# Thoughts on Fingerprint Image Quality and Its Evaluation

NIST Biometric Quality Workshop II
November 7- 8, 2007

Masanori Hara

# Recap from NEC's Presentation at Previous Workshop (2006)

- **Positioning quality:** **a key factor to guarantee common area and matching accuracy**
- **Pattern area:** **good positioning criteria**
- **Quality to predict accuracy:** **matcher dependent (algorithm dependent)**
- **NEC quality metrics:** **better accordance with NEC matchers than NFIQ**

# Contents

# 1. Quality Concepts – Ideal Quality

## What is ideal quality?



**1) good ridge quality**

- dynamic range sufficiently wide
- uniformity evenly distributed density
- linearity gray mid scale reserved
- no saturation (white & black)
- no significant smudge or blur
- sufficient ridge/valley separation

**2) no problematic background noise**

- no Leftover fingerprint or stripe pattern
- background lighter than ridge (foreground)

**3) sufficient size**

- excellent for slap/flat matching
- good for latent cognizant

**4) good positioning and orientation**

- pattern area included & fingertip up

**5) no significant distortion**

***Ideal quality Strong image enhancement NOT required***

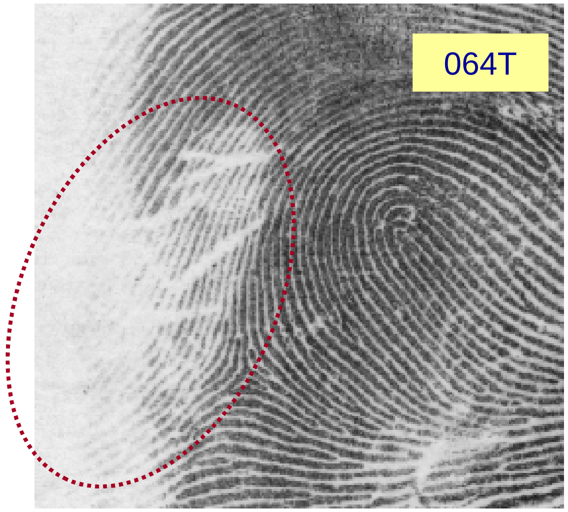
# 1. Quality Concepts – Poor Quality Samples

**Factors related to ridge quality and background noise**

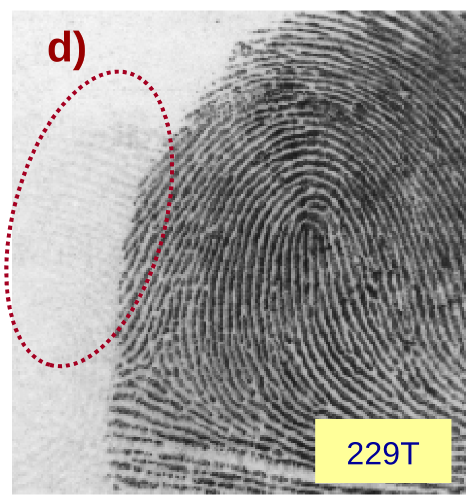
- **a) narrow dynamic range**
- **b) uneven density**
- **c) white saturation**
- **d) leftover fingerprint**
- **e) problematic stripe pattern**

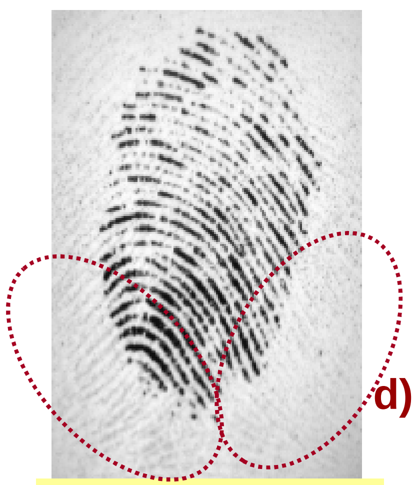
Note: Fingerprint samples are from NIST DB#27 and FVC Data Base.


**a), b), c)**




**a), b)**




FVC2002DB2 006_7

**e)**

FVC2002DB3 001_3

# 2. Requirements for Quality

"Requirements for quality" depend on:

- target filed – **law enforcement (LE)** or **non LE (NLE)**
- operational requirement - **automatic** or **manual intervening**
- image type (flat, slap, rolled, latent)

<u>two major operational categories:</u>

**a) fully automatic matching (NLE/LE)**

a1) Positive ID (cooperative & unsupervised - NLE)

**a2) Negative ID (uncooperative & supervised - NLE)**

**a3) Automatic latent (rolled/latent - LE)**

**b) manual intervening operation (LE)**

**b1) for latent cognizant rolled print**

**b2) for latent print**

*Note: a1) is not discussed here.*

# 2. Requirements for Quality

## Criteria for quality

a) criterion for acceptance/rejection at capture
b) criterion for enrollment or registration
c) criterion for special search (e.g. latent)
d) criterion to predict “matching accuracy”

**1) real-time processing required for criterion a)**

**2) quality metrics specific to capture device is effective for real-time processing**

**e.g. positioning & orientation do NOT have to be checked for identification slap (4-slap) capturing**

***Note: Only offline processing for static image is discussed here.***

# 3. Pure Quality and Predictive Quality

***pure quality*** **– intrinsic quality of image itself**

*matcher independent*

*also Independent of operational needs*

***predictive quality*** **- quality for predicting accuracy**

*matcher dependent*

**a) predictive quality for auto (PQ_A)**

*for automatic operation (NLE, automatic latent)*

***compatibility with human examiners NOT required***

***OK to use with non-minutia-based matching***

**b) predictive quality for manual (PQ_M)**

*for manual intervening operation (LE, manual latent)*

***need to consider compatibility with human examiner's minutia definition***

# 4. Image Enhancement - Quality Metrics

## *- Question -*

## *Image enhancement – OK for quality metrics?*

### 1) contrast enhancement

- contrast stretch (global & local)
- histogram equalization (global & local)
- sharpening, etc.

**strong tools to cope with “narrow dynamic range” and “uneven density”, etc.**

### 2) ridge enhancement

- filtering – contextual, Fourier, Gabor, Wavelet, etc.
  (*) based on ridge direction & pitch
- pore & incipient ridge removal, etc.

**useful to cope with insufficient ridge/valley separation**

# 4. Image Enhancement - Quality Metrics

***- Question -***

***Image enhancement - OK for quality metrics?***

***No, at least for pure quality!***

**image enhancement has side effects such as:**

**1) increasing background noise**

**2) removing gray intermediate (mid) scale**

**3) creating false (ghost) ridges**

**4) removing true ridges**

**5) creating false minutiae and missing true minutiae**

# 4. Image Enhancement – Contrast Enhance

**Which image is of better quality?**

265T

A

NFIQ=50; NEC=61

dynamic range

**narrow < wide**

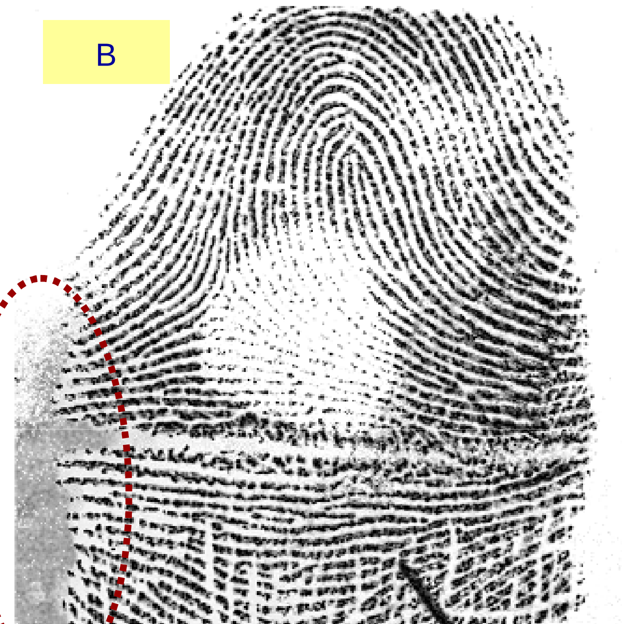
gray mid scale

**equivalent**

background noise

**less > more**

B



NFIQ=50; NEC=60

Image B: contrast-enhanced image of Image A using local contrast stretch

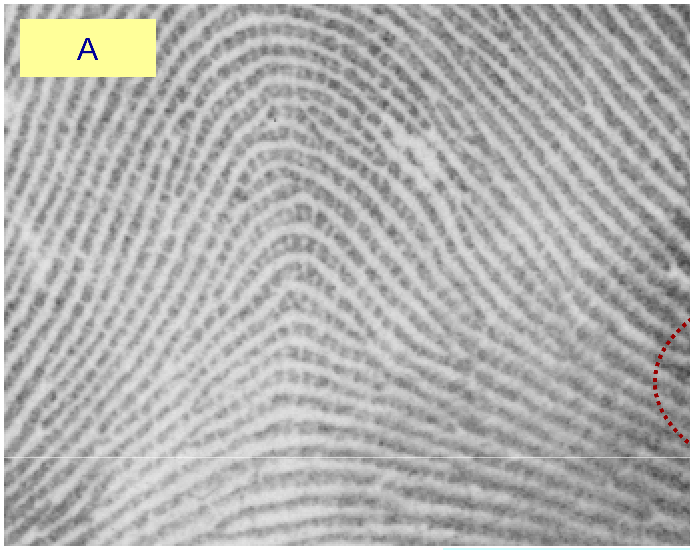
**Transformed NFIQ employed for easier comparison**

| | poorer | | | | better |
|---|---|---|---|---|---|
| NFIQ (original): | 5 | 4 | 3 | 2 | 1 |
| **NFIQ (transformed):** | **0** | **25** | **50** | **75** | **100** |
| NEC quality metrics: | 0 ------------------100 | | | | |

# 4. Image Enhancement – Contrast Enhance

**Which image is of better quality?**

092T

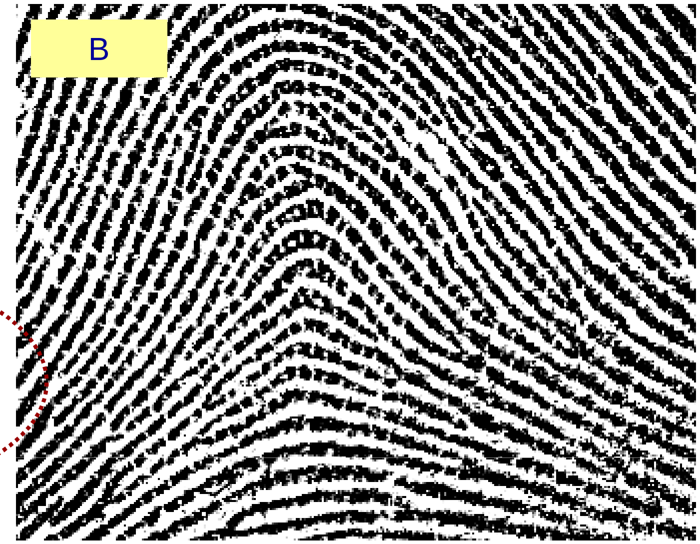



dynamic range

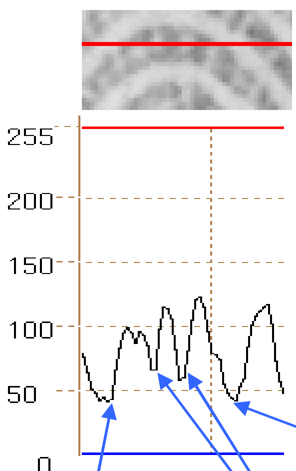
**narrow < wide**

gray mid scale

**some > no**




NFIQ=100; NEC=100




Valley

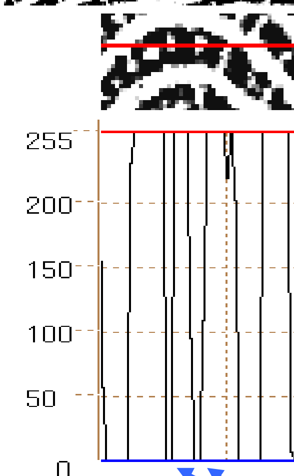
Valley

Pore

Image B: contrast-enhanced image of Image A using sharpening and contrast stretch

