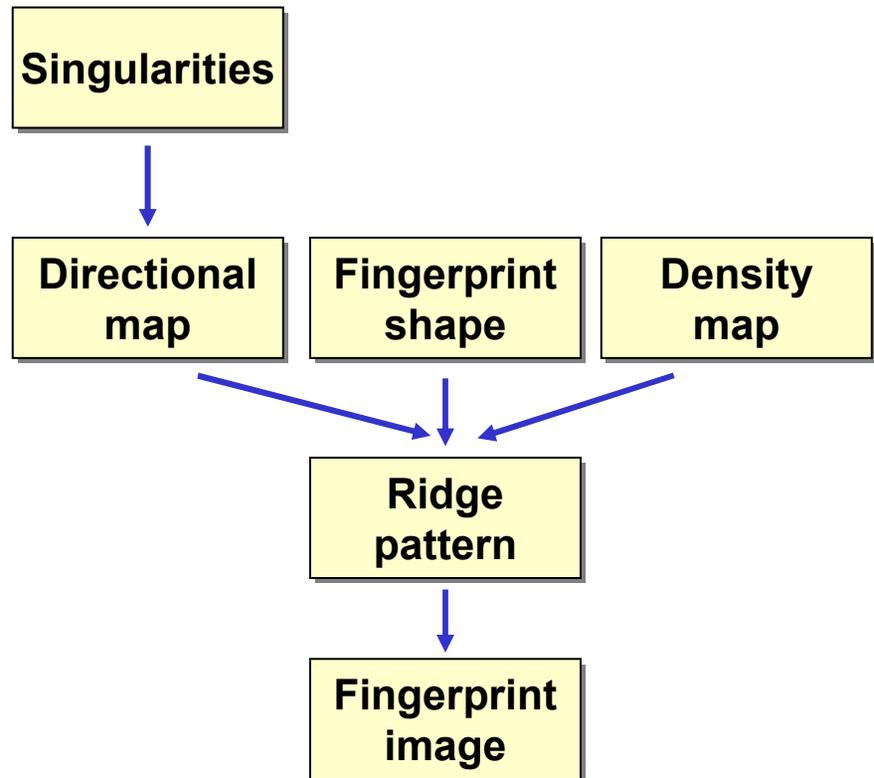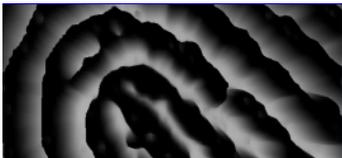
How SFinGe works (2)

Using synthetic data to evaluate the quality of minutiae extraction algorithms

7
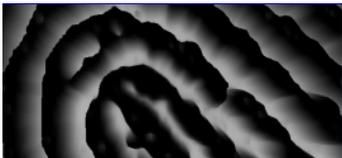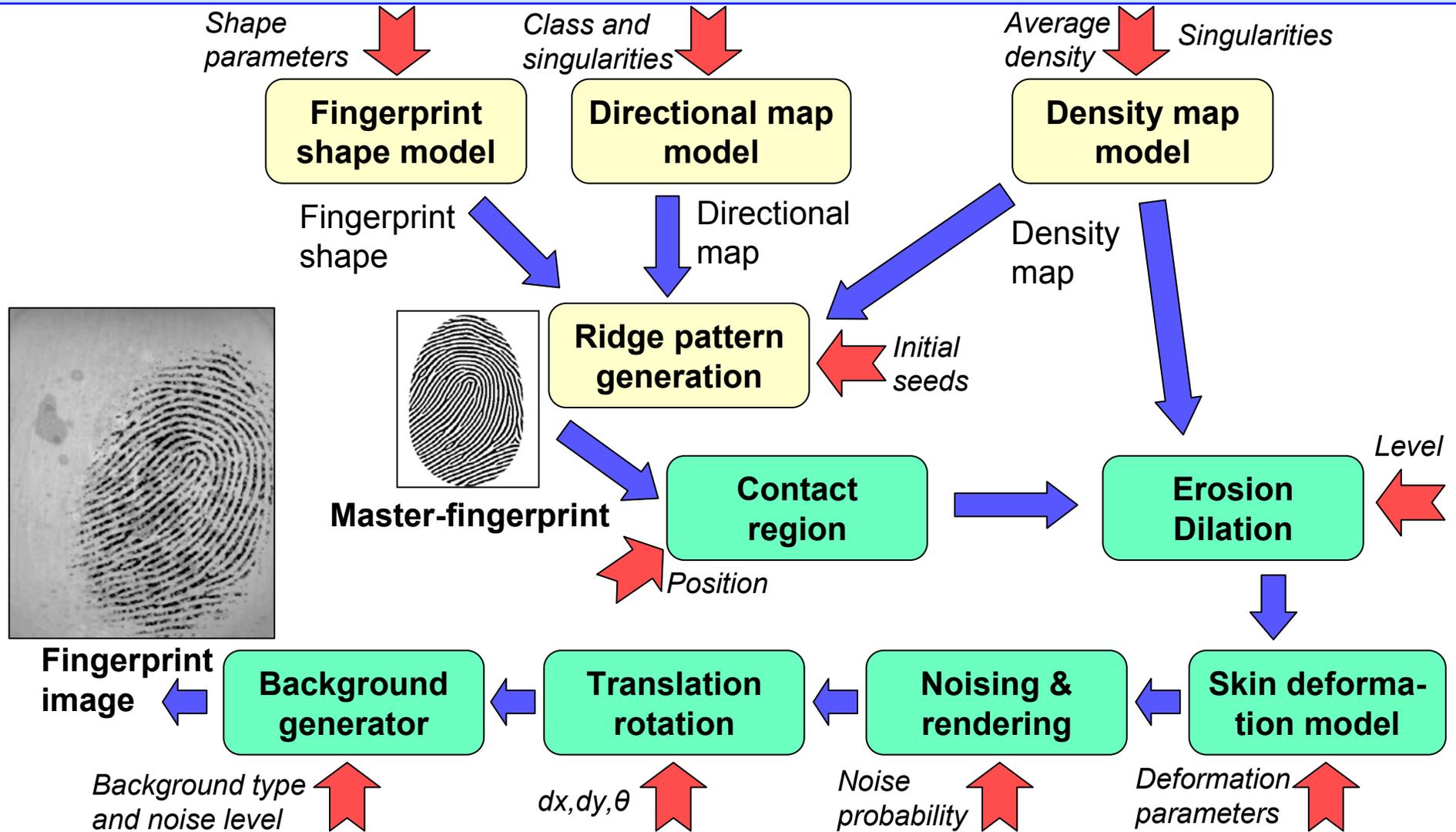
The "master fingerprint": a "perfect" binary ridge pattern

