On The Uniqueness of Fingerprints

Anil K. Jain

Dept. of Computer Science and Engineering

Michigan State University

Jain@cse.msu.edu

Fingerprints

"Leaving aside microscopic peculiarities which are of unknown magnitudes,…..out of the 4000 cubic inches or so of flesh, fat, and bone of a single human body, there are many that are visible with or without the aid of a lens."

"Perhaps the most beautiful and characteristic of all superficial marks are the small furrows with the intervening ridges and their pores that are disposed in a singularly complex yet even order on the under surfaces of the hands and feet."

Francis Galton, Nature, June 28, 1888

Fingerprint Identification

Based on two basic premises

- Persistence: The basic characteristics of fingerprints do not change with time
- Individuality: The fingerprint is unique to an individual

The validity of persistence has been established

The uniqueness of fingerprints has been accepted over time because of lack of contradiction and relentless repetition

Fingerprint Identification Systems

- Scotland Yard adopted fingerprints for identification in 1900
- FBI installed IAFIS in 2000
- Current database size is ~ 50 million 10 prints
- Conducts an average of 50,000 searches/day;
- ~15% of searches are in lights out mode
- Response time: ~2 hours for criminal search

There is an overwhelming amount of discriminatory information in the fingerprints. But, how much?

Are Fingerprints Unique?

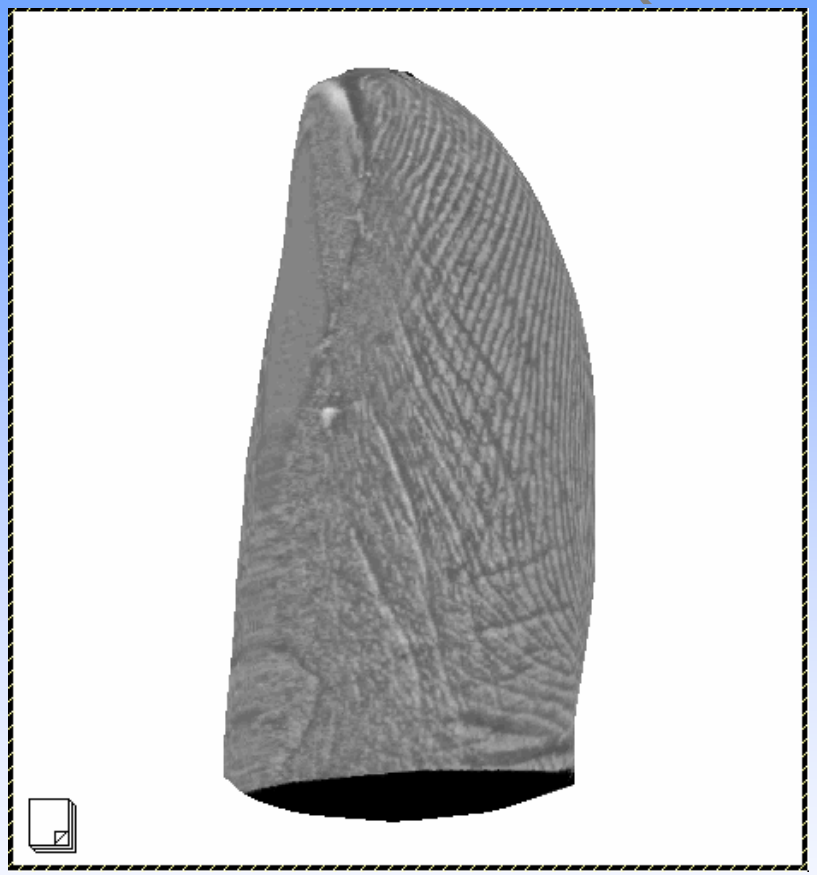
- "Only Once during the Existence of Our Solar System Will two Human Beings Be Born with Similar Finger Markings". *Harper's headline, 1910*

- "Two Like Fingerprints Would be Found Only Once Every 10^{48} Years". *Scientific American, 1911*

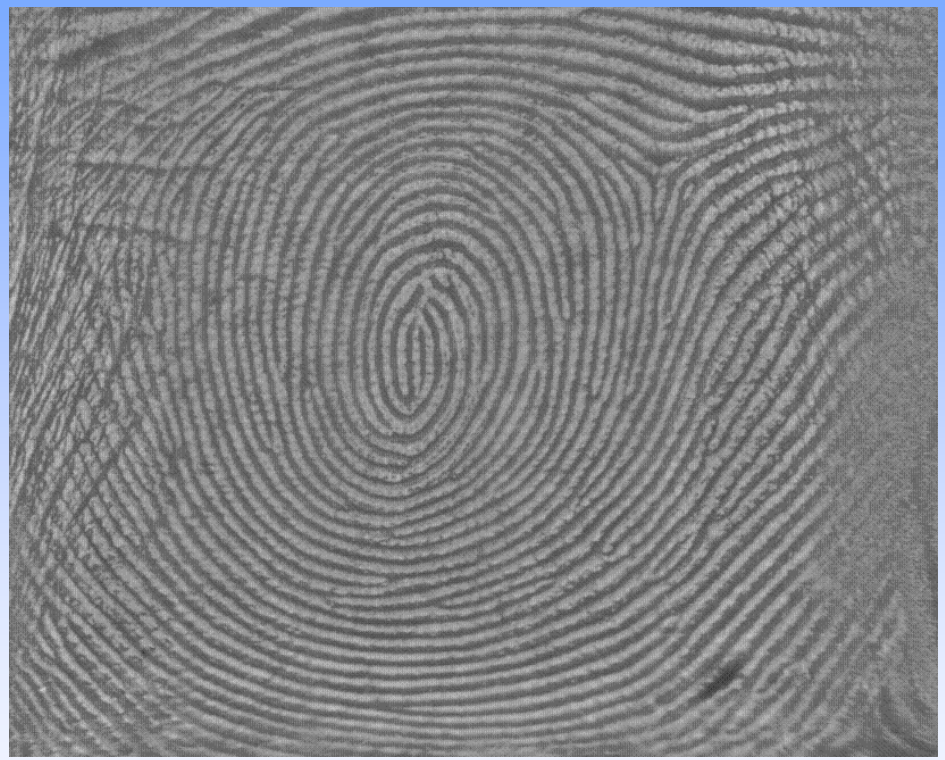
- "They left a mark - on criminology and culture. But what if they're not what they seem?" *Simon Cole, 2001*

"The time is ripe for the traditional forensic sciences to replace antiquated assumptions of uniqueness and perfection with more defensible empirical and probabilistic foundation." Saks and Koehler, Science, Aug 5, 2005

Finger Marks
(Ground Truth)



Touchless 3D image



Touchless "rolled" image

Courtesy: TBS North America, Inc.

Finger Impressions

Are the impressions (prints) true representations of the finger mark?