NFIQ=75

NEC=100



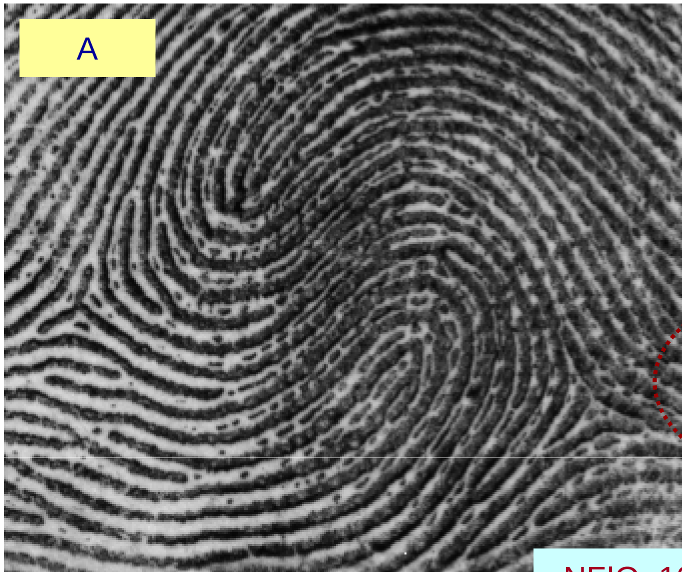

Pore

**no difference between pore and valley!**

# 4. Image Enhancement – Contrast Enhance

**Which image is of better quality?**

201T

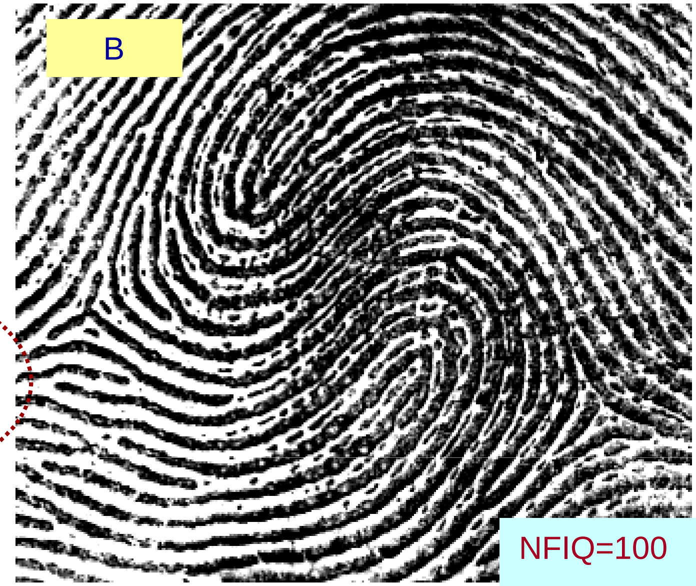



dynamic range

**narrow < wide**

gray mid scale

**some > no**




NFIQ=100; NEC=95

NFIQ=100

NEC=100

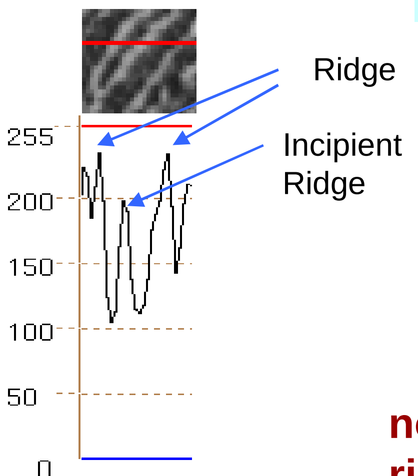
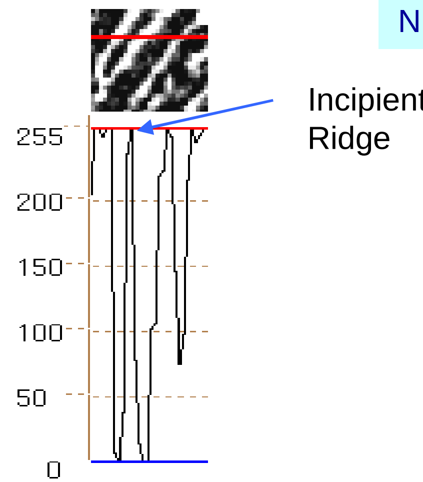



Image B: contrast-enhanced image of Image A using sharpening and contrast stretch




**no difference between incipient ridge and true ridge!**

# 4. Image Enhancement – Contrast Enhance

**Which image is of better quality?**

FVC02DB3 001_3

A

NFIQ=75; NEC=39

dynamic range
**narrow < wide**

gray mid scale
**some > little**

background noise
**more < less**

white saturation
**no > yes**

B

NFIQ=75; NEC=30

Image B: contrast-enhanced image of Image A using *light density removing* and contrast stretch

# 4. Image Enhancement – Ridge Enhance

**Which image is of better quality?**

265T

A

NFIQ=50; NEC=61

dynamic range

**narrow < wide**

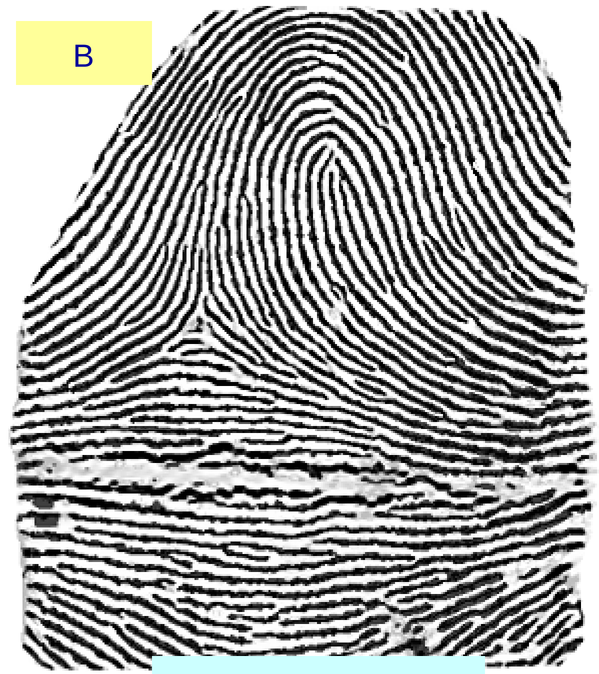
gray mid scale

**yes > not reliable**

fidelity

**Yes > No**

B



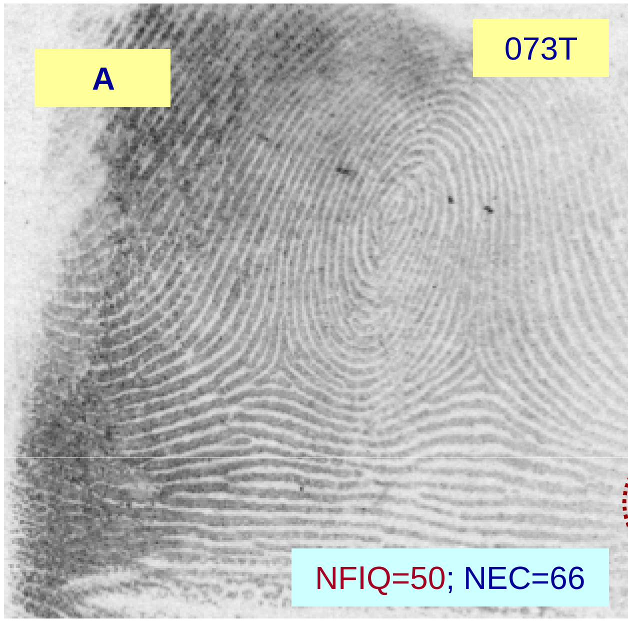
NFIQ=75; NEC=80

**Image B: ridge-enhanced image of Image A using contextual filtering**

**Ridge-enhanced image does NOT ALWATYS represent the original image.**

# 4. Image Enhancement – Ridge Enhance

**Which image is of better quality?**




dynamic range

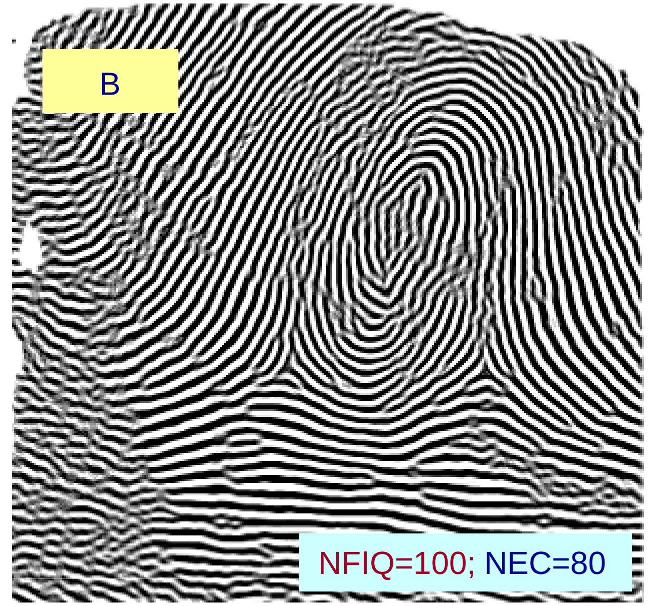
**narrow < wide**

gray mid scale

**yes > not reliable**

fidelity

**Yes > No**



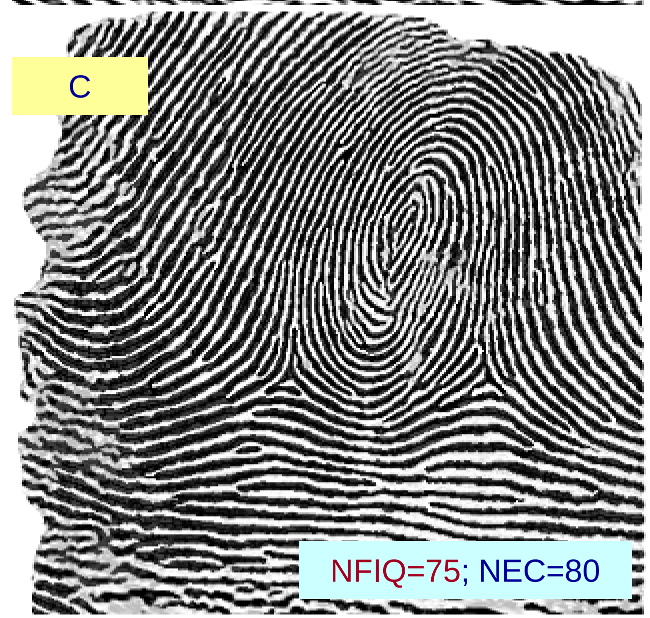




**Image B: ridge-enhanced image using fixed-width pitch data ( false ridges)**
**Image C: ridge enhanced image using variable-width pitch data (locally estimated)**

