Biometrics

Lucía Montero Sanchis

Tuesday 30th January, 2018

Question 1 – Leading Biometric Technology

Physiological Characteristics

01. Fingerprint

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

02. Face (2D and 3D)

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

03. Iris

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

04. Periocular, Retina, Ear

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

05. Hand Geometry

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

06. Palmprint and Palm Veins

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

Behavioral Characteristics

07. Voice (02, 03)

Generalities:

- **Voice biometric** combines physiological and behavioral characteristics. Useful for remote-access transactions over telecommunication networks. Voice is subject to many sources of variability.
- **Disambiguation** – Voice recognition can refer to *Speaker recognition* (who is speaking) or *Speech recognition* (what is being said).
- **Perceptual Cues**: *High-level* (learned behaviors – Semantic, Dialogic, Idiolectal that depend on status, education, place of birth) vs *Low-level* (physical characteristics – Spectral, Prosodic, Phonetic that depend on anatomical structure)

Sensing:

- **Microphones** (e.g. in smartphones)
- **Speech signal** is real, continuous, non-stationary, 4-dimensional (4D), has finite energy. It's complex and variable over time. Sometimes periodic (pseudo-periodic) for voiced sounds; Sometimes random for fricative sounds; Sometimes impulsive in explosive phases of occlusive sounds.
- **Bandwidth** – Frequency Band of Telephone Speech: 300 Hz – 3.4 kHz
- **Coding Bands**:
  - Hi-Fi: 20 Hz – 20 kHz (sampling frequency 44.1 kHz)
  - Wideband: 50 Hz – 7 kHz (sampling frequency 16 kHz)
  - Narrow band: 300 Hz – 3.4 kHz (sampling frequency 8 kHz)
- **Sampling** and Uniform **Quantization**:

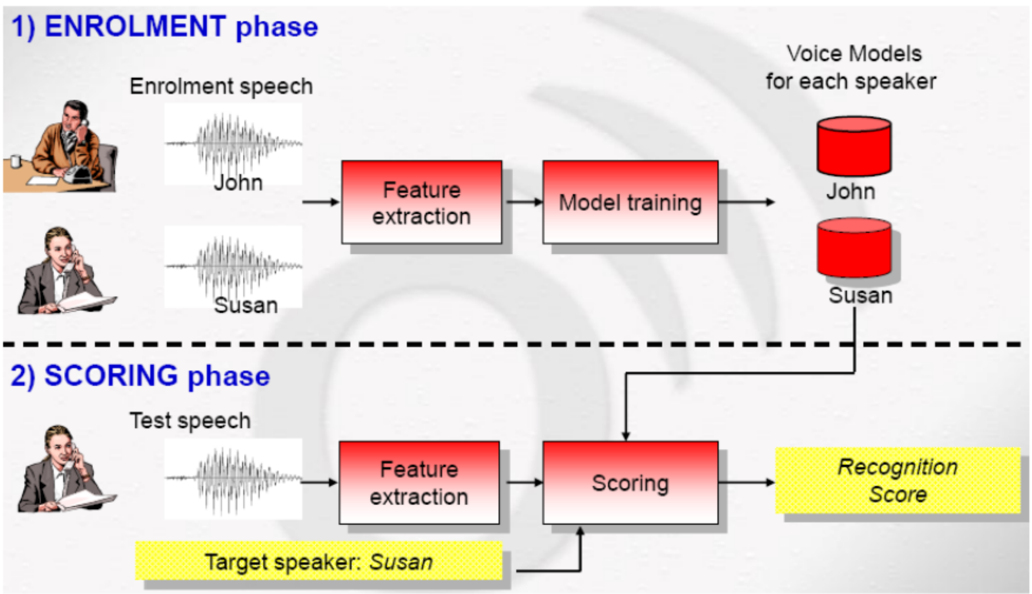


Figure 1: Enrolment and Scoring in Speaker Recognition

Feature extraction: Spectral envelope, MFCC

this??