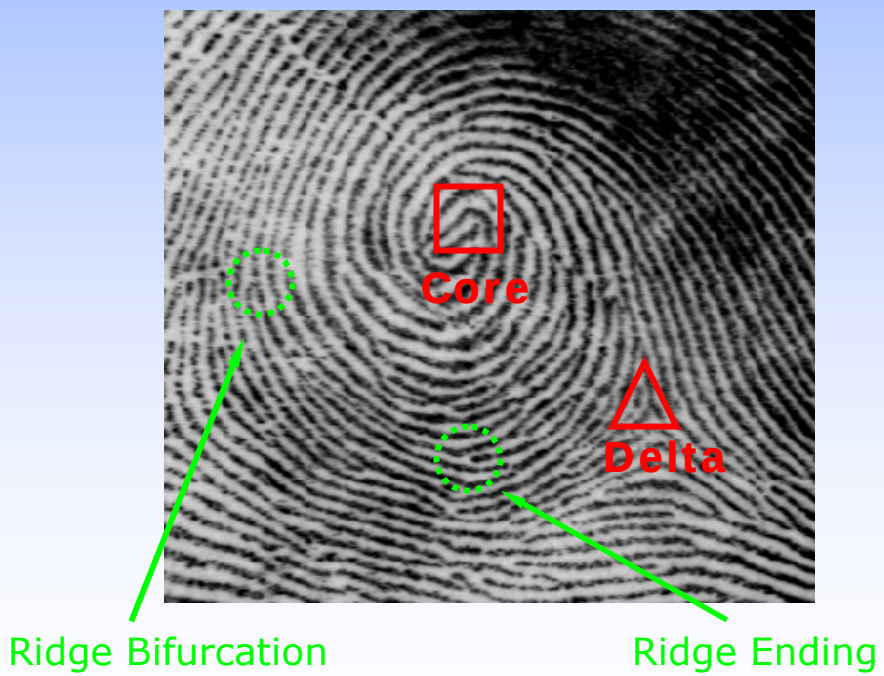
Cross Match
500 dpi - Rolled

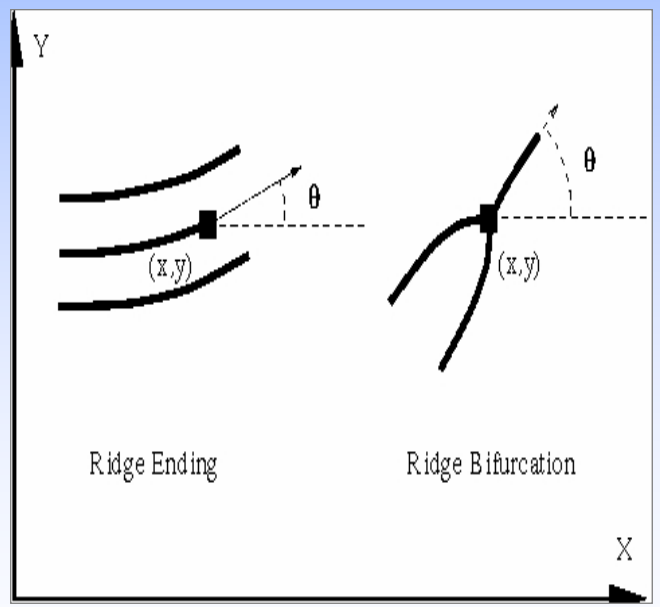
Cross Match
500 dpi - Slap

Cross Match
1000 dpi - Slap

Fingerprint Representation

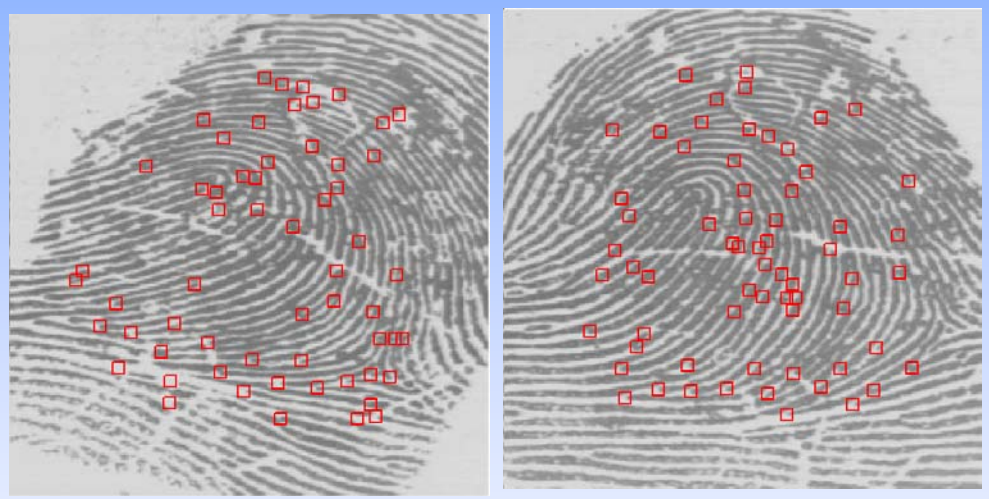
- Local characteristics (minutiae): ridge ending and bifurcation
- Singular points (core and delta): discontinuity in ridge orientations



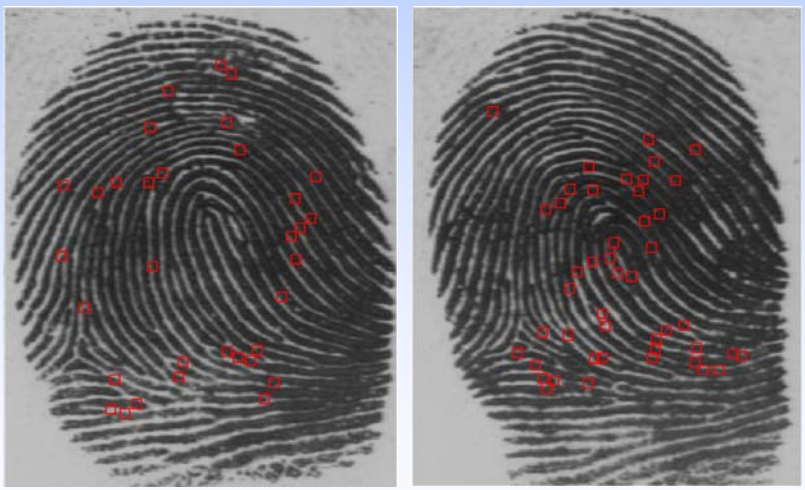


Fingerprint Matching

- Find the similarity (proportional to the no. of matched minutiae pairs) between two fingerprints



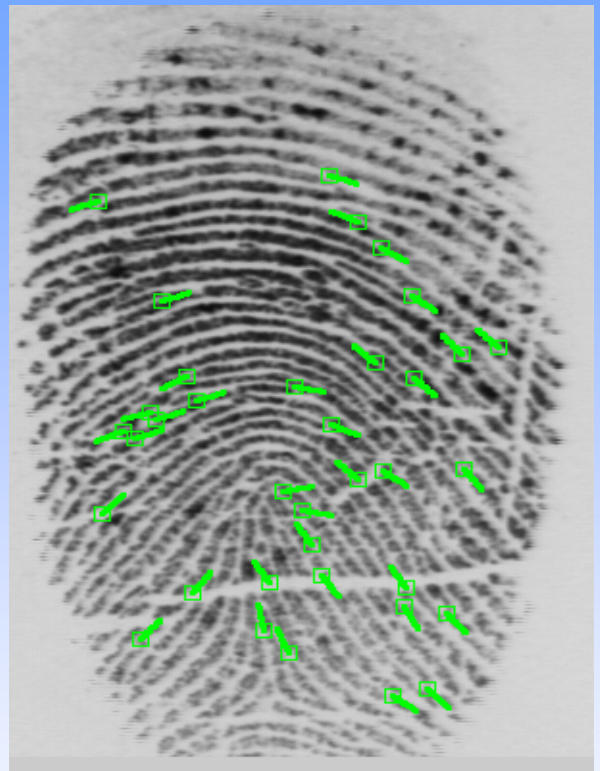
Fingerprints from the same finger
(intra-class variability)