**Ridge-enhanced image does NOT ALWATYS represent the original image.**

# 4. Image Enhancement – Predictive Quality

***Image enhancement - OK for predictive quality?***

## a) predictive quality for auto (PQ_A)

***no problem to apply any image enhancements***

*reasonable to use equivalent method to matcher*

*(1) false minutia per examiners' definition be useful as long as such "feature" is consistent*

*(2) consistently miss-extracting true (but unstable) minutiae be more favorite than inconsistent extraction*

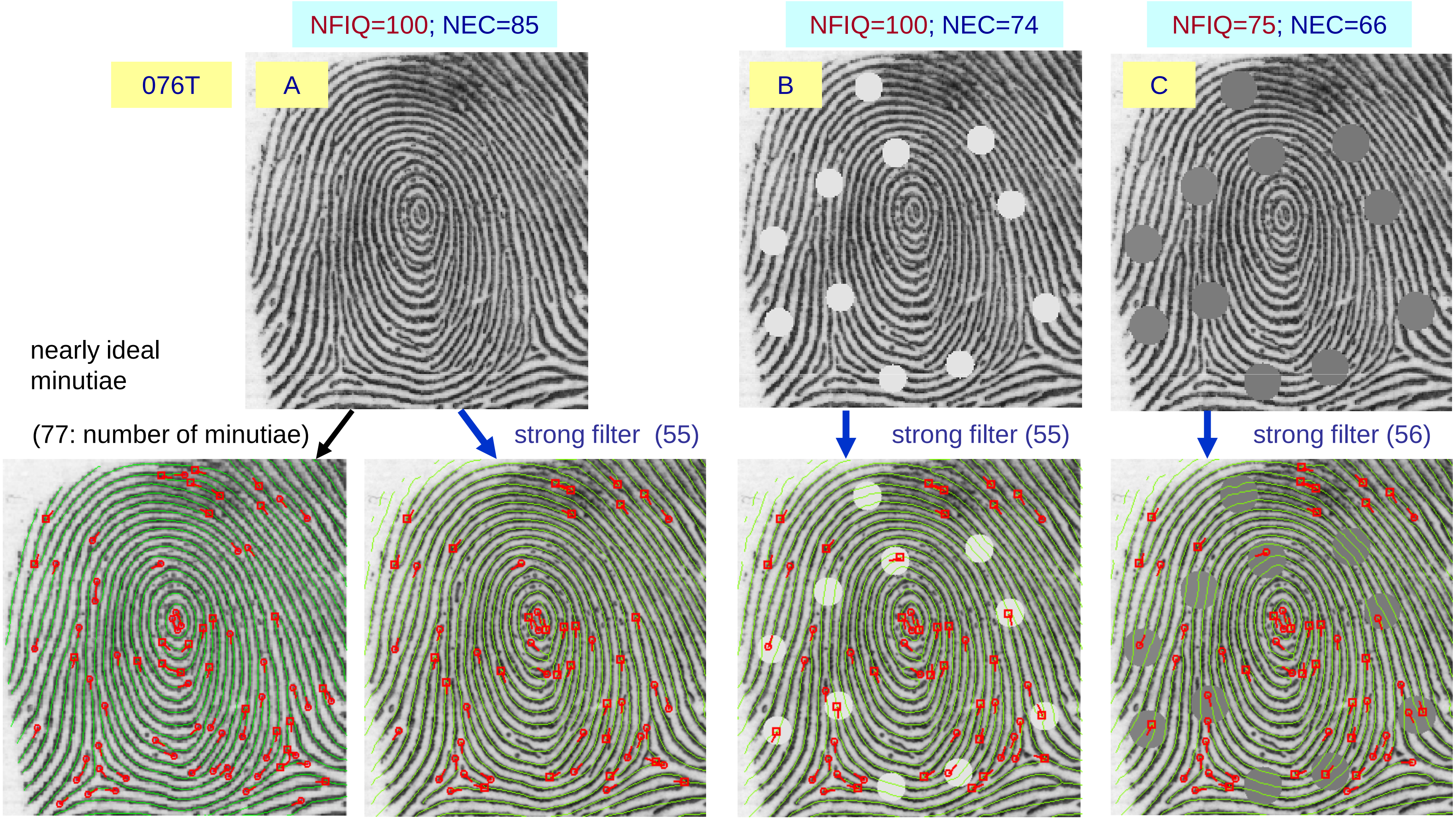
## b) predictive quality for manual (PQ_M)

***limited use suggested***

*strong filtering in ridge enhancement tends to create false minutiae or to remove true minutiae*

# 4. Image Enhancement – Predictive Quality

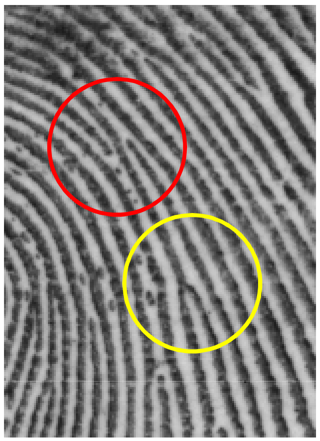
## *PQ_A: strong filtering - robust to low quality image*

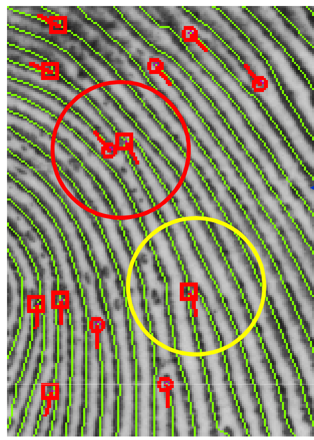


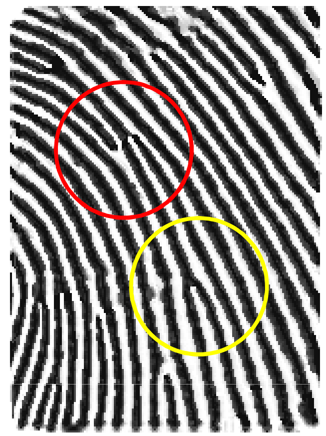

**Ignoring *unstable minutiae* increases *consistency of minutia* on auto process**

# 4. Image Enhancement – Predictive Quality

## *PQ_A: strong filtering - not desired for manual latent*





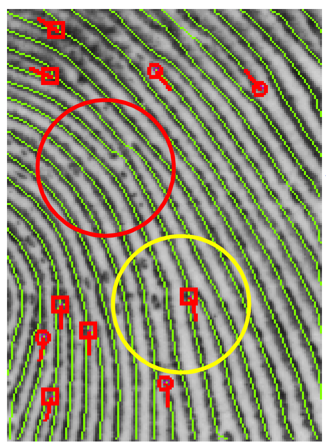


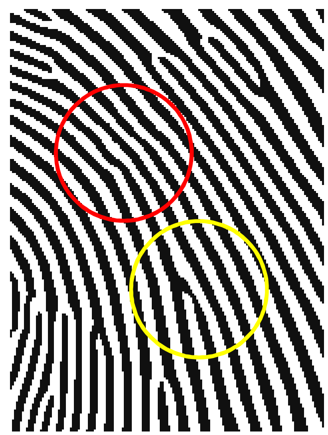
**weak filtering in ridge enhancement**

**good for manual latent**

**both unstable and stable minutiae extracted**

**both unstable minutia (e.g. crossover in red circle) and stable minutia (in yellow circle) are important as latent cognizant features for manual latent**





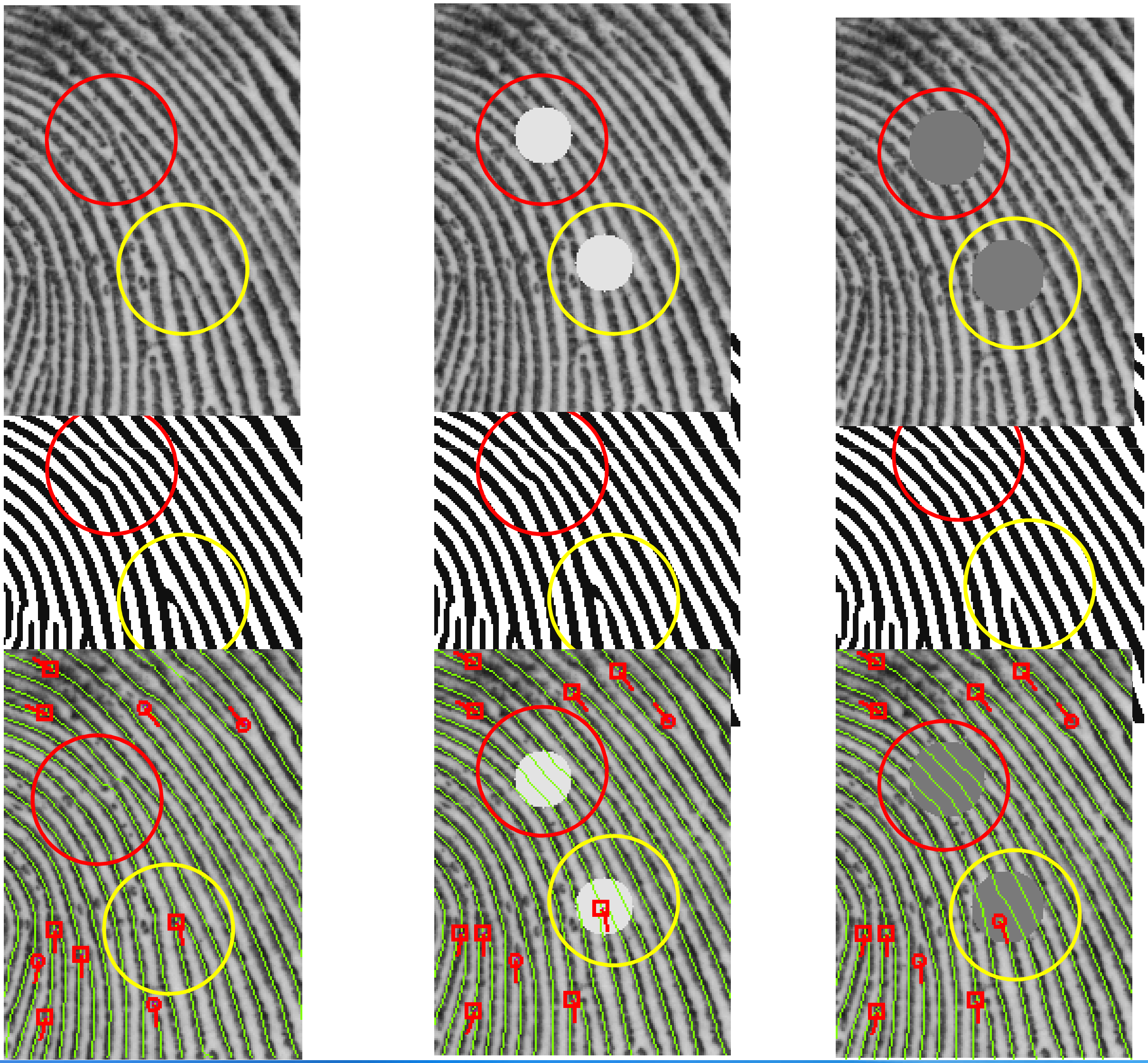
**strong filtering in ridge enhancement**