- **Mel-Frequency Cepstral Coefficients (MFCCs)**: Mel-Frequency Cepstrum (MFC) is a representation of spectral energy of a sound on the mel scale, and is made up by the MFCCs (coefficients). In an MFC the frequency bands approximates the human auditory system's response – better representation of sound.
  - GMMs are used to capture the distribution of MFCCs in the feature space.
  - We enroll a speaker by adapting the UBM using the speaker's input.
  - Are obtained through FFT (Short-term transform), Logarithm, Deconvolution source/tract (cepstre). The acoustic vector consists of cepstrum, delta-cepstrum and delta-delta-cepstrum
  - **Short-Term Feature Extraction**:

    Window, frame, feature vector (acoustic vector)
  - **Short-term Processing**:

    window N, frame M, window and frame duration definitions
  - **Short-term DFT**
  - **Spectral envelope**

    (+spectrogram)
  - **Real Cepstrum**
  - **Dynamic features**
- Recent research on supervectors and i-vectors.

Templates:

- **Template** is a compact, electronic representation of a biometric sample that is created at the time of enrollment and stored in the system database for future reference and comparison. The process of creating a template and storing it in the database is called *enrollment*.
- **Speech Modalities**:
  - *Text-dependent*: System knows text spoken (e.g. find fixed phrase, prompted phrase). Used for strong control over user input. Improved system performance.
  - *Text-independent*: Not know text spoken (e.g. conversation). Less control over user input. More flexible and more difficult.
- **Vector Quantization**

  explain VQ

	Deterministic	**Statistical**
Text-dependent	Dynamic Time Warping (DTW)	Hidden Markov Model (HMM)
Text-independent	Vector Quantization (VQ)	Gaussian Mixture Model (GMM)

Table 1: *Templates and Models* in Speaker Recognition

Matching (comparison):

- **Comparative analysis** is a comparison between two biometrics, typically a tested sample and an enrollment (reference) template or model. The output of the comparison is a distance or score.
- **Speaker Recognition Systems**:

  - *Conventional Speaker Verification System*
    Enrollment:
    GMM training → Speaker-dependent GMM → Database
  - *SV with Verbal Information Verification*
    Automatic enrollment:
    Verbal Information Verification ⇒ Save for training → Verified pass-phrases for training → HMM training → Speaker-dependent HMM → Database
  - *Text-prompted Speaker Recognition*: Uses **speaker-specific phoneme models** as basic acoustic units. New prompted sentence every time – can't cheat with pre-recordings.

- **Prerequisites for good performance in Speaker Recognition**: Speakers must not disguise their voices; Recording conditions and signal processing techniques are known or controlled; Speech recorded in conditions similar to those of the test signal is available; Reference values for similarity measures established in similar conditions as the test signal; Decision thresholds calibrated according to reference values depending on the application.

08. Dynamic Signature

Generalities:

- *Behavioral* characteristic. Combines *knowledge* and *biometric*. Not *permanent* (invariant over time). Currently can be digitalized and is in ID cards.
- **Applications**: Signature forensics, Signature authentication, Signature surveillance, Digital Rights Management, Biometric cryptosystems.
- **Off-line** or **Static** – scanned from paper documents, written conventionally.
- **On-line** or **Dynamic** – written with electronic device. Dynamic information (pen tip location through time) usually available at high resolution, even if pen not in contact with paper.

Sensing:

- Tablets, smartphones, IKEA's and Sunrise's SignPad, UPS and SwissPost...

