Towards Standardised Fingerprint Matching Robustness Assessment: The StirMark Toolkit – Cross-Feature Type Comparisons

J. Hämmerle-Uhl, M. Pober, A. Uhl

Department of Computer Sciences
University of Salzburg

September 25th, 2013

Overview

Outline

Fingerprint recognition robustness

- Sample image quality impacts on recognition accuracy
- Skin conditions (e.g., dryness, moisture, dirt, cuts and bruises, ageing)
- Sensor conditions (e.g., dirt, noise, size)
- User cooperation, crime scene preservation
- Benchmarking frameworks: FVC, BioSecure, SFinGe
- **Problem:** Additional acquisition conditions require re-enrolment; results from different datasets hard to compare.
- **Here:** Propose standardised tool to simulate a wide class of acquisition conditions, applicable to any given dataset. StirMark used as a first example. **Difference to IH&MMSec'13 paper**: Different StirMark Tools, non-minutiae matchers.

Outline

The Stirmark Toolkit

Basic idea: The StirMark Benchmark is a generic benchmark test for evaluating the robustness of digital image watermarking methods, developed by Fabien A. P. Petitcolas et al.

- Additive noise: Actual dust on the fingerprint contact area, sensor noise, grainy surface a latent fingerprint has been taken off
- Median Cut filtering: Simulates blur in fingerprint images, e.g. smudgy fingerprints (too much moist) etc.
- Remove Lines and Columns: Sensor errors, esp. sweep sensors can be affected by line removal (examples are shown)
- Rotation: Omnipresent challenge in fingerprint recognition
- Stretching: Simulates a higher force being applied when pressing the finger onto the contact area, in forensics a soft or flexible surface can be the reason
- Shearing: Simulates a setting where the applied pressing force is not perpendicular to the contact area
- Random distortions: Combination of several distortion types, modelling e.g. unevenly distributed pressure exercised on the

StirMark Examples

(a) Noise (level 15)

(b) Median Cut Filter (size 9)

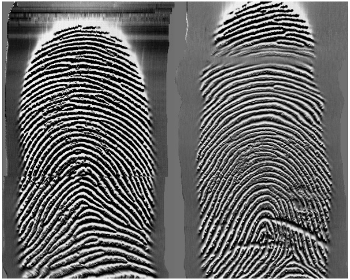
(c) Rotation of -15°

(d) Stretching ($d = 1.350$)

(e) Shearing ($b = c = 0.20$)

(f) Random Dist. (lrnddist 4.2)

Real Examples



(a) Missing lines

(b) Warping effects

Figure: Examples for distortions from actual acquisition problems.

Types of Fingerprint Matchers

Correlation-Based Matcher These approaches use the fingerprint images in their entirety, the global ridge and furrow structure of a fingerprint is decisive. Images are correlated at different rotational and translational alignments.

Ridge Feature-Based Matcher Matching algorithms in this category likewise deal with the overall ridge and furrow structure in the fingerprint, yet in a localised manner. Characteristics like local ridge orientation or local ridge frequency are used.

Minutiae-Based Matcher The set of minutiae within each fingerprint is determined and stored as list, each minutia being represented (at least) by its location and direction. The matching process then basically tries to establish an optimal alignment between the minutiae sets.

Outline

Experimental Settings

- DB1 - DB3 from the FVC2004 data are used in a verification setting employing the evaluation protocol as specified by FVC
- Fingerprint matching software
  - "NIST Biometric Image Software" (**NBIS**) package (*mindtct* and *bozorth3*)
  - *Phase only correlation (POC)*, custom implementation. First, the normalised cross spectrum (or cross-phase spectrum) of the DFT of the two images is computed. The POC is then obtained by taking the inverse DFT of the normalised cross spectrum.
  - *Fingercode (FC)*, custom implementation. A Gabor filter bank is applied to the orientation image resulting in a "Ridge Feature Map" which is translationally and rotationally aligned for matching.
- EER is used to assess and compare the impact of StirMark distortions

Results on plain data (without Stirmark distortions)

Table: EERs for the considered fingerprint recognition schemes when applied to the original, “undistorted” sample image databases within the StirMark framework.

	NBIS (%)	FC (%)	POC (%)
DB1	14.81	12.54	22.60
DB2	11.12	9.60	9.69
DB3	6.68	8.98	15.07

$\longrightarrow$ ranking of the algorithms is heavily dependent on the used dataset !

Results on Additive Noise: DB2

Table: EERs: All matchers affected, FC is best, POC second, NBIS third.

Noise Level	NBIS (%)	FC (%)	POC (%)
unperturbed	11.12	9.60	9.69
03	10.86	11.85	10.65
07	15.03	14.22	14.36
11	20.54	17.74	20.22
15	30.78	21.80	26.94

Results on Additive Noise: DB3

Table: EERs: NBIS hardly affected, POC is second and well performing, FC worst and severly affected worst and severly affected.

Noise Level	NBIS (%)	FC (%)	POC (%)
unperturbed	6.68	8.98	15.07
03	7.05	9.25	15.28
07	7.19	10.50	15.16
11	7.08	14.79	15.71
15	7.91	24.99	17.46

Results on Median Cut Filtering: DB1

Table: EERs: Performance on unperturbed data cannot predict EER under filtering !

Filter Size	NBIS (%)	FC (%)	POC (%)
unperturbed	14.81	12.54	22.60
03	15.50	12.90	23.63
05	17.69	13.52	24.92
07	32.17	16.55	30.71
09	46.88	28.26	38.11

Results on Line Removal: DB2

Table: EERs: POC cannot cope with increased amount of missing lines !

k	NBIS (%)	FC (%)	POC (%)
unperturbed	11.12	9.60	9.69
90	11.04	9.71	10.00
70	11.60	9.73	10.24
40	11.99	9.47	11.18
20	12.92	9.97	14.75

Results on Rotation: DB3

Table: EERs: NBIS very high robustness against rotations, while FC and POC are severely affected.

Rotation	NBIS (%)	FC (%)	POC (%)
unperturbed	11.12	9.60	9.69
-15	11.00	14.13	14.22
-5.5	11.28	12.04	10.89
13	10.59	13.57	13.01
20	10.94	16.27	18.08

Results on Shearing (Y-direction): DB1

Table: EERs: NBIS and FC hardly affected up to medium strength, POC severly affected.

Configuration	NBIS (%)	FC (%)	POC (%)
unperturbed	14.81	12.54	22.60
1	13.85	12.57	22.64
2	13.88	12.76	24.79
3	14.88	13.30	27.43
4	16.26	14.15	29.90
5	17.96	14.71	37.73
6	21.46	15.82	40.22

Results on Small Random Distortions: DB2

Table: EERs: NBIS with good robustness up to medium strenth, FC better as POC.

Factor	NBIS (%)	FC (%)	POC (%)
unperturbed	11.12	9.60	9.69
0.6	10.78	12.42	10.89
1.0	11.35	12.34	11.49
1.8	11.75	12.40	13.23
2.6	12.57	12.61	16.34
3.4	13.23	13.57	19.00
4.2	14.82	14.05	21.96

Outline

Conclusion

Lessons learnt:

- Significant variability of robustness properties across different types of matching schemes **and** across different datasets
- Results underline the need for a standardised tool in fingerprint recognition robustness assessment
- StirMark manipulations are a **first model** for a wide class of fingerprint acquisition conditions (including some forensic settings)
- More accurate modelling of actual fingerprint acquisition condition required
- **Example**: Accurate modelling of forensic conditions to resolve the urgent demand for realistic forensic testdata

Thank you for your attention!

Questions?