Iteratively created by applying Gabor filters whose orientation and frequency are locally adjusted according to the directional and density maps.
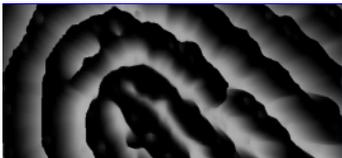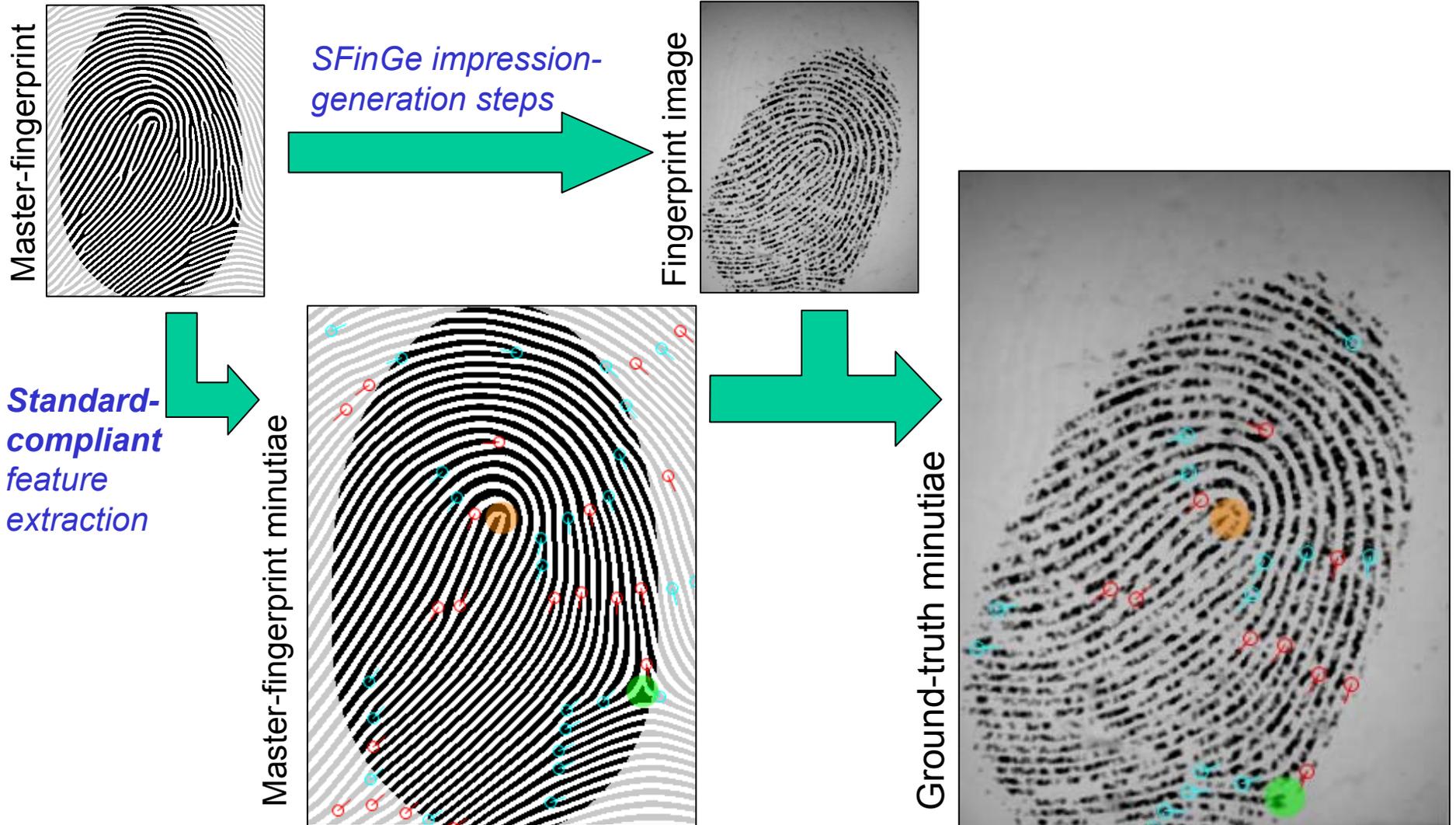
Realistic minutiae appear at random positions