**not desired for manual latent**

**unstable minutiae missing**

# 4. Image Enhancement – Predictive Quality

## *PQ_A: strong filtering – why robust?*



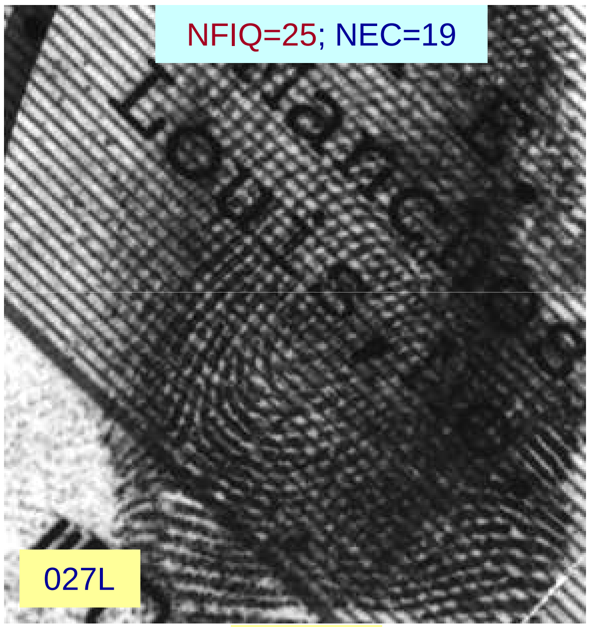
**Strong filtering in ridge enhancement creates stripe patterns even though there is no real ridge information on the input image.**

**This process also tends to create pseudo ridges (maybe false ridges) or to remove true ridges.**

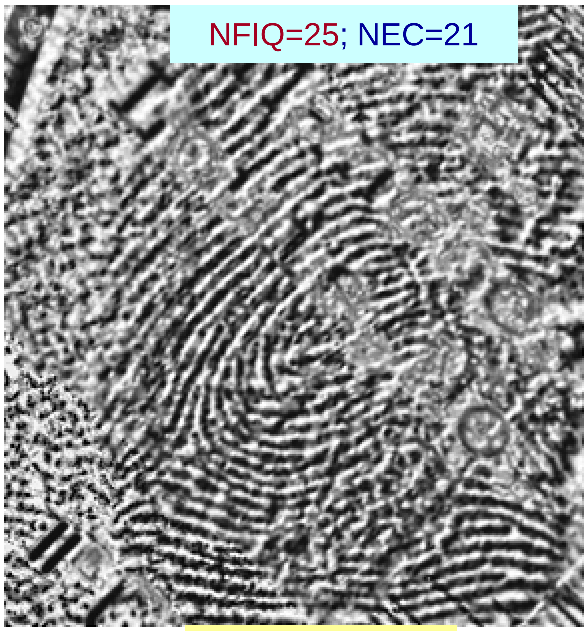
# 4. Image Enhancement – Latent

pure quality for latent print    not practical to be evaluated
predictive quality for manual latent    dependent on human examiner

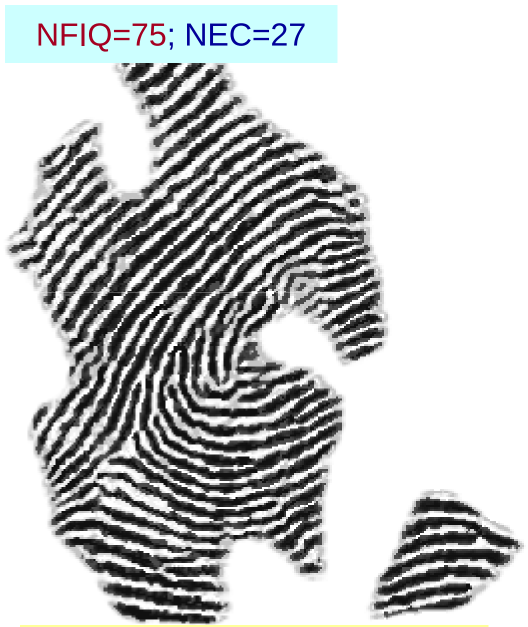
## *PQ_A: noise reduction - necessary for auto latent*



original



noise reduced



trimmed & ridge enhanced

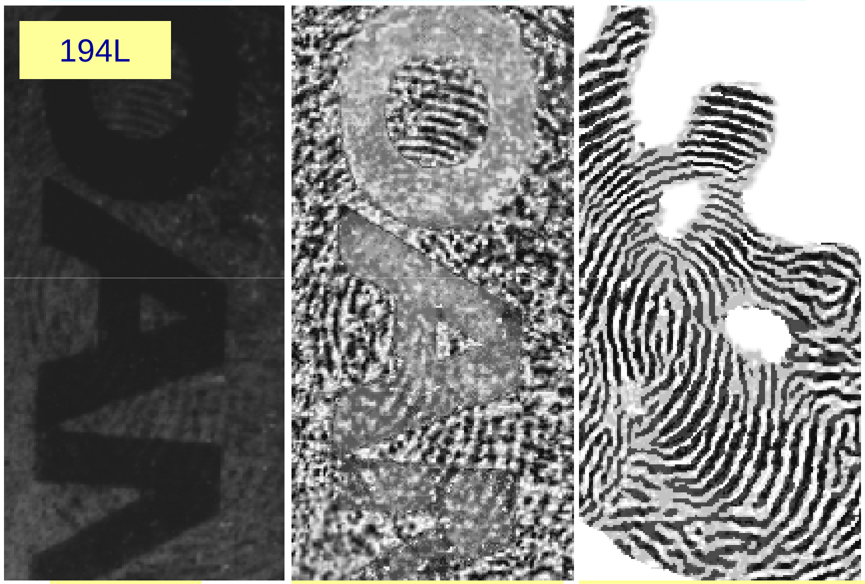
**Some latent prints have severe background noise. Quality metrics as well as matching expectation for those latent prints depend on *noise reduction performance***

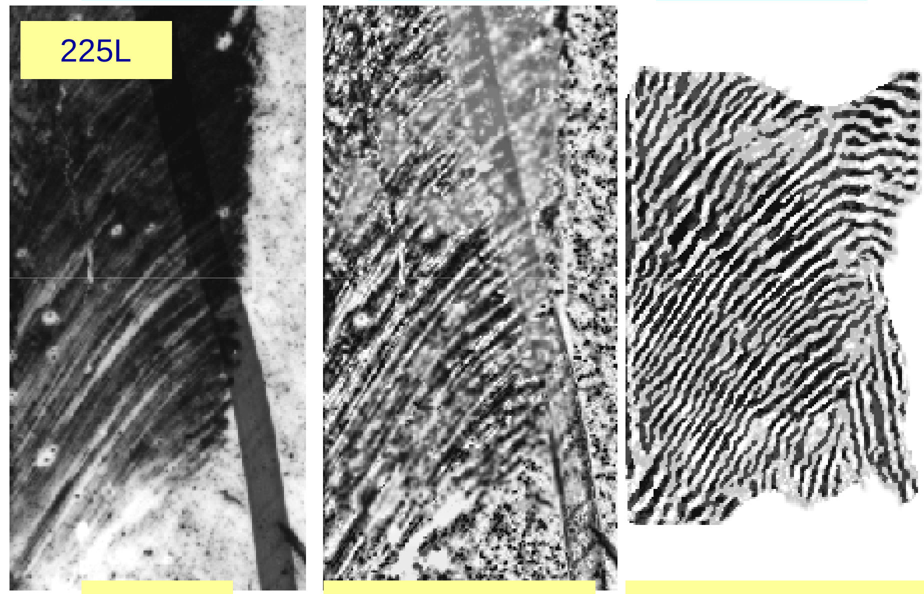
# 4. Image Enhancement – Latent

## *PQ_A: filtering - necessary for auto latent*

non fingerprint

| NFIQ=0 NEC= 0 | NFIQ=0 NEC= 0 | NFIQ=50 NEC=54 | NFIQ=0 NEC= 0 | NFIQ=0 NEC= 0 | NFIQ=100 NEC=30 |
|---|---|---|---|---|---|




| original | noise reduced | trimmed & ridge enhanced | original | noise reduced | trimmed & ridge enhanced |
|---|---|---|---|---|---|

**filtering robust to noise**

**filtering – tends to create false ridge even from non fingerprint pattern such as smudge**

# 5. Fingerprint Properness Analysis

*Fingerprint properness* and *of matching feature sufficiency*, etc. are important for predictive quality metrics.

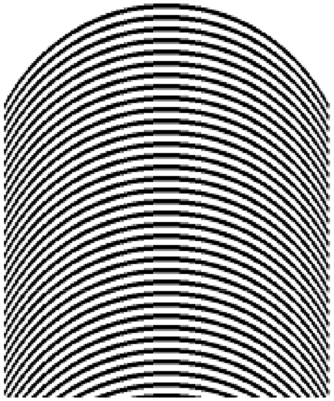
### *1) fingerprint pattern?*

**non fingerprint**

NFIQ=75

NEC=30

### *2) matching feature sufficient?*



**no minutia pattern**

NFIQ=75

NEC= 0

### *3) rotated?*

NFIQ=100

NEC=33

NFIQ=100

NEC=85

**Accepting rotated images will unnecessary increase matching cost and unnecessary decrease matching accuracy**

# 5. Fingerprint Properness Analysis

***4) distortion?***

FVC2004_DB1 028_3

NFIQ=100

NEC=80

FVC2004_DB1 028_4

NFIQ=100

NEC=69

***difficult to evaluate from static image!***

**<effective countermeasures>**

**1) identification slap (4-slap)  to prevent “intentional” distortion**

**2) matcher algorithm improvement  but, increase in cost involved**

Note: As for positioning and size issues, please see our presentation at 2006 Workshop

# 6. Factors to Degrade Quality

| | | Possible Causes | | Effective Countermeasures | Recapture Useful? |
|---|---|---|---|---|---|
| Factors to Degrade Quality | 1) poor ridge quality | device performance:<br>capture operation:<br>device problem:<br>nature of skin: | 40%<br>30%<br>20%<br>10% | better device<br>supervision<br>periodical replacement<br>nothing | to some extent |
| | 2) background noise | device performance:<br>capture operation: | 50%<br>50% | better device<br>maintenance (clean up) | to some extent |
| | 3) insufficient size | device performance:<br>capture operation: | 20%<br>80% | large area capture device<br>4-slap (3/2-slap) capture<br>supervision | yes |
| | 4) Improper positioning or orientation | device performance:<br>capture operation: | 20%<br>80% | 4-slap (3/2-slap) capture<br>large area capture device<br>supervision | yes |
| | 5) distortion | device performance:<br>capture operation: | 50%<br>50% | 4-slap (3/2-slap) capture<br>supervision | limited |

# 6. Factors to Degrade Quality

<u>effective countermeasures</u>

**1) better device most effective countermeasure**

**identification slap capture (4-slap)**

- **to consistently capture pattern area**
- **to solve rotation problem (fingertip up)**
- **to avoid distortion**
- **to avoid wrong finger capture**

Ref. T. Hopper; Identification Flats (NIST Fingerprint Standard Workshop)

***(*) 3 or 2-slap capturing is also effective***

**2) better supervision and capture operation**

- **to reduce background noise (by platen clean up, etc.)**
- **to capture sufficient size, etc.**