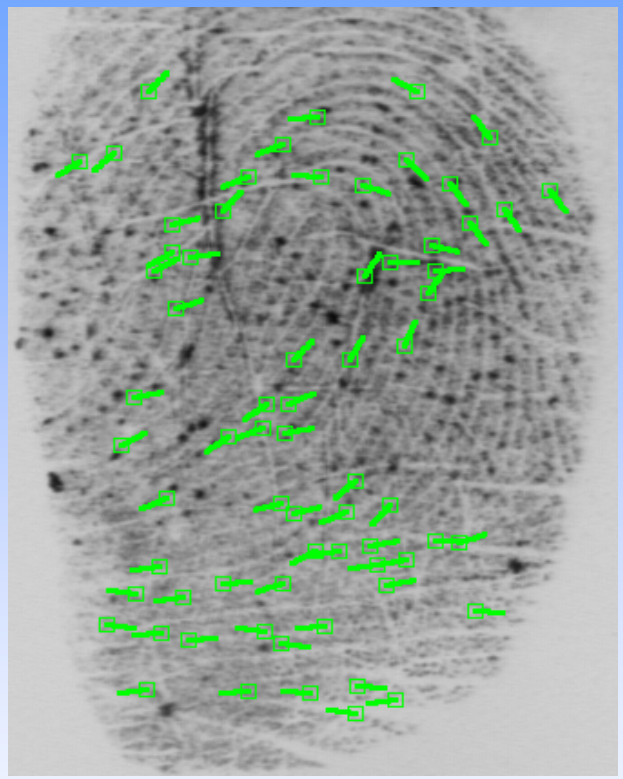
Fingerprints from two different fingers
(inter-class similarity)

Image Quality

No. False Minutiae = 0



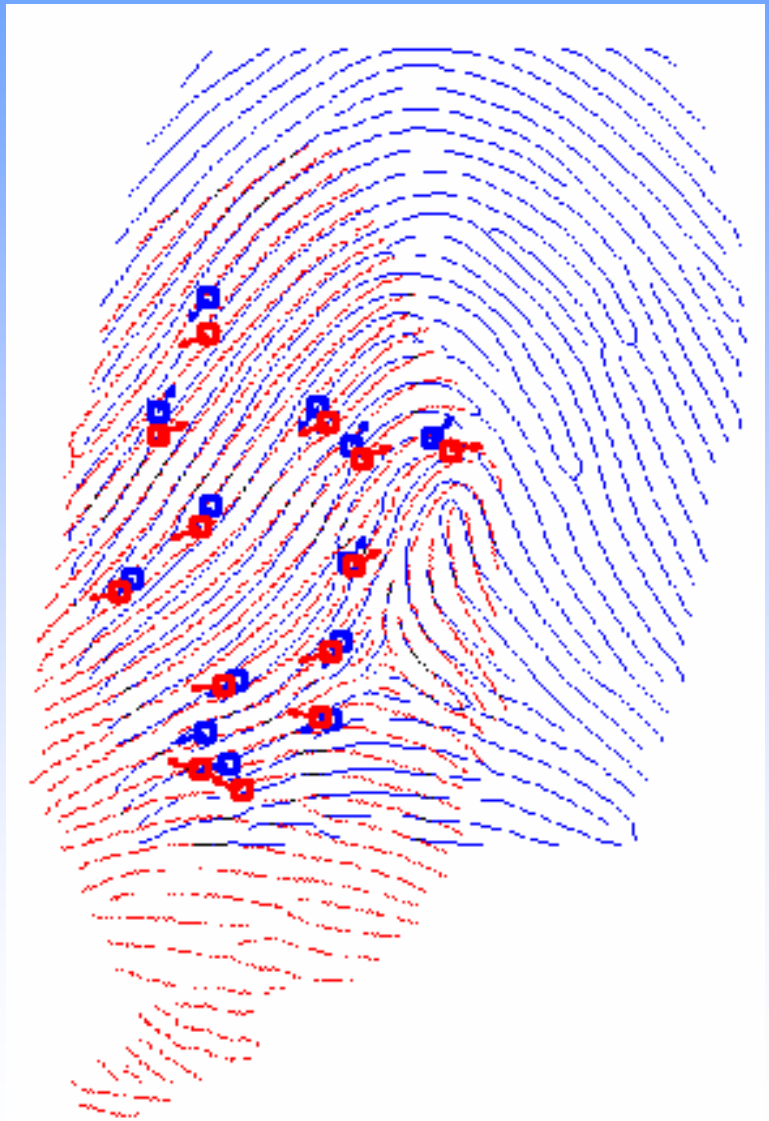
No. False Minutiae = 7



No. False Minutiae = 27

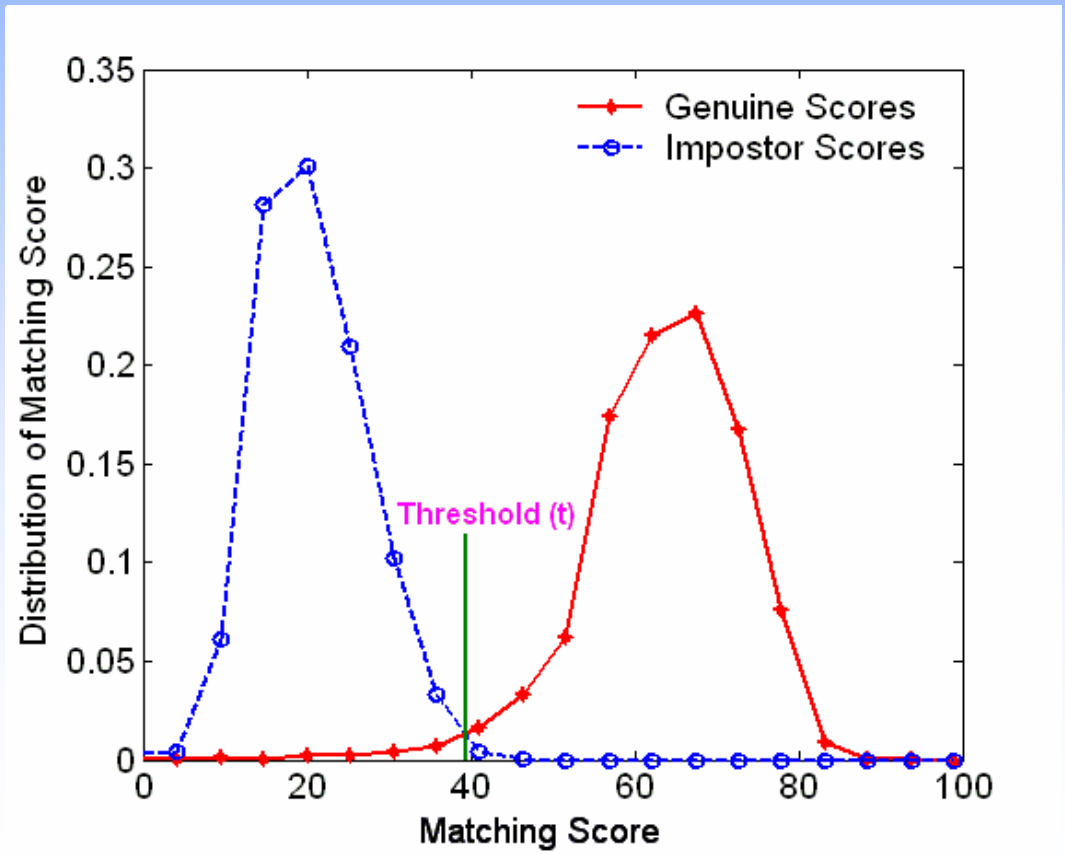
Poor image quality leads to missing and spurious minutiae