SfinGe "master fingerprints" are well-suited for applying the precise minutiae extraction procedures that are being proposed as ANSI and ISO standards.

Automatic generation of the *ground-truth*

Master-fingerprint

SFinGe impression-generation steps

Fingerprint image

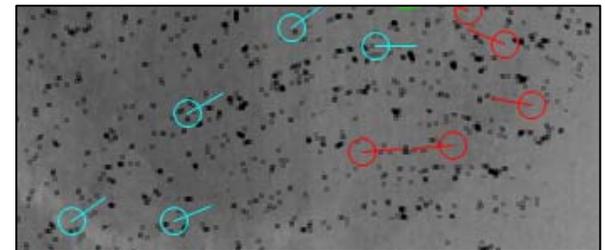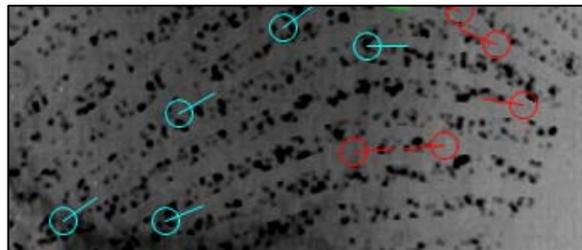***Standard-compliant*** *feature extraction*

Master-fingerprint minutiae

Ground-truth minutiae

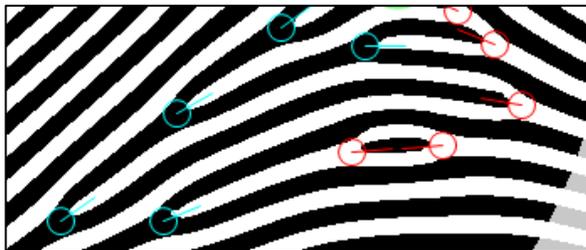Using synthetic data to evaluate the quality of minutiae extraction algorithms

Some advantages

- Automatic generation of large fingerprint databases with ground-truth minutiae
  - Features can be extracted by applying the standard procedures easily and without ambiguities (extraction occurs on a binary image without noise)
- The main fingerprint characteristics can be controlled
  - e.g. Fingerprint class, ridge line density, finger placement, skin distortion, fingerprint quality, ...
  - Datasets to test the impact of a given parameter (e.g. fingerprint quality) can be easily generated
- The ground truth is always unique and sound, even when the quality of the final image is very low
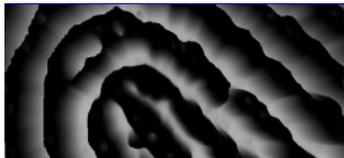
Using synthetic data to evaluate the quality of minutiae extraction algorithms

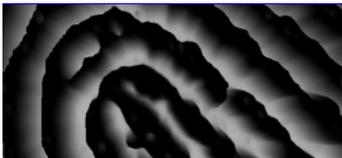# Evaluating minutiae extractors quality

- Datasets:
  - Dataset 0: images generated by SFinGe with no noise (binary images – maximum quality)
  - Datasets 1,2,3,4,5: obtained from a larger dataset generated by SFinGe (with background and noise), by partitioning the images according to the NIST Fingerprint Quality Index