# 7. Quality Metrics Evaluation

evaluation method - different per type of quality metrics

- pure quality metrics evaluation

  *no straightforward method* ***not good for contest***

  ***r****ecommendation*

  ***specific criteria be evaluated by specific algorithms***

- predictive quality metrics evaluation

  *matcher dependent* ***tied up with matcher***

  *recommendation for contest*

  ***RRG(99.9): Rejection Rate to Guarantee 99.9%***

  ***quality and matcher integrated evaluation***

# 7. Quality Metrics Evaluation

## pure quality metrics evaluation

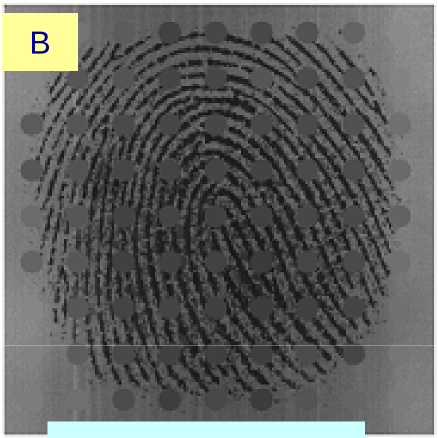
FVC2002DB3 22_6

A

NFIQ=75; NEC=28

FVC2002DB3 22_2

B



NFIQ=50; NEC=26

***Which image is of better quality?***

**A) good ridge quality but fingertip only**

**B) poor ridge quality but pattern area exits**

***difficult to define overall quality rating!***

***Can you rank these images?***

# 7. Quality Metrics Evaluation

## pure quality metrics evaluation

*specific criteria be evaluated by specific algorithms*

1) ridge quality    specific check tool
   *simple method suggested    NIST (or public domain) open source*
2) background noise    specific check tool
   *simple method suggested    NIST (or public domain) open source*
3) size    specific check tool
   *simple method suggested    NIST (or public domain) open source*
4) positioning and orientation    specific check tool
   *good candidate for contest*
   *reference (correct) data needed (manual coded) for contest*
5) distortion    difficult to check
   *difficult to evaluate from static image*

# 7. Quality Metrics Evaluation

**predictive quality metrics evaluation**

*1) PQ_A evaluation is relatively Simple.*
*This is discussed here.*

*2) However, PQ_M evaluation is difficult.*
*Method for pure quality evaluation is also practical for PQ_M evaluation.*

# 7. Quality Metrics Evaluation

recommended criteria for PQ_A metrics evaluation

## RRG(X): Rejection Rate to Guarantee X% Accuracy

**Given:**

***1) A set of fingerprint images***

***2) Its accuracy is less than X% (e.g. 99.9%)***

**Question:**

***How much proportion of the poorest quality images need to be rejected in order to guarantee X% (e.g. 99.9%) accuracy?***

RRG(X) – straightforward criteria to evaluate the predictive quality (PQ_A) metrics and matching performance at the same time

# 7. Quality Metrics Evaluation

## RRG(99.9) - recommended evaluation method

**1) contestants provide three programs**

a) a quality program to produce quality metrics (e.g. Q: 0-100)

b) a feature extraction program to produce templates

c) a matching program to produce score (e.g. 0-9999)

**2) NIST conducts test at the NIST facility**

a) produces quality metrics for search and file ($Q_{search}$ , $Q_{file}$)

b) TAR (or first rank hit) considered for simplicity

c) determine RRG(99.9) as follows

reject a mate if ($Q_{search} < Q_{th}$) or ($Q_{file} < Q_{th}$); $Q_{th}$ : $Q_{threshold}$

calculate $R(Q_{th})$ - percentage the rejection (a function of Q)

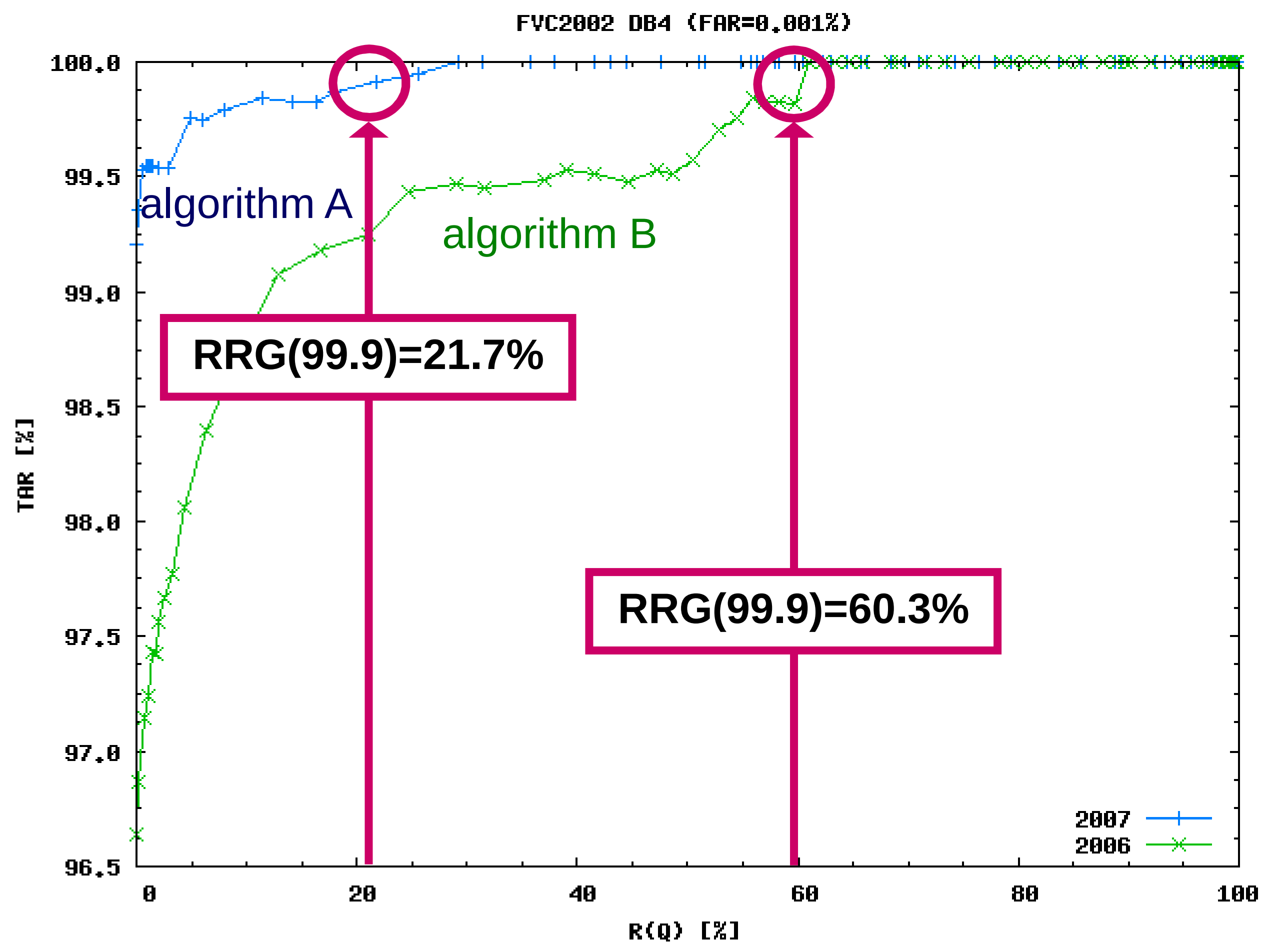
find RRG(99.9) – minimum R to achieve 99.9% accuracy

*Note: 99.9% (instead of 100%) is recommended as target accuracy in order to avoid undesired side effect from the exceptional data*

# 7. Quality Metrics Evaluation

## <u>Sample of RRG(X): R(Q) vs. TAR</u>




**straightforward method to compare different algorithms**

# 8. Conclusion and Suggestion

**1) image enhancement for pure quality**

- shall NOT be used or limited to moderate contrast enhancement

**2) image enhancement for predictive quality**

- for PQ_A (automatic) no restriction
- for PQ_M (manual) limited use suggested

**3) evaluation for pure quality**

- specific algorithm be developed by NIST (or public domain)
- not appropriate for contest

**4) evaluation for predictive quality (PQ_A)**

- RRG(99.9) criteria suggested
- appropriate for contest with proprietary matcher

**5) practical solution for negative identification system**

- identification slap capture (4-slap capture)

# Appendix: Improvement on NEC's Predictive Quality Metrics (1/13)

## NEC quality metrics　　PQ_A

- Rated on a 0-100 scale, where 0 is the lowest quality and 100 is the highest quality
- Nonlinear combination of four independent indices
  - ridge quality with its area size
  - high-confidence minutiae count
  - positioning quality for common area
  - distortion tolerance

Note: Appendix is prepared by Amane Yoshida from his research.

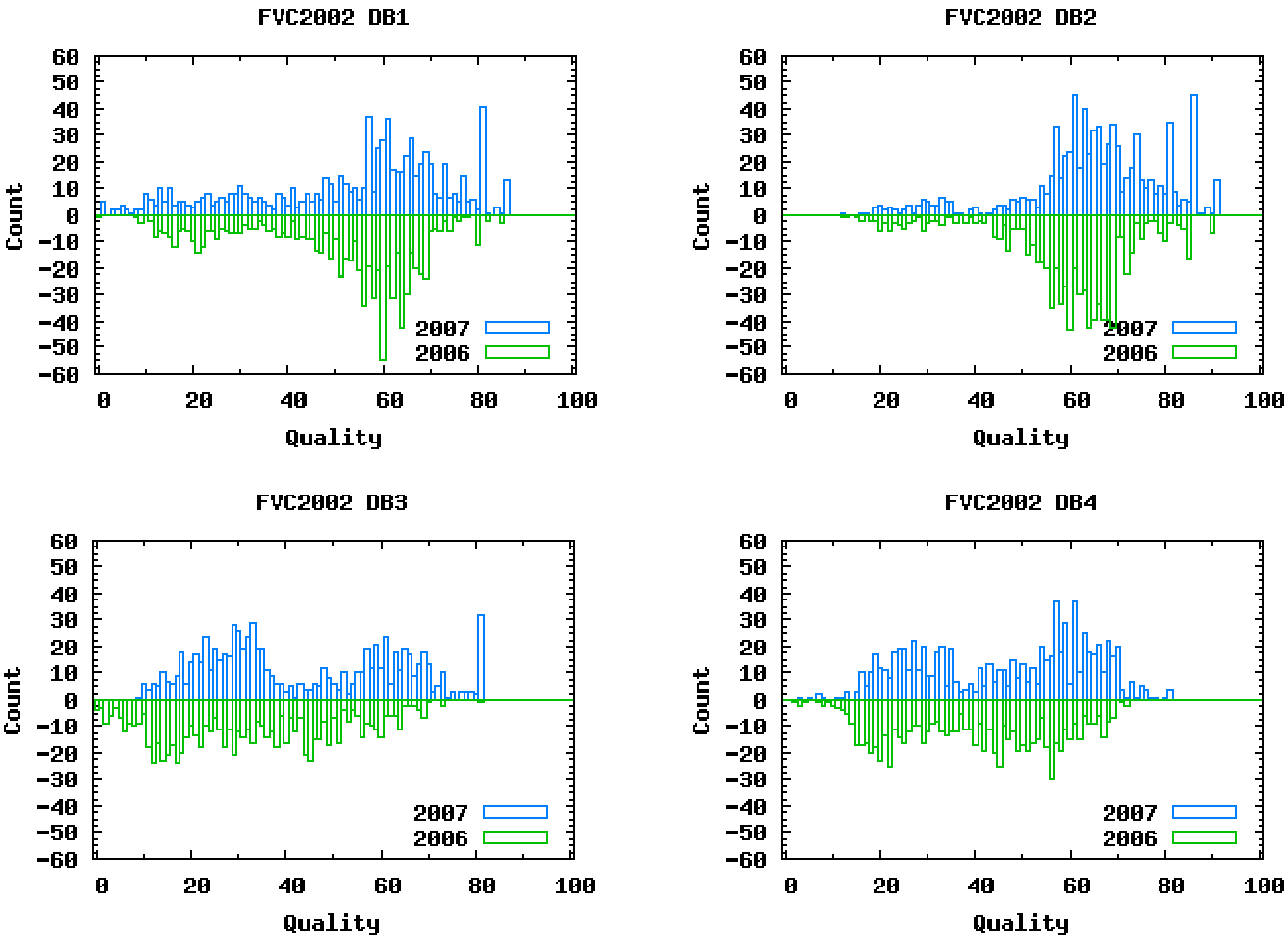
# Appendix: FVC2002 (2/13)