Alignment



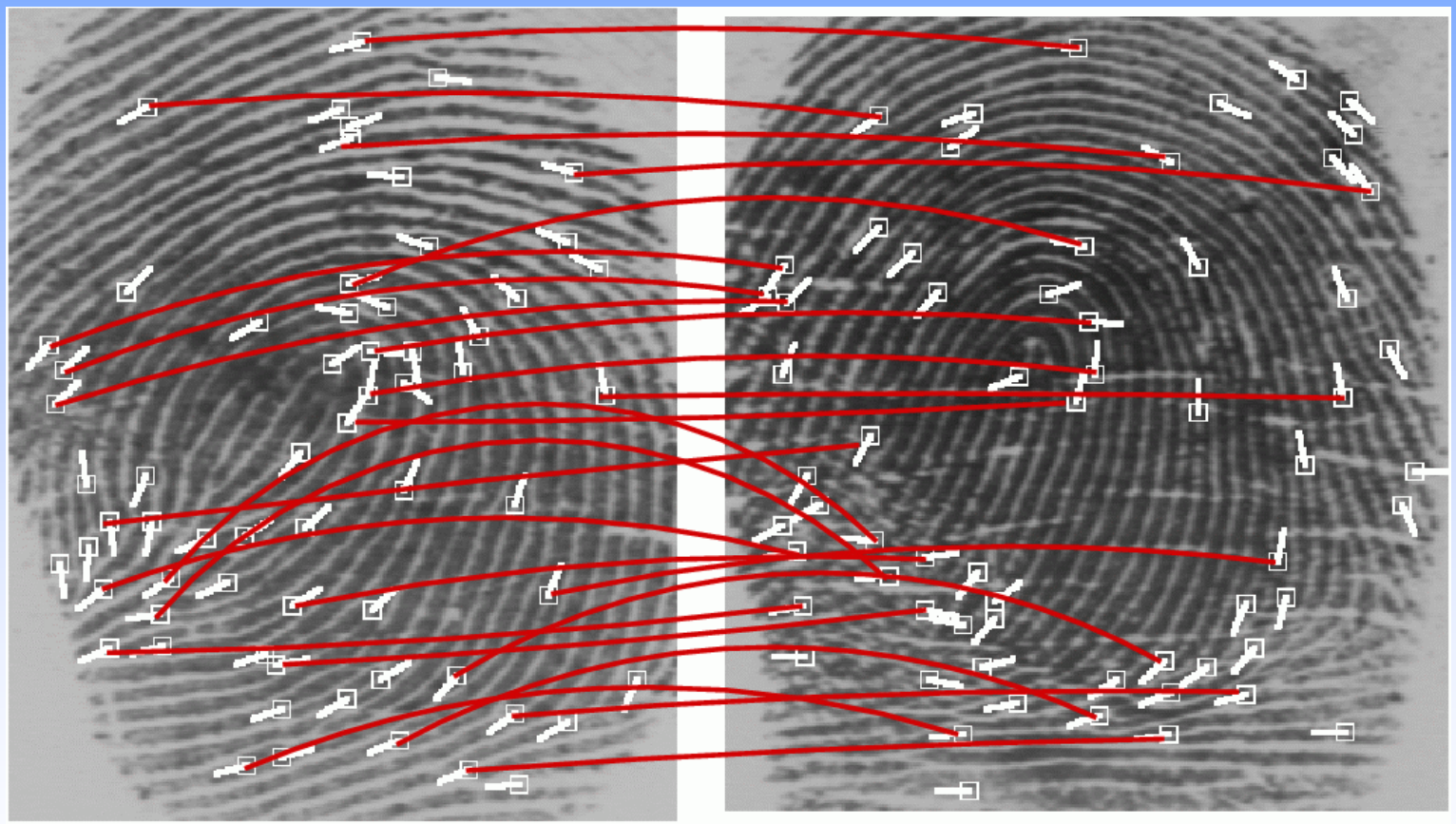
Matching Score Distributions

- Performance depends on the database. FVC2002 Database
- For FAR = 0.1% (1 in 1000), GAR = 97.1%
- EER = 1.65%; at 0% False Accept, FRR = 4%



Fingerprint Individuality

Question: Given a fingerprint query, what is the probability of finding a sufficiently similar fingerprint in a target population?



What is the probability of finding w false correspondences between two fingerprints containing m and n minutiae?

Approaches to Fingerprint Individuality

- Empirical Approach: (i) Collect representative samples of fingerprints; (ii) choose a fingerprint matcher; (iii) accuracy of the matcher on the samples provides an indication of the uniqueness of the fingerprints w.r.t. matcher
- Theoretical Approach: (i) Model all realistic phenomena affecting intra-class and inter-class fingerprint variations; (ii) given the similarity metric, theoretically estimate the probability of a false correspondence

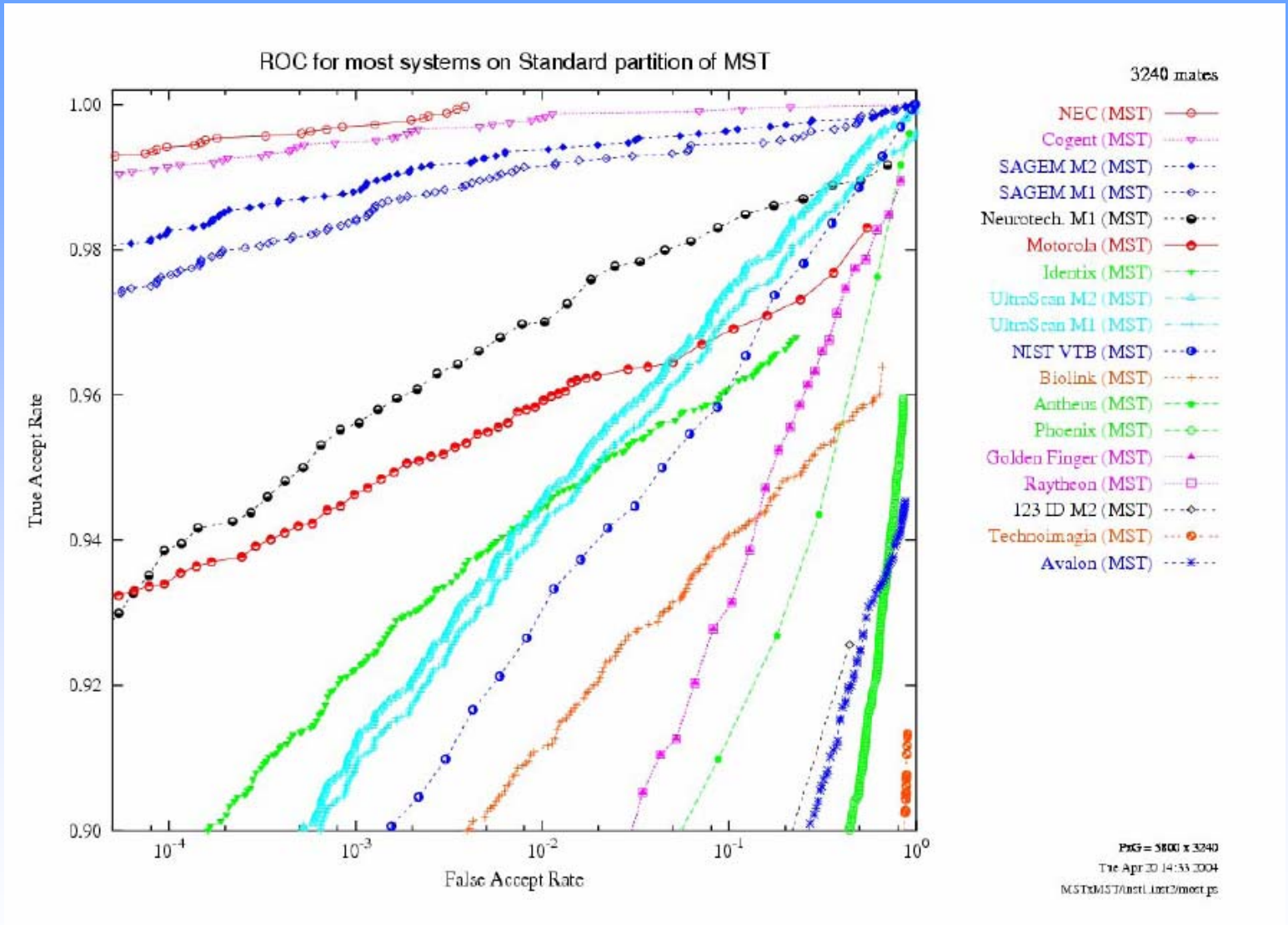
How well does the model conform to reality?