| 0 | 1 | 2 | 3 | 4 | 5 |
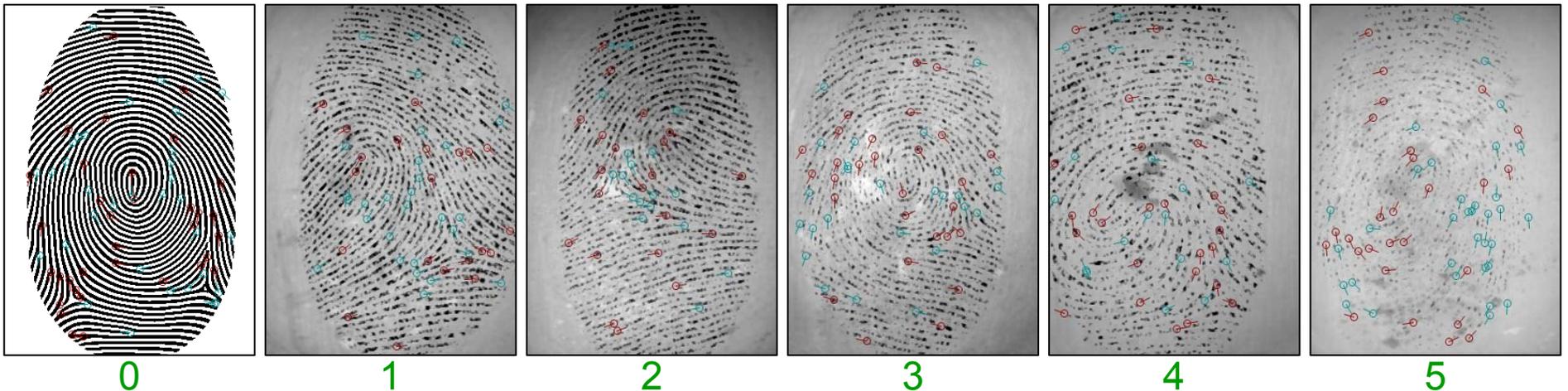
  - Minutiae ground-truth generated according to the ISO/IEC 19794-2 standard

- Algorithms
  - 5 feature extractors able to save ISO/IEC 19794-2 records have been tested

- The minutiae template created by each algorithm is compared to the ground truth

- The "corresponding" minutiae are determined using the Hungarian algorithm to maximize the number of matched minutiae (assignment problem)

- Interesting statistics can be gathered:
  - False and missed minutiae, Type errors, Position errors, Direction errors
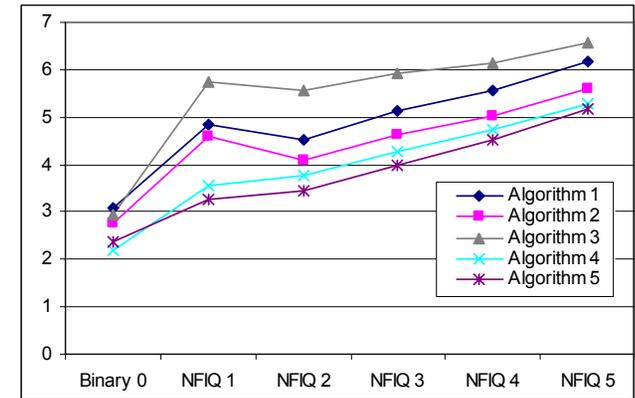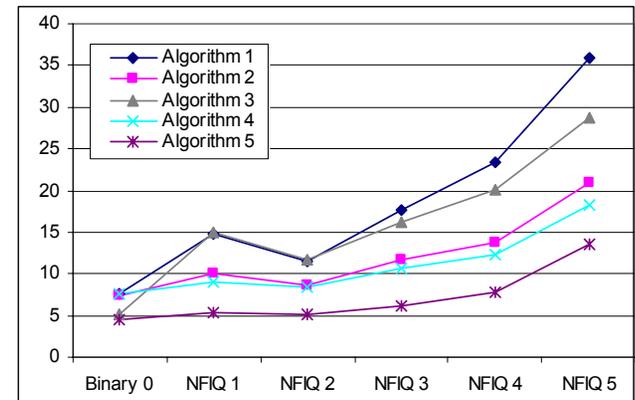