- **Accuracy Improvement**

| | Speed | | TAR at FAR=0.01% [%] | | | |
|---|---|---|---|---|---|---|
| | Match | FE | DB1 | DB2 | DB3 | DB4 |
| **SDK H4 (2007)** | **H-equiv.** | **Slow** | **99.73 (+0.09)** | **99.93 (+0.18)** | **99.07 (+0.69)** | **99.80 (+1.09)** |
| SDK H3 (2006) | H-equiv. | Slow | 99.64 | 99.75 | 98.38 | 98.71 |
| SDK H2 | H-equiv. | H-equiv. | 99.45 | 99.79 | 95.18 | 97.38 |
| SDK H | See NISTIR7151 | | 99.02 | 99.68 | 92.13 | 96.36 |

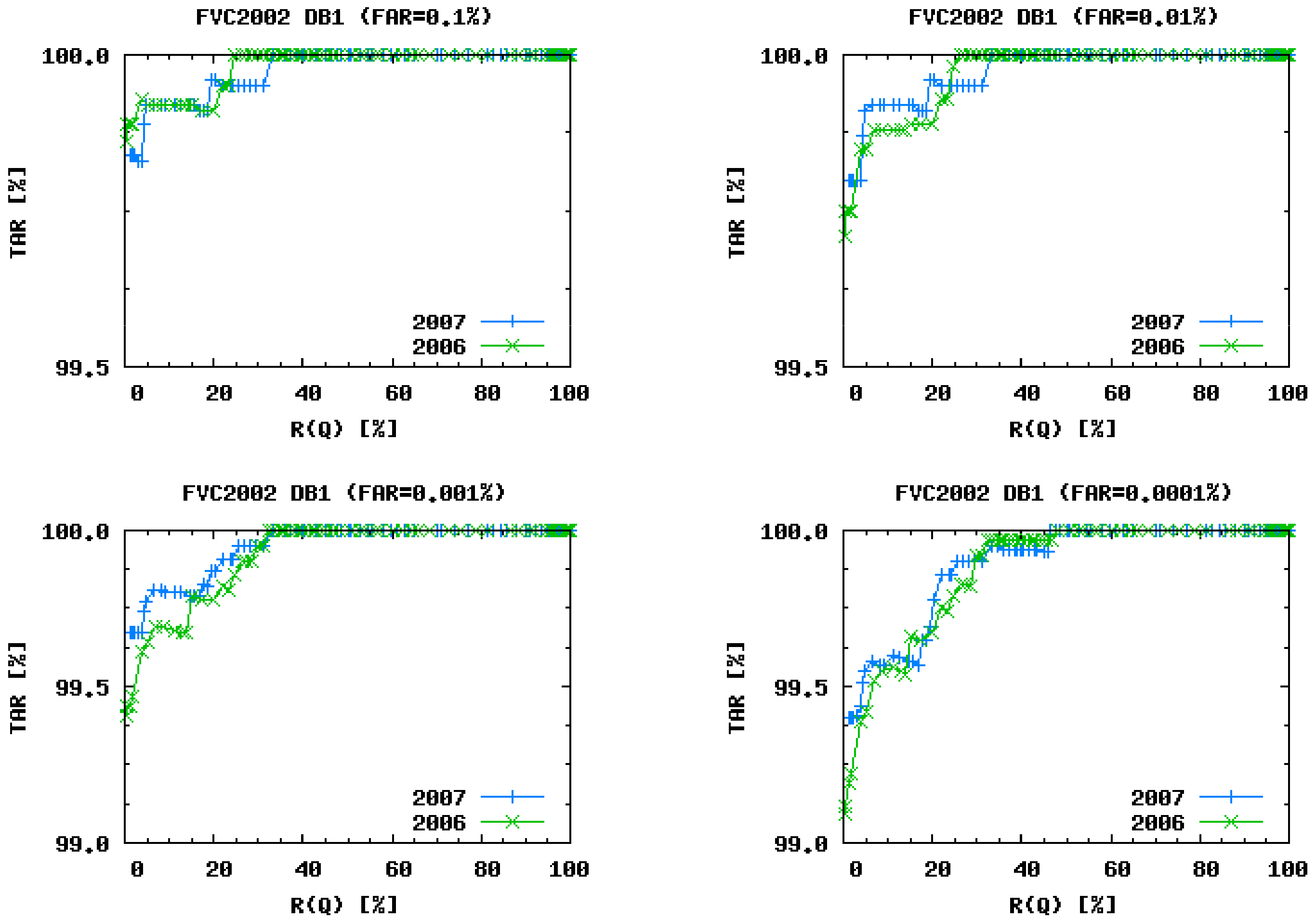
# Appendix: FVC2002 (3/13)

- **Quality Distributions**



# Appendix: FVC2002 (4/13)

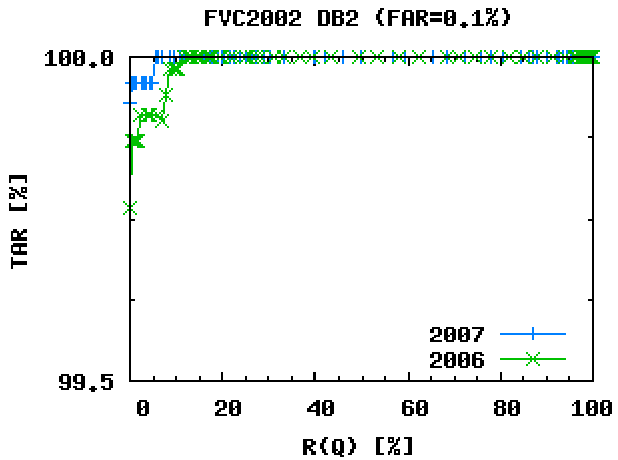
- **DB1: RRG comparison over varying FAR**

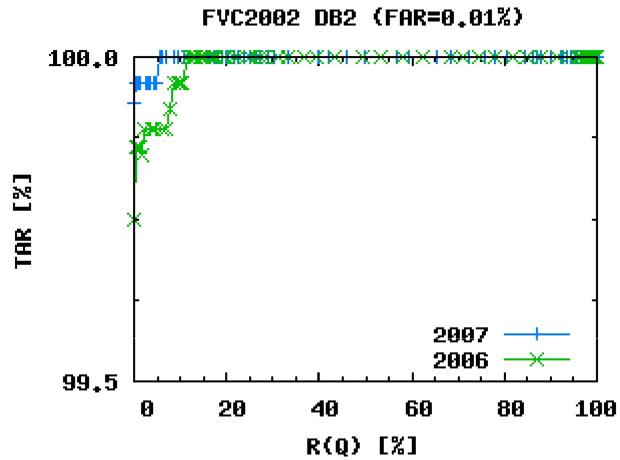


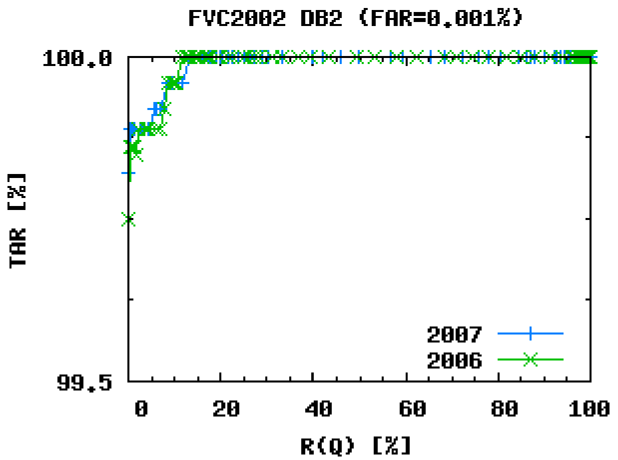

  

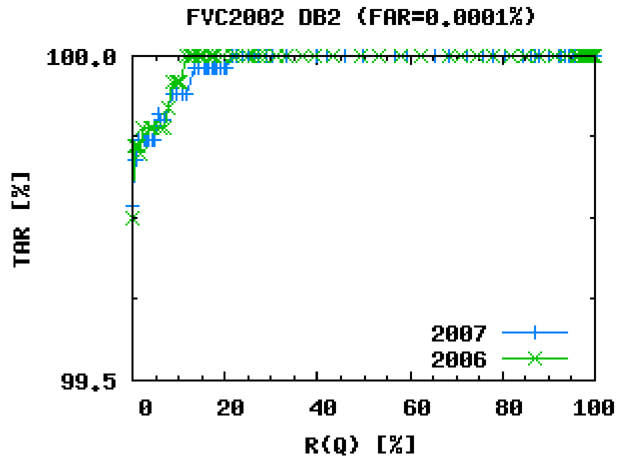
# Appendix: FVC2002 (5/13)