Empirical Approach
(Fingerprint Vendor Technology Evaluation)

Test	Compares	Database Size	# Comparisons
Large Scale Test (LST)	Sets of 1-10 images (Flat, Slap, Rolled; various combinations of fingers)	Approx. 48,000 fingerprint sets	1.044 billion set-to-set comparisons
Medium Scale Test (MST)	Single images (Flat and Slap Right Index)	10,000 images	100 million single image comparisons
Small Scale Test (SST)	Single images (Flat Right Index; subset of MST)	1,000 images	1 million single image comparisons

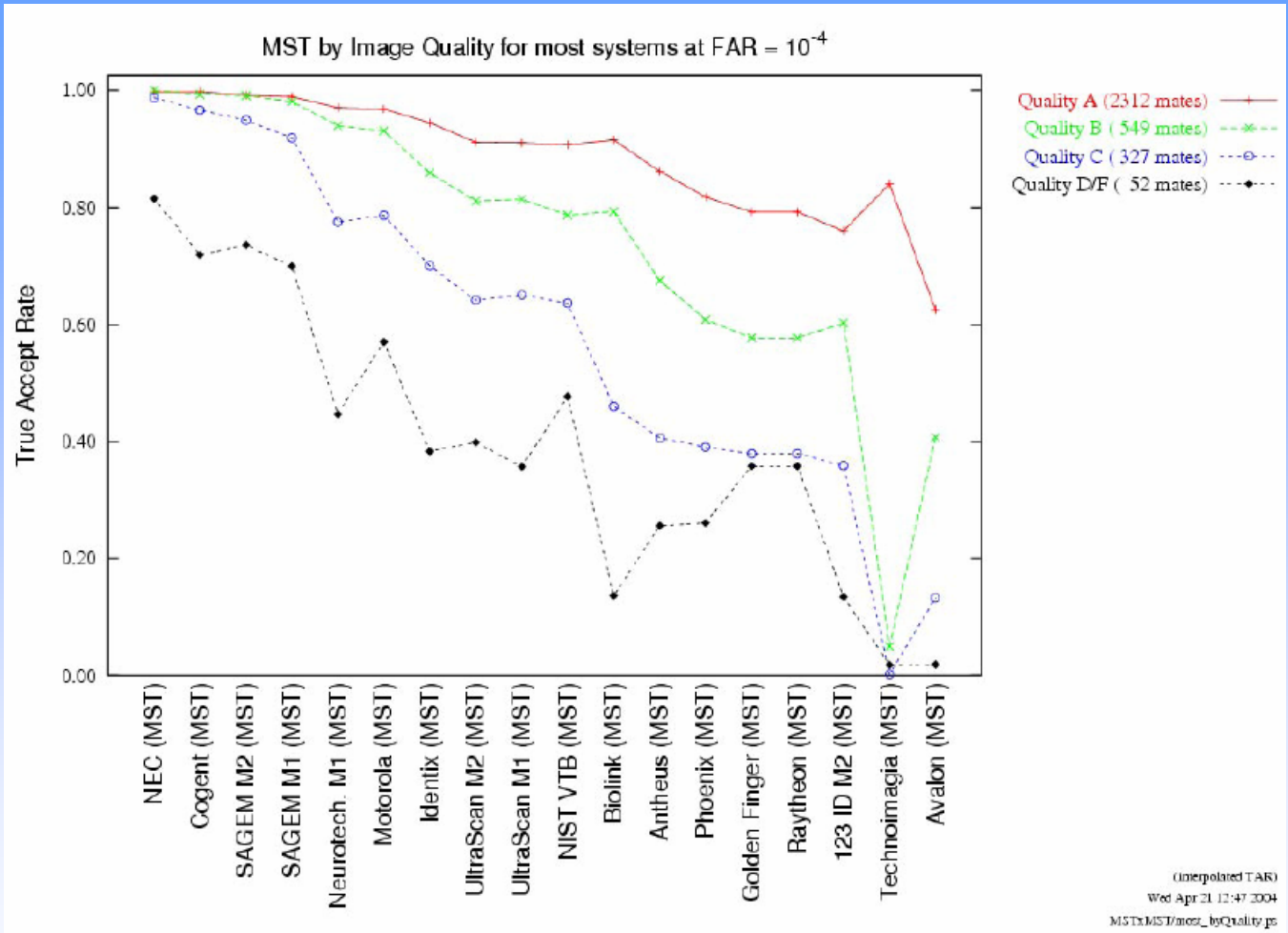
http://fpvte.nist.gov

MST Verification Performance



The order of the top 4 systems (NEC, Cogent, Sagem M1 and Sagem M2) was stable and clearly separated (> 97 % TAR at 0.01% FAR)

Effect of Image Quality



True Accept Rate (TAR) at 0.01% False Accept Rate (FAR)

Without exception, accuracy on good quality images was much higher than accuracy on poor quality images. Some systems were extremely sensitive to image quality. Low quality mostly led to false non-matches

Theoretical Approach

- The total no. of degrees-of-freedom of the minutiae configuration space does not directly relate to the discriminability of different fingers
- There are several sources of variability in multiple impressions of fingerprint that lead to detection of spurious minutiae or missing genuine minutiae and deformation of genuine minutiae
- Most of the earlier approaches did not account for this intra-class variation in their models and hence overestimated fingerprint individuality (gave a lower prob. of random correspondence)

Probability of Fingerprint Configuration

Author	P(Configuration)	P(Conf.)
Galton (1892)	$\frac{1}{16}\times\frac{1}{256}\times\left(\frac{1}{2}\right)^{R}$	1.45×10^{-11}
Pearson (1930)	$\frac{1}{16}\times\frac{1}{256}\times\left(\frac{1}{2}\right)^{R}$	1.09×10^{-41}
Henry (1900)	$\left(\frac{1}{4}\right)^{N+2}$	1.32×10^{-23}
Balthazard (1911)	$\left(\frac{1}{4}\right)^{N}$	2.12×10^{-22}
Bose (1917)	$\left(\frac{1}{4}\right)^{N}$	2.12×10^{-22}
Wentworth & Wilder (1918)	$\left(\frac{1}{50}\right)^{N}$	6.87×10^{-62}
Cummins & Midlo (1943)	$\frac{1}{31}\times\left(\frac{1}{50}\right)^{N}$	2.22×10^{-63}
Gupta (1968)	$\frac{1}{10}\times\frac{1}{10}\times\left(\frac{1}{10}\right)^{N}$	1.00×10^{-38}
Roxburgh (1933)	$\frac{1}{1000}\times\left(\frac{1.5}{10\times2.412}\right)^{N}$	3.75×10^{-47}
Trauring (1963)	$(0.1944)^{N}$	2.47×10^{-26}
Osterberg et al. (1980)	$(0.766)^{M-N}\times(0.234)^{N}$	1.33×10^{-27}
Stoney (1985)	$\frac{N}{5}\times0.6\times\left(0.5\times10^{-3}\right)^{N-1}$	1.20×10^{-80}

Probability of a fingerprint configuration using different models. We assume that an average fingerprint has R=24 Galton regions, M=72 Osterburg regions, and N=36 minutiae.