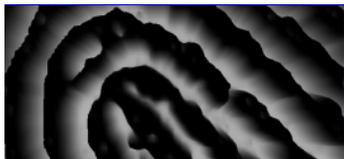Average position error
(in pixels)
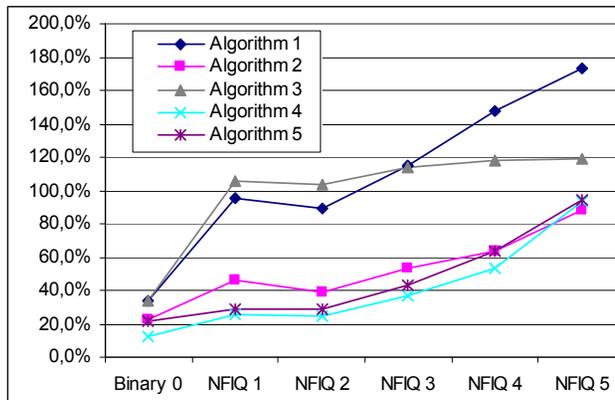
False+Missed minutiae
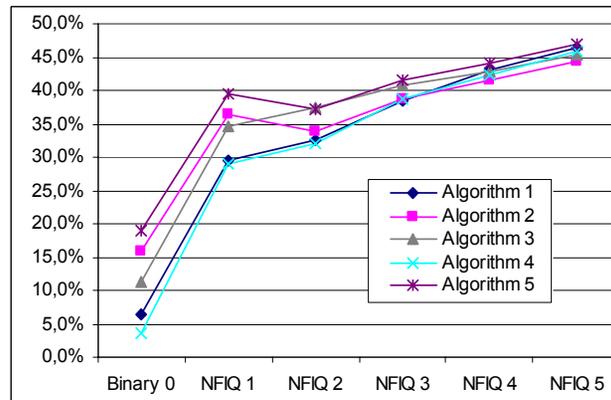
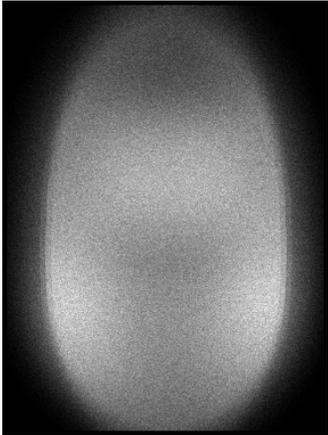Type errors
(ridge ending/bifurcation)

Average direction error
(in degrees)

Using synthetic data to evaluate the quality of minutiae extraction algorithms
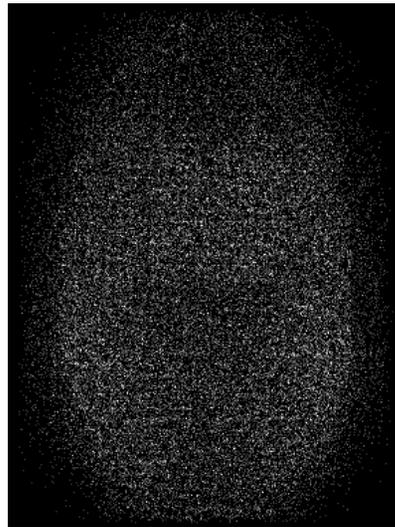
Experiment 2: Minutiae density

Intensity $I[x,y]$ is proportional to the estimated likelihood that a minutia will generated by SFinGe at (x,y).

The density function has been estimated from 200,000 minutiae templates generated by SFinGe.
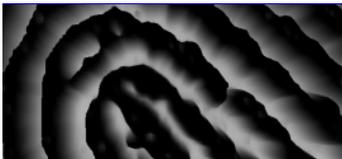

Ground truth

Algorithm 5

$I[x,y]$ is proportional to the estimated likelihood that a minutia will be found at (x,y) by the feature extraction algorithm.
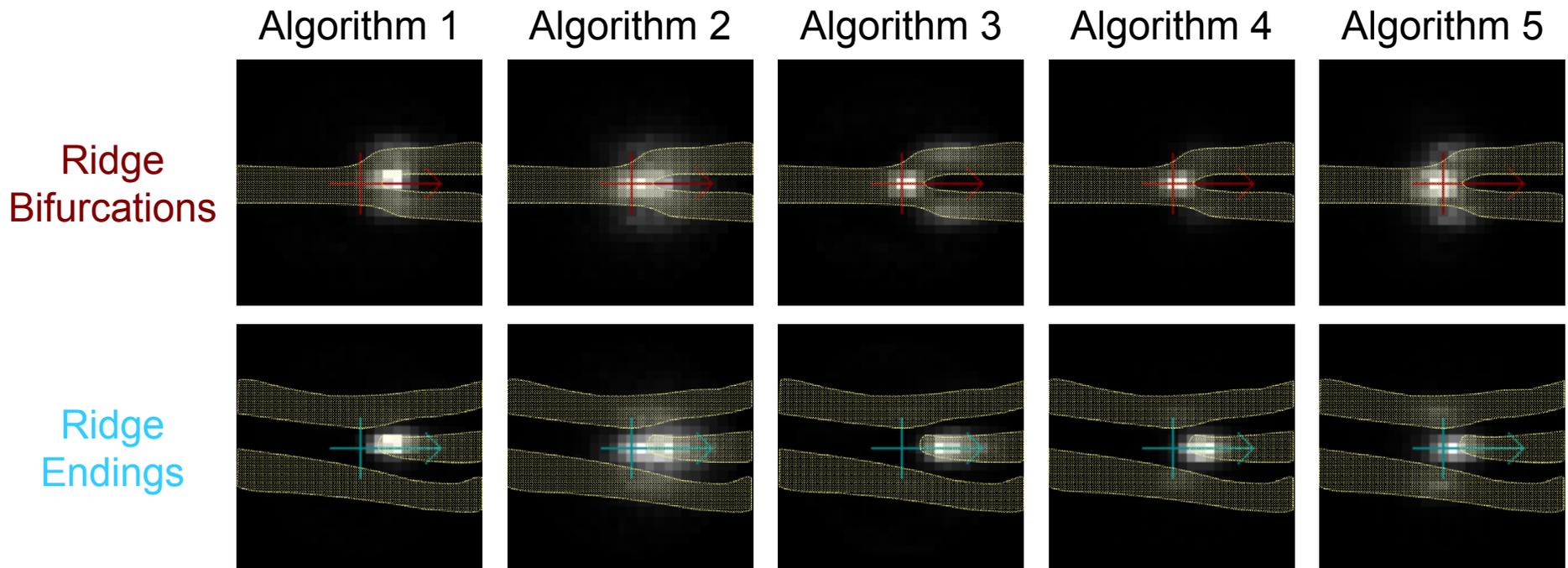
The density function has been estimated from 1,000 fingerprints [288x384 pixels] in dataset 0.
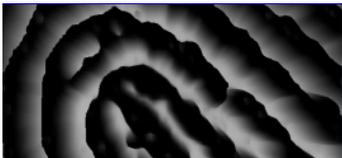
Experiment 3: Analysis of minutiae placement