- **DB2: RRG comparison over varying FAR**



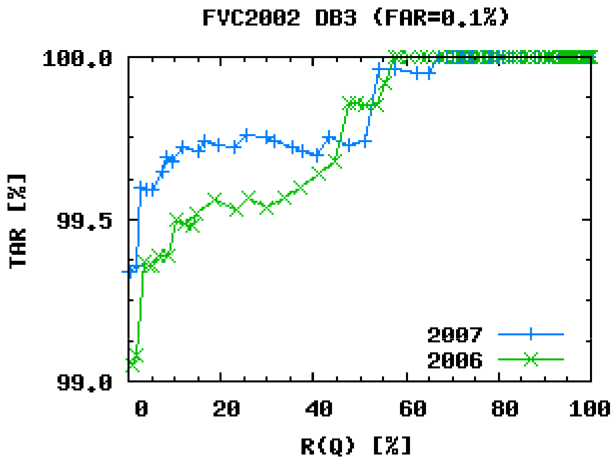



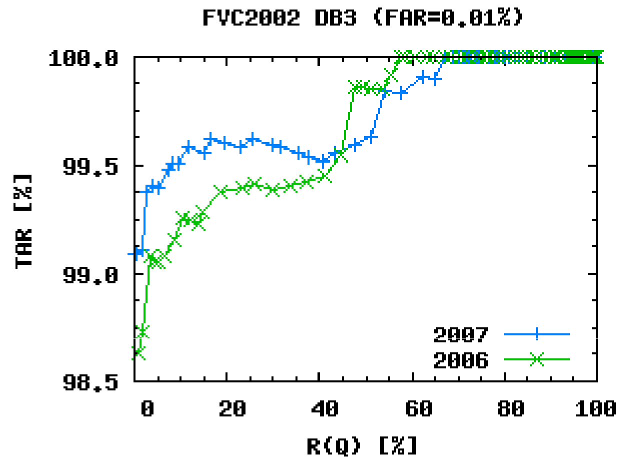



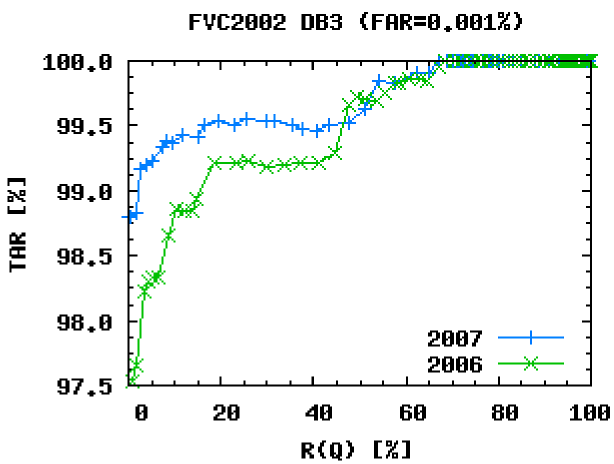



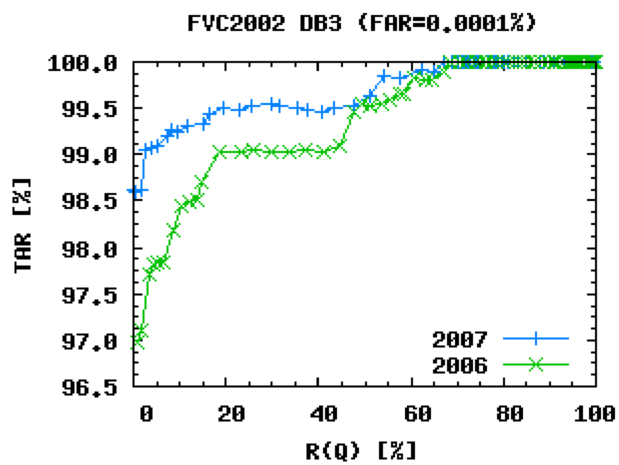
# Appendix: FVC2002 (6/13)

- **DB3: RRG comparison over varying FAR**



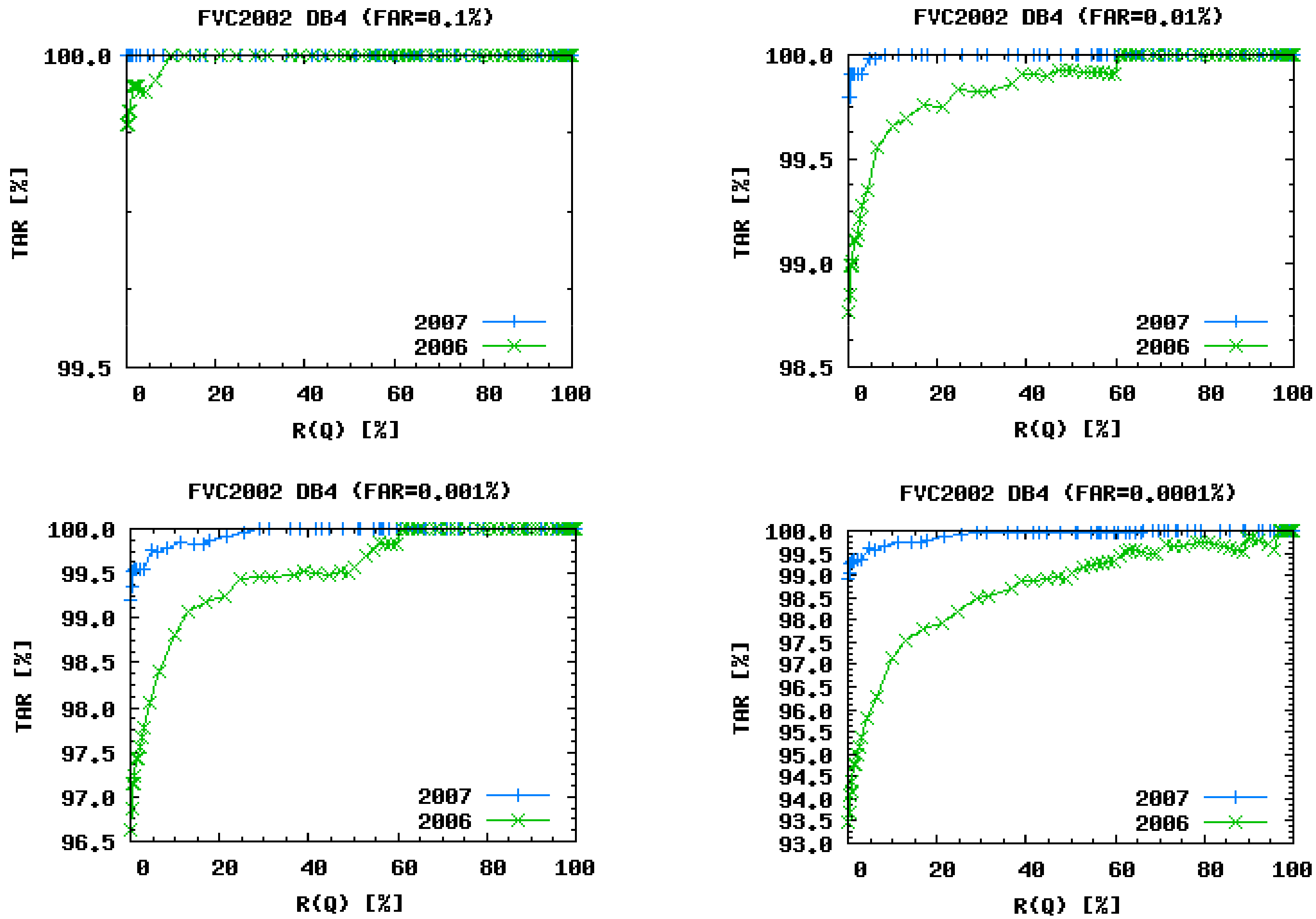



# Appendix: FVC2002 (7/13)

- **DB4: RRG comparison over varying FAR**



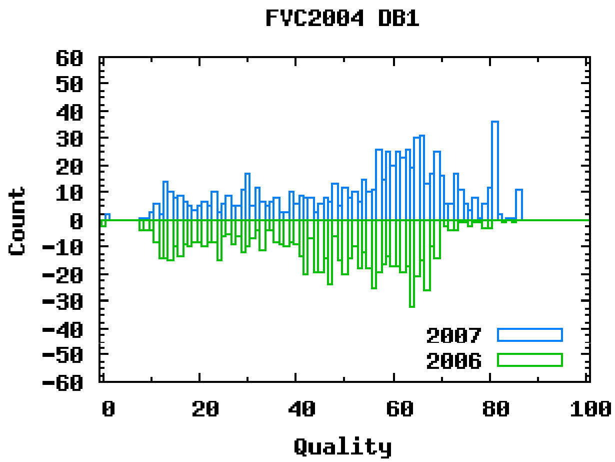
# Appendix: FVC2004 (8/13)

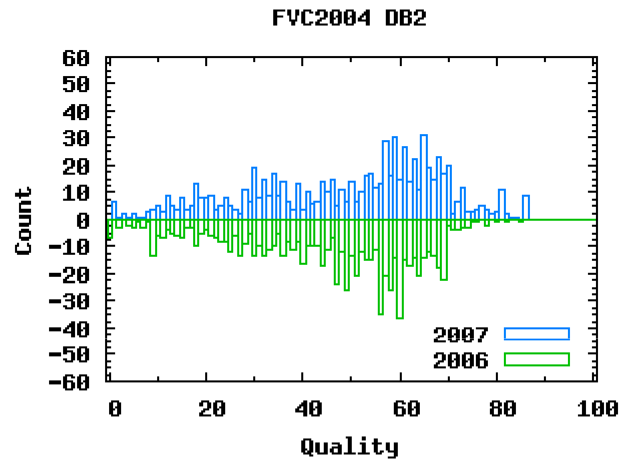
- **Accuracy Improvement**

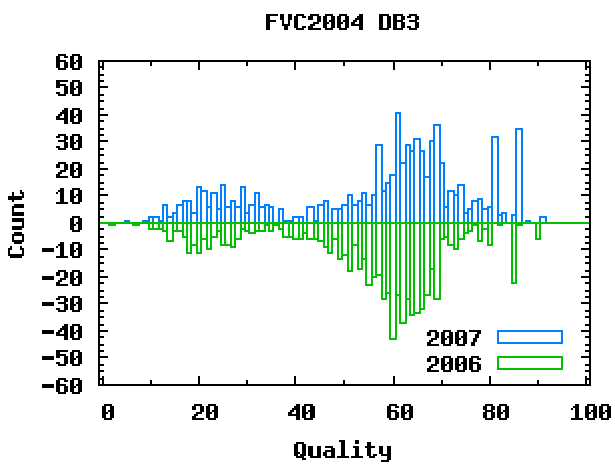
| | Speed | | TAR at FAR=0.01% [%] | | | |
|---|---|---|---|---|---|---|
| | Match | FE | DB1 | DB2 | DB3 | DB4 |
| **SDK H4 (2007)** | **H-equiv.** | **Slow** | **96.68 (+0.93)** | **97.13 (+0.58)** | **99.14 (+0.07)** | **99.46 (+0.69)** |
| SDK H3 (2006) | H-equiv. | Slow | 95.75 | 96.55 | 99.07 | 98.77 |
| SDK H2 | H-equiv. | H-equiv. | 95.66 | 95.09 | 98.70 | 97.96 |
| SDK H | See NISTIR7151 | | 93.63 | 94.88 | 97.79 | 97.02 |

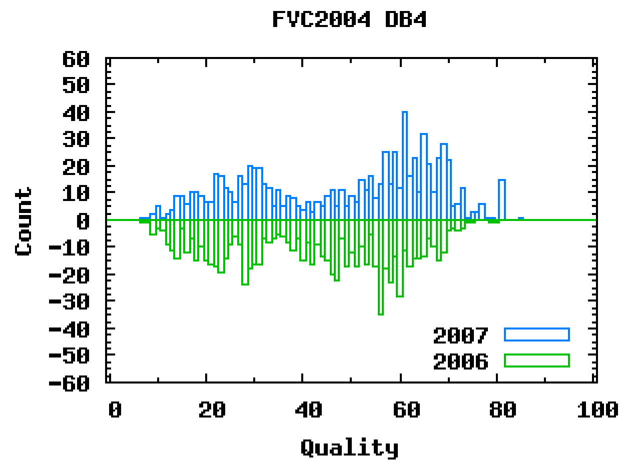
# Appendix: FVC2004 (9/13)