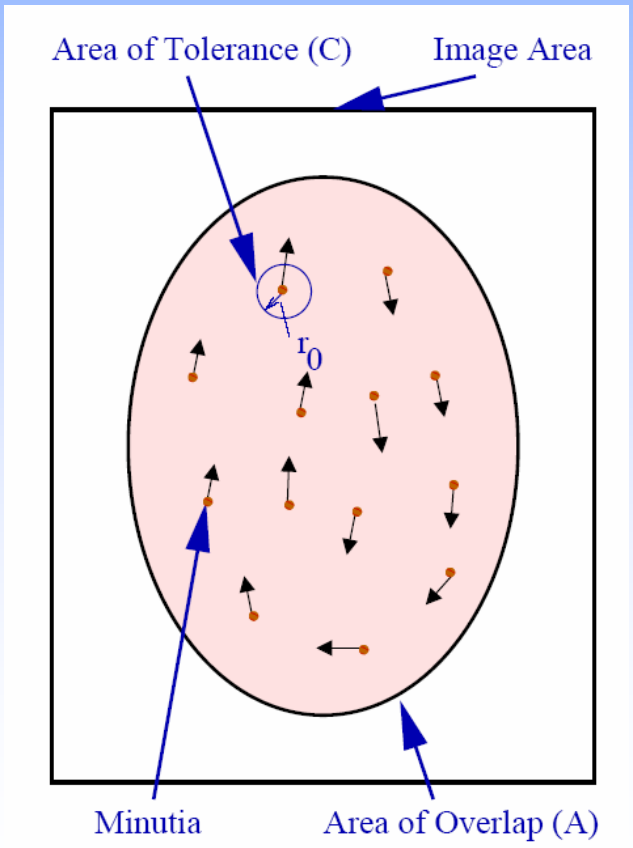
Uniform Model

- Considered only minutiae features
- Minutiae locations and directions are independent
- Minutiae locations are uniformly distributed
- Correspondence of a minutiae pair is an independent event
- "Quality" is not explicitly taken into account
- Ridge frequency is assumed to be constant
- Hyper-geometric distribution for the number of corresponding minutiae based on location alone
- Binomial distribution for the number of corresponding minutiae based on orientation alone

Pankanti, Prabhakar and Jain "On The Individuality of Fingeprints" IEEE Trans. PAMI, Vol. 24, No. 8, pp. 1010-1025, 2002

Probability of Random Correspondence

$$PRC(M,m,n,w) = \sum_{\rho=w}^{\min(m,n)} \left(\frac{\binom{m}{\rho}\binom{M-m}{n-\rho}}{\binom{M}{n}} \times \binom{\rho}{w} (l)^{w} (1-l)^{\rho-w} \right)$$



M = (A /d) / ($2r_0$)

m = no. of minutiae in template

n = no. of minutiae in query

w = no. of matching minutiae based on location and direction

ρ = no. of matching minutiae based on location alone

A = area of overlap between input and template

d = ridge period

r_0 = tolerance in minutiae location

l = $P(\min(|\theta'_i-\theta_j|, 360-|\theta'_i-\theta_j|) \leq \theta_0)$

θ_0 = tolerance in minutiae location

M=52, m=n=w=26, PRC = 2.40×10^{-30}

M=52, m=n=26, w=10, PRC = 5.49×10^{-4}

12-point Guideline

The effects of the fingerprint matcher misjudgments in using the *12-point guideline* is shown here. The source of error could be in underestimating the minutiae detected in the latent print (n) or overestimating the correct number of matched minutiae (q); m=12 for all entries.

Except for (m=12, n=12, q=12) entry, all other entries represent incorrect judgments by the fingerprint expert. For instance, the (m=12, n=14, q=8) entry in the table indicates that although the fingerprint examiner determined that all the 12 input minutiae matched, there were indeed 14 minutiae in the input and only 8 correctly matched with the template.

n \ q	8	9	10	11	**12**
12	6.19×10^{-10}	4.88×10^{-12}	1.96×10^{-14}	3.21×10^{-17}	$\mathbf{1.22 \times 10^{-20}}$
13	1.58×10^{-9}	1.56×10^{-11}	8.42×10^{-14}	2.08×10^{-16}	1.58×10^{-19}
14	3.62×10^{-9}	4.32×10^{-11}	2.92×10^{-13}	9.66×10^{-16}	1.11×10^{-18}
15	7.63×10^{-9}	1.06×10^{-10}	8.68×10^{-13}	3.60×10^{-15}	5.53×10^{-18}
16	1.50×10^{-8}	2.40×10^{-10}	2.30×10^{-12}	1.45×10^{-14}	2.21×10^{-17}

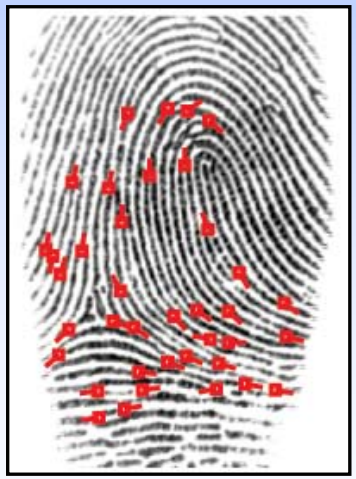
Mixture Model

- Given: N fingers, L impressions/finger
- Construct a master template for each finger by aligning minutiae from L impressions
- Fit a mixture model to the master template
- Simulate minutiae sets (with m or n minutiae) using the mixture models
- Obtain empirical distribution of the no. of impostor minutiae matches
- Fit a parametric distribution to the no. of impostor matches
- Compute the probability of false correspondence

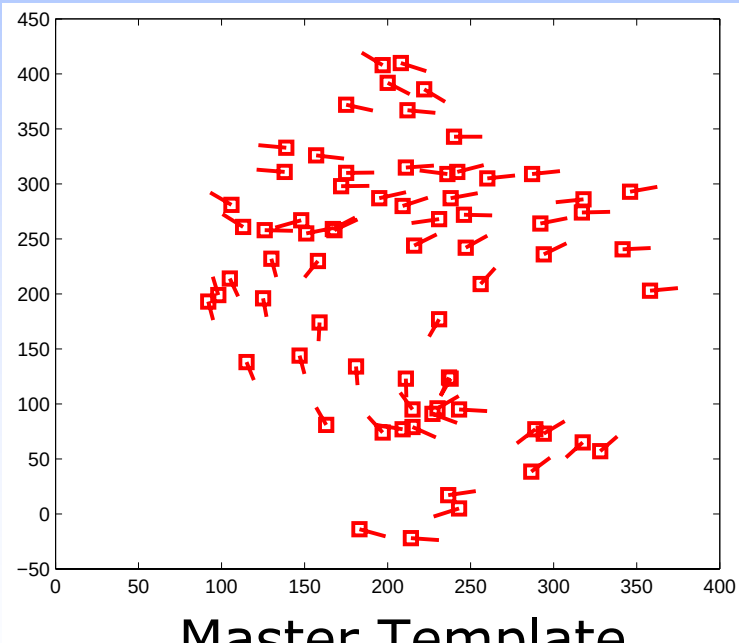
Dass, Wang and Jain 2005

Master Fingerprint

- Extract minutiae from all L impressions of a finger
- Choose a reference print (based on quality and area of overlap)
- Align minutiae from the other L-1 impressions to the reference



Reference



Master Template

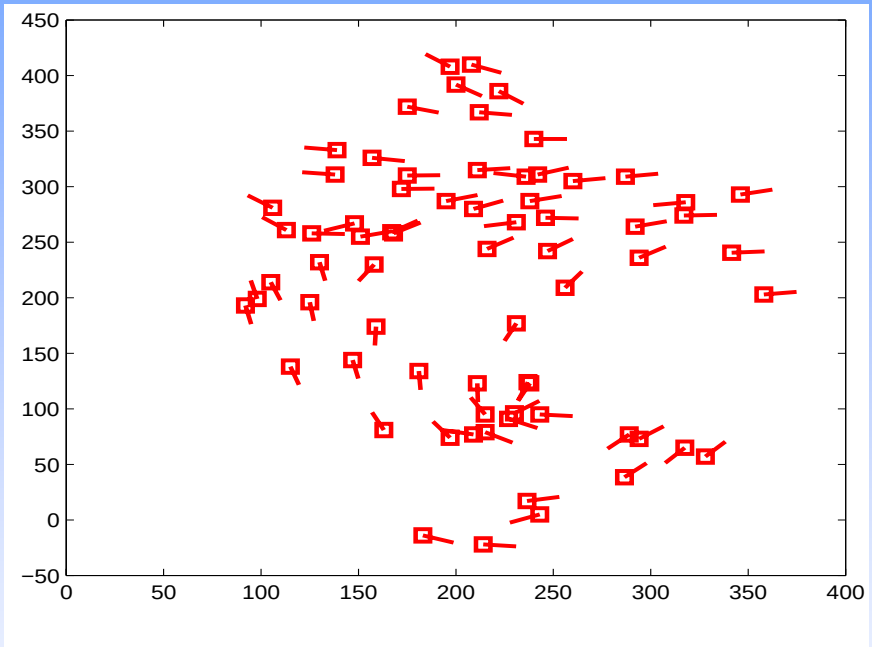
Minutiae Location & Direction Model

- To capture the clustered nature of the minutiae, fit a mixture model to minutiae location and orientation