Intensity $I[x,y]$ is proportional to the estimated likelihood that a minutia will be found by a feature extraction algorithm at position (x,y) *with respect to the ground truth minutia direction* (denoted by the arrow).


Dataset used: 1 (NFIQ=1)

Using synthetic data to evaluate the quality of minutiae extraction algorithms

Experiment 4: Minutiae placement correction

Offset added to each minutia along the minutiae direction reported by the algorithms		
Minutiae detected as	*Alg. 4*	*Alg. 5*
Bifurcations	-1 pixel	+1 pixel
Endings	-4 pixels	-1 pixel

Alg. 4   Alg. 5

Ridge Bifurcations

Ridge Endings
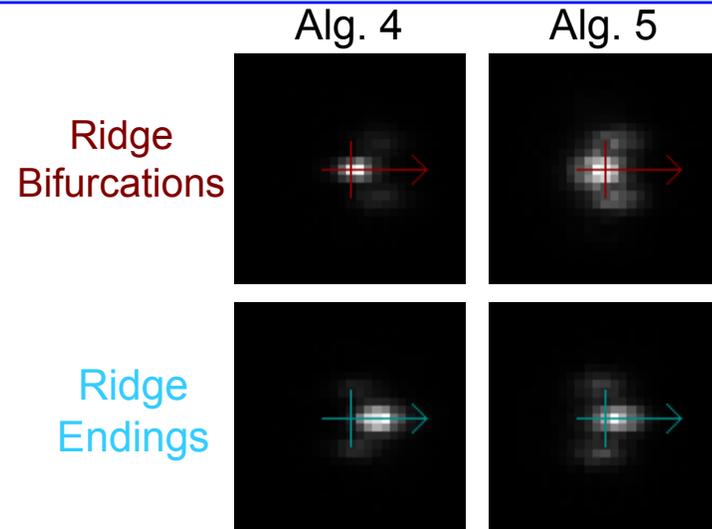
Experiment on FVC2006 DB2 (1680 fingerprints)
A real (non-synthetic) fingerprint database

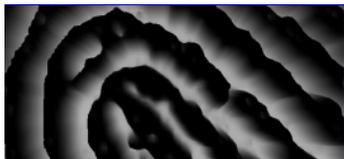| Without correction		
FNMR at FMR=0.1%	Alg. 4	Alg. 5
Alg. 4	0.108	1.212
Alg. 5	0.801	1.537

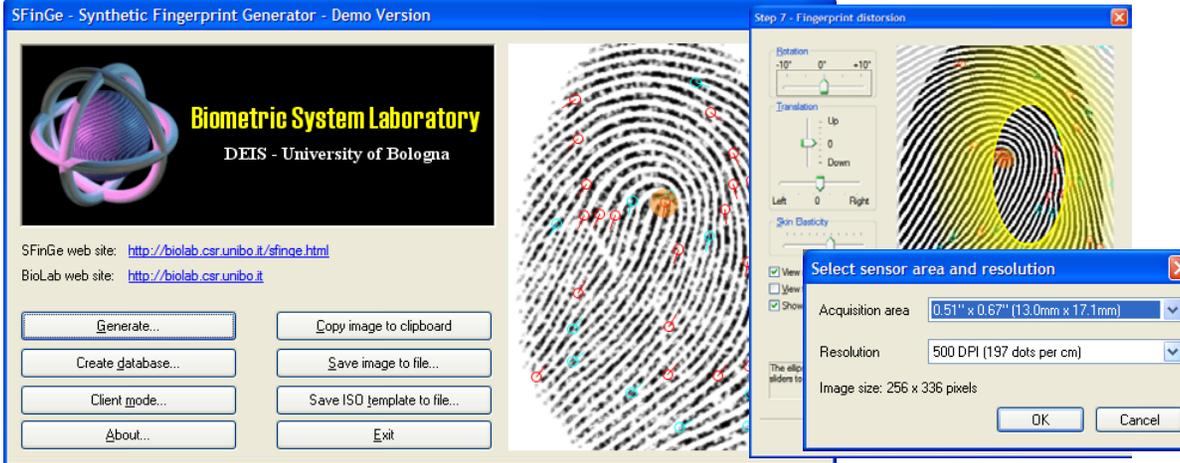| With correction		
FNMR at FMR=0.1%	Alg. 4	Alg. 5
Alg. 4	0.108	0.747
Alg. 5	0.703	1.558