- Quality Distributions



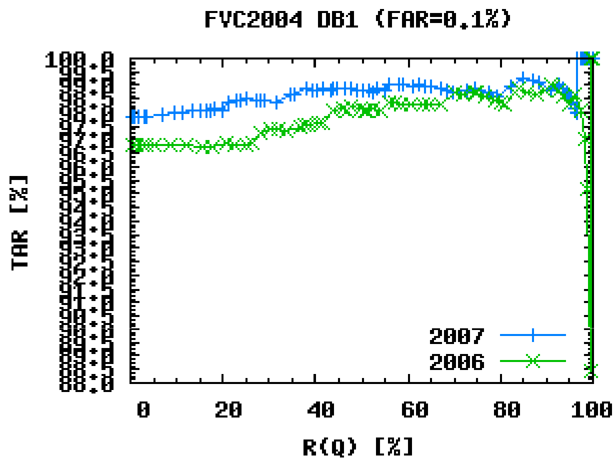



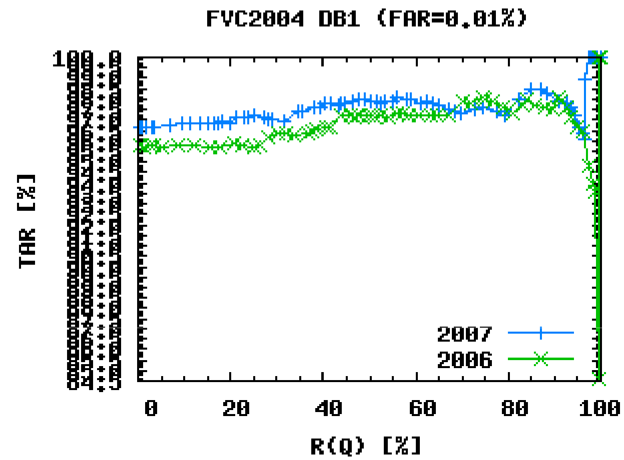



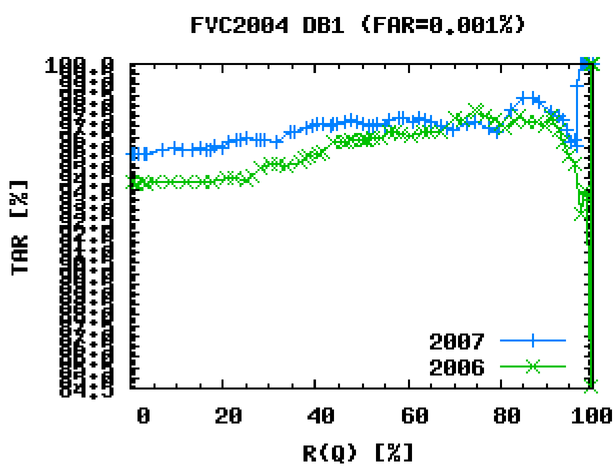



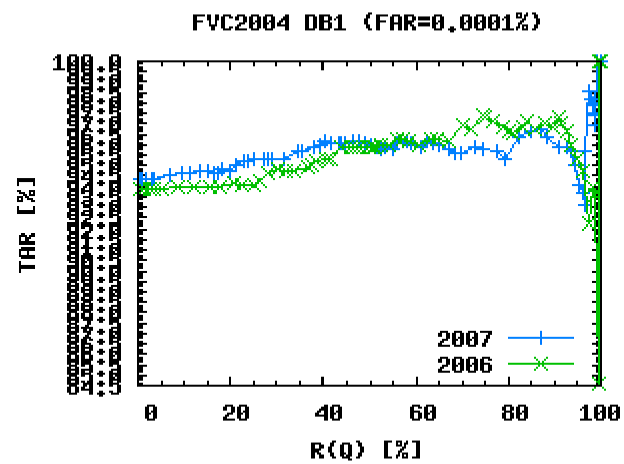
# Appendix: FVC2004 (10/13)

- **DB1: RRG comparison over varying FAR**



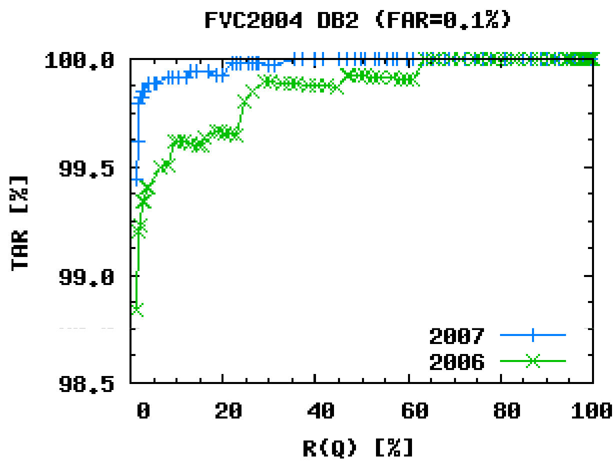



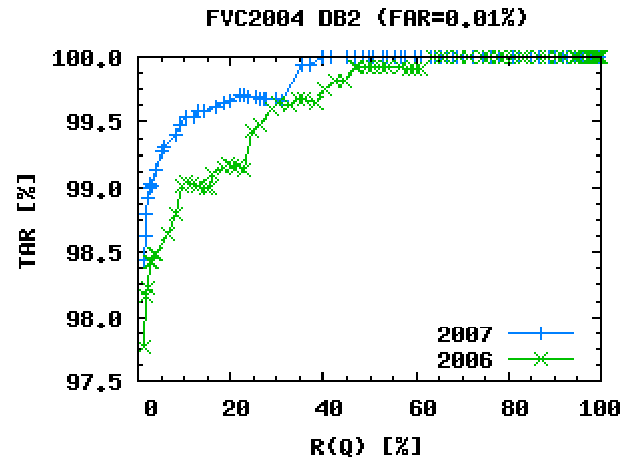



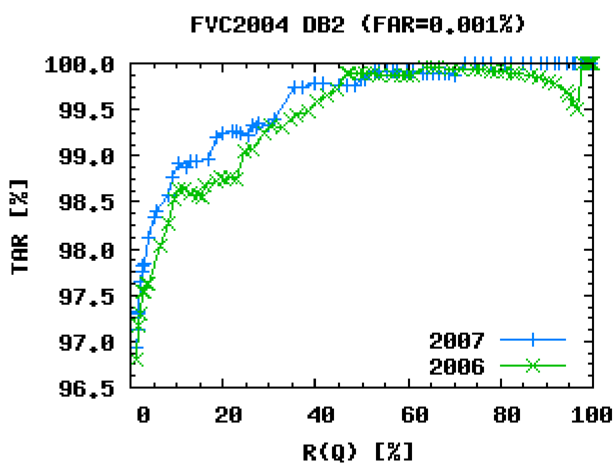

  

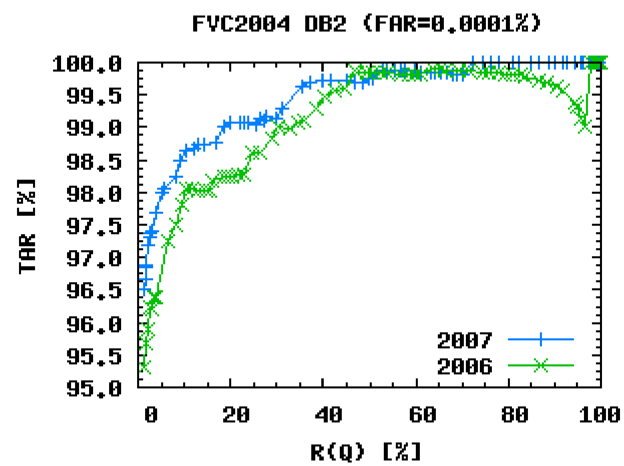
# Appendix: FVC2004 (11/13)

- **DB2: RRG comparison over varying FAR**



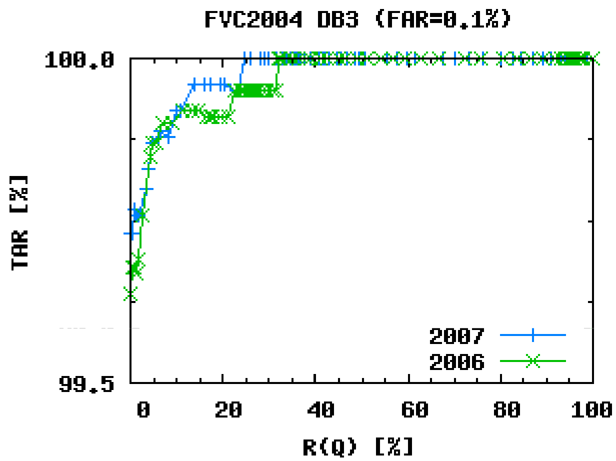



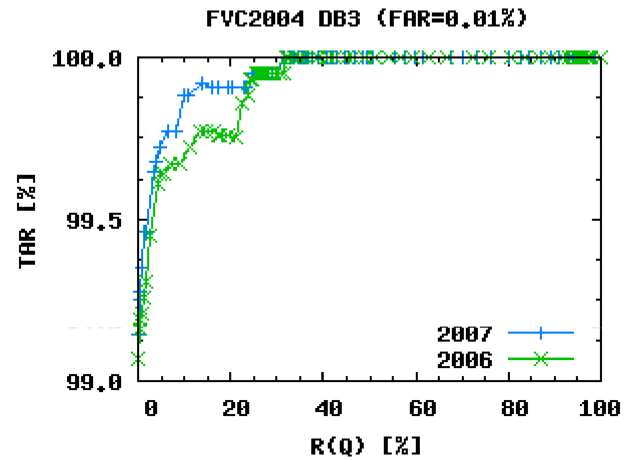



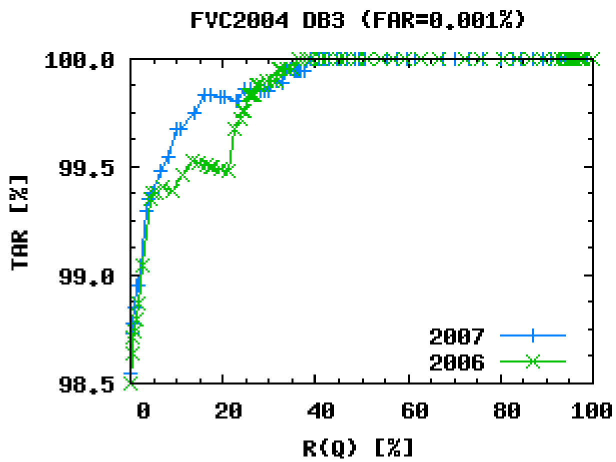

   

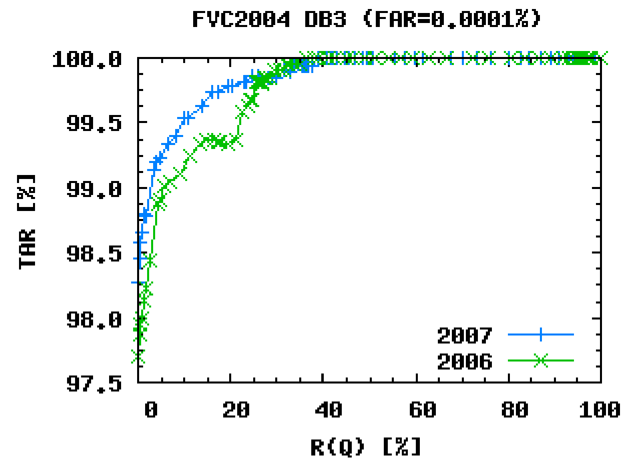
# Appendix: FVC2004 (12/13)

- **DB3: RRG comparison over varying FAR**



# Appendix: FVC2004 (13/13)

- **DB4: RRG comparison over varying FAR**