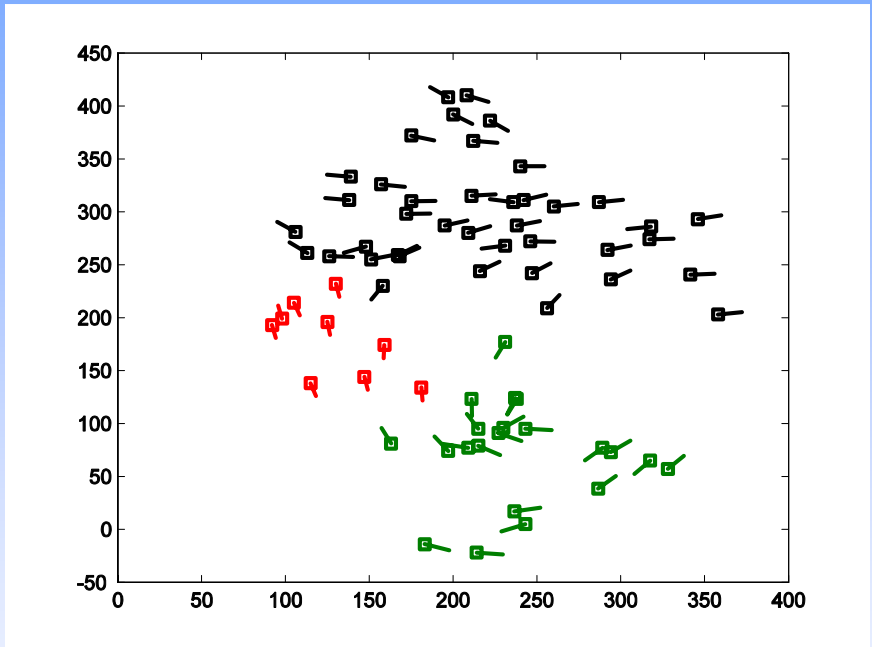
$$f(s,\theta \mid \Theta_G) = \sum_{g=1}^{G} \tau_g f_g^X \left(s \mid \mu_g, \Sigma_g\right) f_g^D \left(\theta \mid \nu_g, \kappa_g\right)$$

- Minutiae location is modeled using bivariate Gaussian distribution
- Minutiae direction is modeled using Von-Mises distribution
- EM algorithm is used to fit the mixture model; G is selected according to the Bayes Information Criterion

Mixture Model Fitting



Master Template

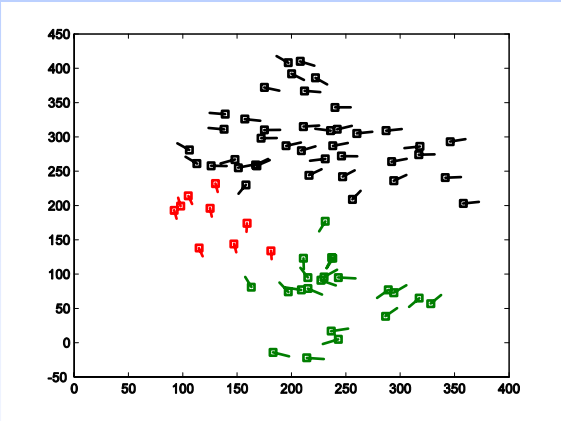


Fitted Mixture Model (G=3)

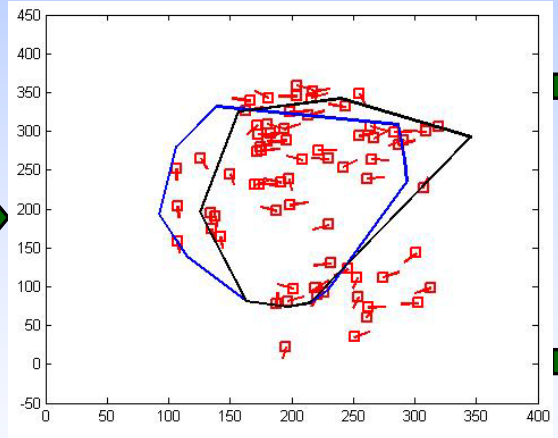
Mixture model is used to simulate prints with a specified no. of minutiae

Simulations From Mixture Model

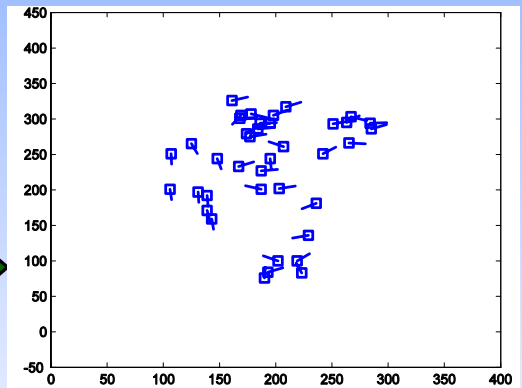
- Simulate a master print from the mixture model
- Generate L impressions from the synthetic master; Convex hulls come from original L prints
- Compare no. of matches based on simulated and true minutiae