-12%

-38%

+1%

Using synthetic data to evaluate the quality of minutiae extraction algorithms

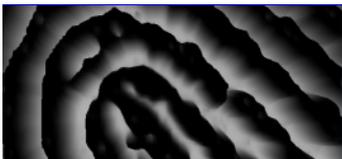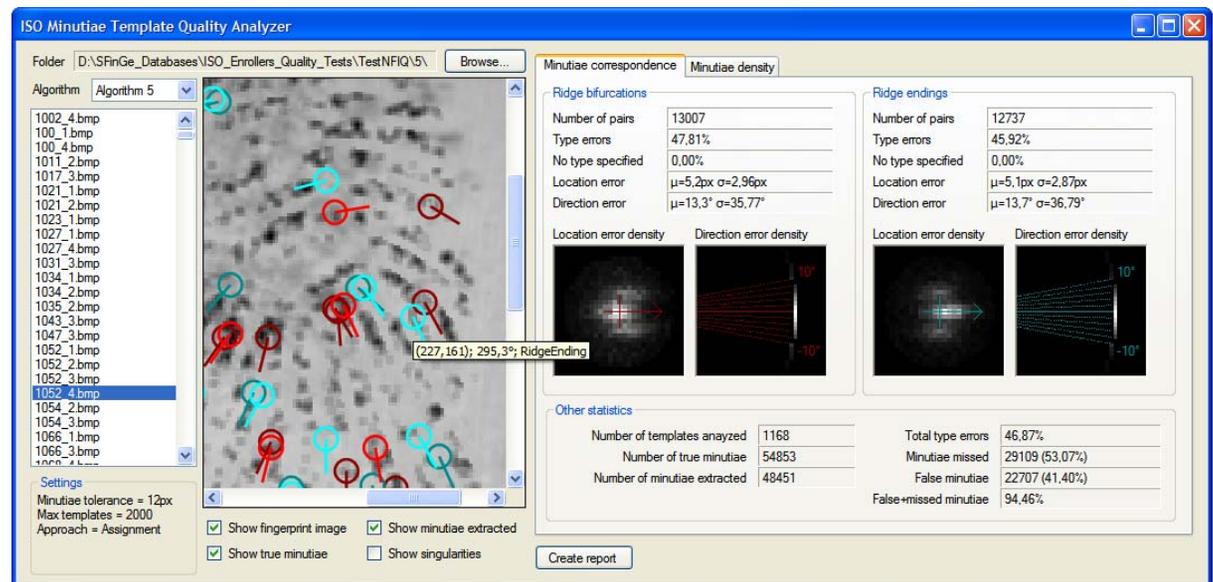# Software tools that are being developed at BioLab

New version of SFinGe
- Generates fingerprints (500/1000ppi) with *ground truth minutiae*
- Distributes the generation among multiple computers

Minutiae Template Analyzer
- Compares ground truth with extracted minutiae
- Generates statistics

Using synthetic data to evaluate the quality of minutiae extraction algorithms

16

Conclusions

- Syntactic conformance to a standard minutiae format is not enough for interoperability: *semantic conformance* is needed

- SFinGe can be used to generate large fingerprint databases with *ground-truth minutiae*
  - The minutiae can be extracted by applying the standard procedures easily and without ambiguities (extraction occurs on a binary image without noise)
  - Datasets to test the impact of a given parameter (e.g. fingerprint quality) can be easily generated