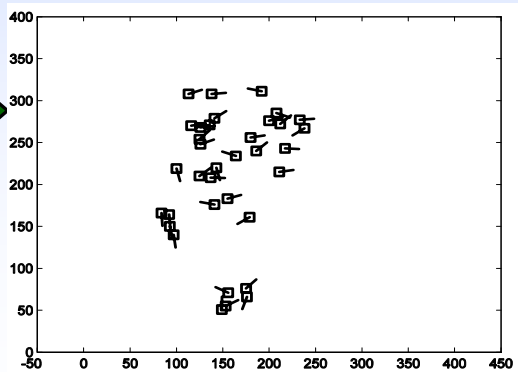


Original master



Synthetic master from the fitted mixture model



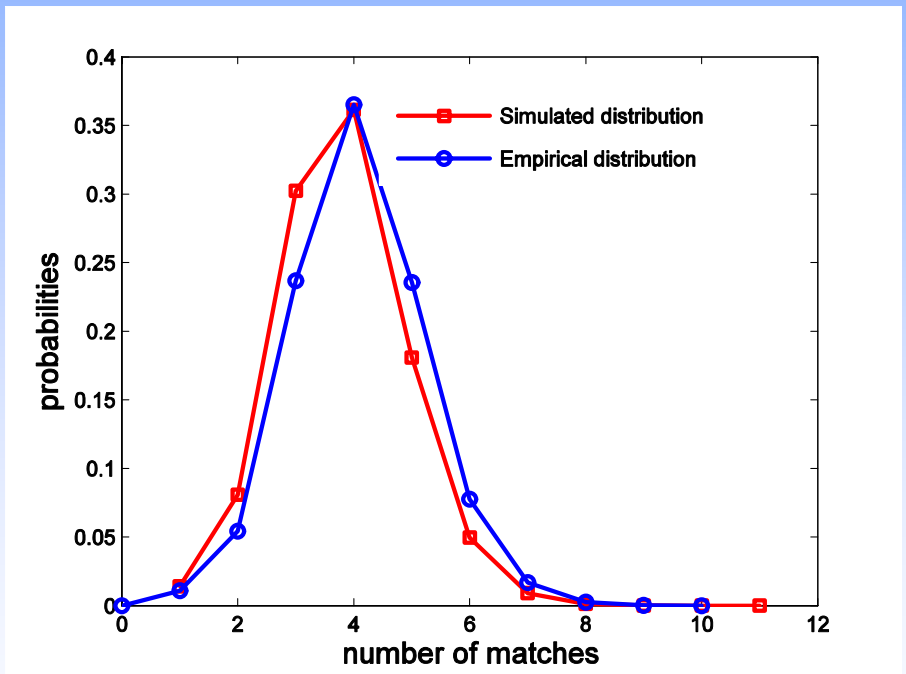


Sampled minutiae sets

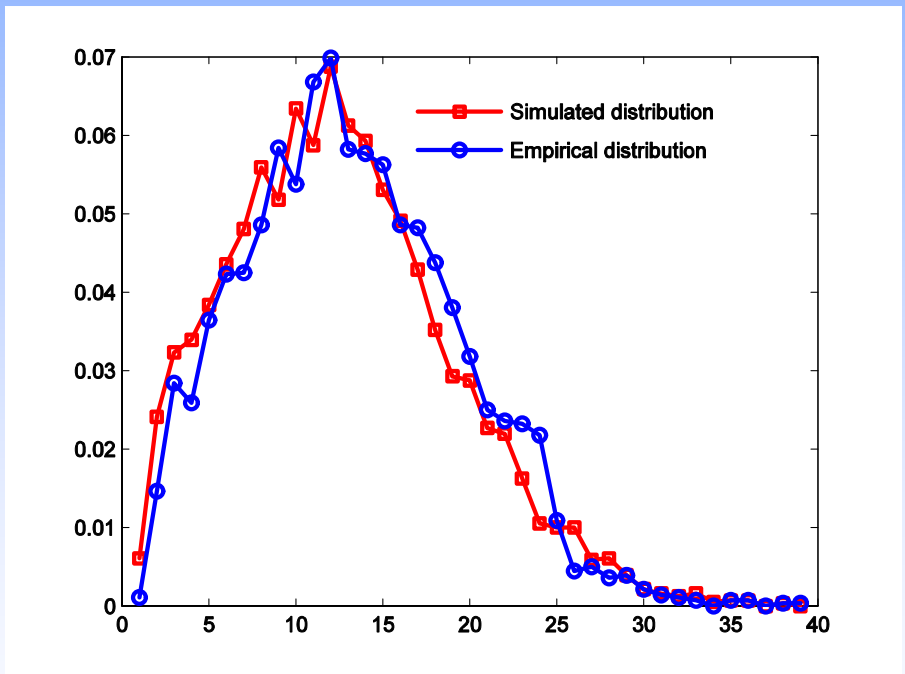
Validation of Mixture Model

- FVC 2002 DB1 database: 100 fingers, 8 impressions/finger; avg. no. of minutiae per impression = 27

63,3600 impostor matches

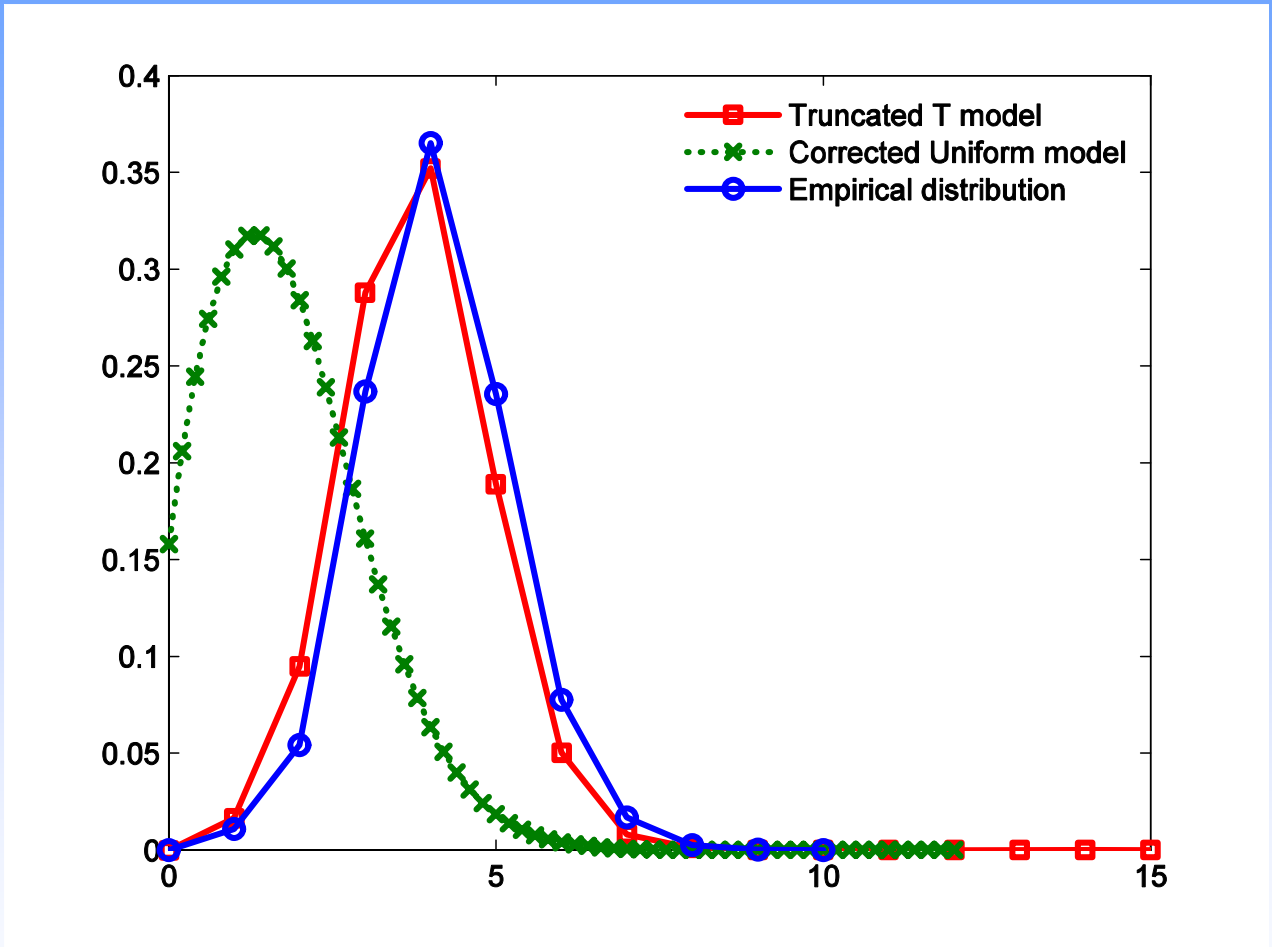


5,600 genuine matches



Histograms of the number of matches based on simulated minutiae (using mixture models) and true minutiae are similar

Modeling No. of Impostor Matches



Fit a truncated t distribution with p degrees of freedom and location (μ) and scale (σ) parameters to model the no. of impostor matches

Estimated PRC Values

(m, n, w)	PRC_u	PRC_m
(12, 12, 12)	1.4×10^{-12}	6.5×10^{-9}
(26, 26, 12)	1.2×10^{-9}	3.8×10^{-8}
(36, 36, 12)	4.0×10^{-7}	2.9×10^{-6}
(46, 46, 12)	2.1×10^{-5}	1.4×10^{-4}
(12, 12, 5)	4.9×10^{-4}	2.9×10^{-5}
(26, 26, 5)	1.2×10^{-2}	4.9×10^{-2}
(36, 36, 5)	5.1×10^{-2}	4.6×10^{-1}
(46, 46, 5)	1.2×10^{-1}	1.8×10^{-1}

PRC_u and PRC_m are the PRC values estimated from the uniform model (Pankanti, Prabhakar and Jain, 2003) and the mixture model, respectively

Summary

- What is the inherent discriminatory information available in fingerprints?
- Large scale data-dependent empirical performance evaluations are beginning to be conducted
- Challenge is to (theoretically) model not only the total variation present in the fingerprints, but also the variations of fingerprints of the same individual
- Only a few efforts have been made in statistical estimation of fingerprint individuality
- The available results can only be viewed as a first-order approximation to the answers that are needed