- From the preliminary experimental results:
  - Fingerprint quality heavily affects the quality of the extracted minutiae (false/missed minutiae, placement errors, direction errors, type errors)
    - Some algorithms are more affected by low quality than others
  - The distribution of minutiae position errors with respect to the ground truth:
    - Clearly reveals cases of semantic non-conformance
    - Depends on the minutiae type and is characteristic of each extraction algorithm
    - Its analysis may be useful to improve interoperability
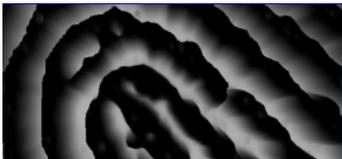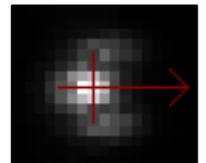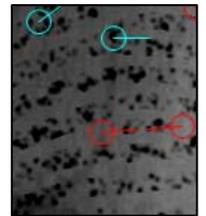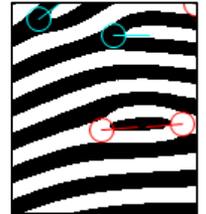
Biometric System Laboratory
University of Bologna - ITALY

http://biolab.csr.unibo.it
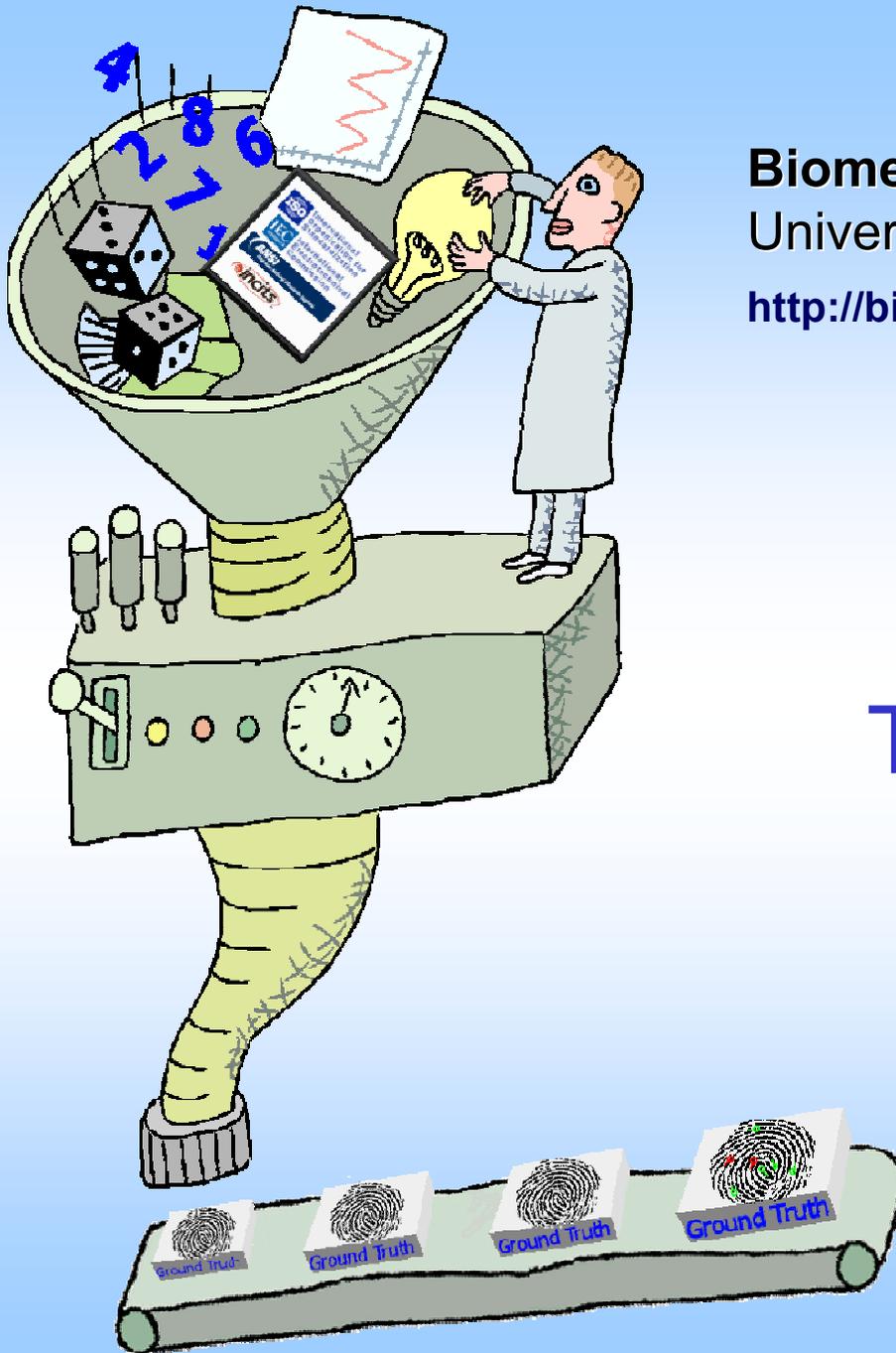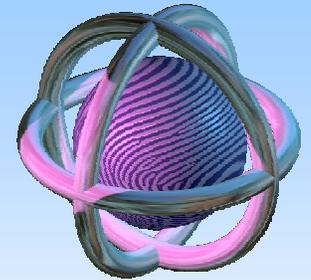
Thank you

Raffaele Cappelli

cappelli@csr.unibo.it