Feature extraction:

- Basic (Local) Features:
  1. X, Y coordinates
  2. Velocity
  3. Acceleration
  4. Pen azimuth (0 - 359 deg)
  5. Pen altitude (0 - 90 deg)
  6. Pressure
  7. First and second derivatives of feature
- Global features (more than 150)
  - Signature length, height, weight
  - Total signature time
  - Total pen-down and pen-up time
  - Avg., max. and min. velocity
- Pre-processing: Smoothing; Segmentation (determine beginning and end); Initial point alignment.
- Forgery: Zero-effort; Home-improved (based on static image); Over-the-shoulder (observe signing); Professional (skilled individuals). Over-the-shoulder + Home-improved combo is called *skilled.*

Templates (Models are DTW, GMM and HMM)

Matching (comparison):

- DET curves to compare different GMMs, HMM...
- Resistant to imposters, non-invasive, can be changed by users, fast and intuitive enrolment, fast verification, independent of native language user.
- Inconsistent signatures increase error rates, limited applications, for good accuracy a 5D pen is needed (costly), some people can't sign.

09. Gait, Typing Rhythm

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

Biological Traces

10. DNA

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

Synthetic Biometric Data Generation

11. Synthesis of Fingerprints

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

Multimodal Biometrics

12. Multimodal Biometrics

Generalities:

-

Sensing:

-

Feature extraction:

-

Templates:

-

Matching (comparison):

-

Miscellaneous

13. Quality and Ageing in Classication of Biometric Data

Question 2 – Other Topics

Fundamentals of Biometrics (01)

01. Identity and Biometrics

The role of Biometrics is to recognize a person by their body traits and link the body to an externally assigned identity.

Biometrics:

- *Biometrics* – automated recognition of individuals based on biological and behavioral characteristics
- *Biometry* – statistical and mathematical methods applicable to data analysis problems in the biological sciences
- *Biometric system* – automatic pattern recognition system that recognizes a person by verifying the authenticity of a specific biological and/or behavioral characteristic (biometric modality) they possess
- (forensic, judicial) *Anthropometry* – (identification of criminals by) measurement techniques of human body and its specific parts

Identity:

- *Identity* – whatever makes something the same or different.
- *Authentication* (identity verification) – process to link a physical person with a certain identity

People are identified by **three basic means**:

- Something they *have* – identity document or card, passport, birth certificate, token...
- Something they *know* – password, PIN, name, date of birth
- Something they *are* – human body

Security level of each solution:

$$\text{Know} \quad < \quad \text{Know} + \text{Have} \quad < \quad \text{Know} + \text{Are} \quad < \quad \text{Know} + \text{Have} + \text{Are}$$

Advantages of Biometric identifiers: *Security*; *Convenience*; *Audit trial*; *Avoid fraud*; *De-duplication*. Examples: automated comparison process occurs in seconds; can replace passwords (often forgotten, lost, or misappropriated); identity justification without paperwork.

Ideal biometric identifier: *Universality* (every person has it); *Uniqueness* (different for every person); *Permanence* (invariant in time); *Collectability* (measurable, practical); *Acceptability* (public has no strong objections).

Challenge of biometrics: *Scalability*, *Usability*, *Accuracy*.

Identifiable biometric characteristics: *Biological traces* (DNA, blood, saliva); *Biological (physiological) characteristics* (fingerprint, iris, retina, hand palm, hand veins, hand geometry, facial geometry); *Behavioral characteristics* (dynamic signature, gait, keystroke dynamics, lip motion); *Combined* (voice).

Comparison of biometric techniques (Cost and Accuracy):

Cost:	Voice	$<$	Face	$\approx$	Fingerprint	$<$	Signature	$<$	Hand	$<$	Iris
Accuracy:	Voice	$\approx$	Face	$\approx$	Signature	$<$	Hand	$<$	Fingerprint	$<$	Iris

02. Recognition, Verification, Identification and Authentication (01)

Recognition – used when we do not distinguish between verification, identification and authentication.

Verification – performs one-to-one comparison of a submitted biometric characteristic (sample) set against a specified stored biometric reference, and returns the comparison score and decision (deciding whether a sample belongs to a specified person) – *Is this person who he claims to be?*

Identification – performs one-to-many comparison/search to determine the identity of the user from a known set of identities – *Who is this person?*

Authentication – the user claims an identity and the system verifies whether the claim is genuine (link a person with a chosen identity).

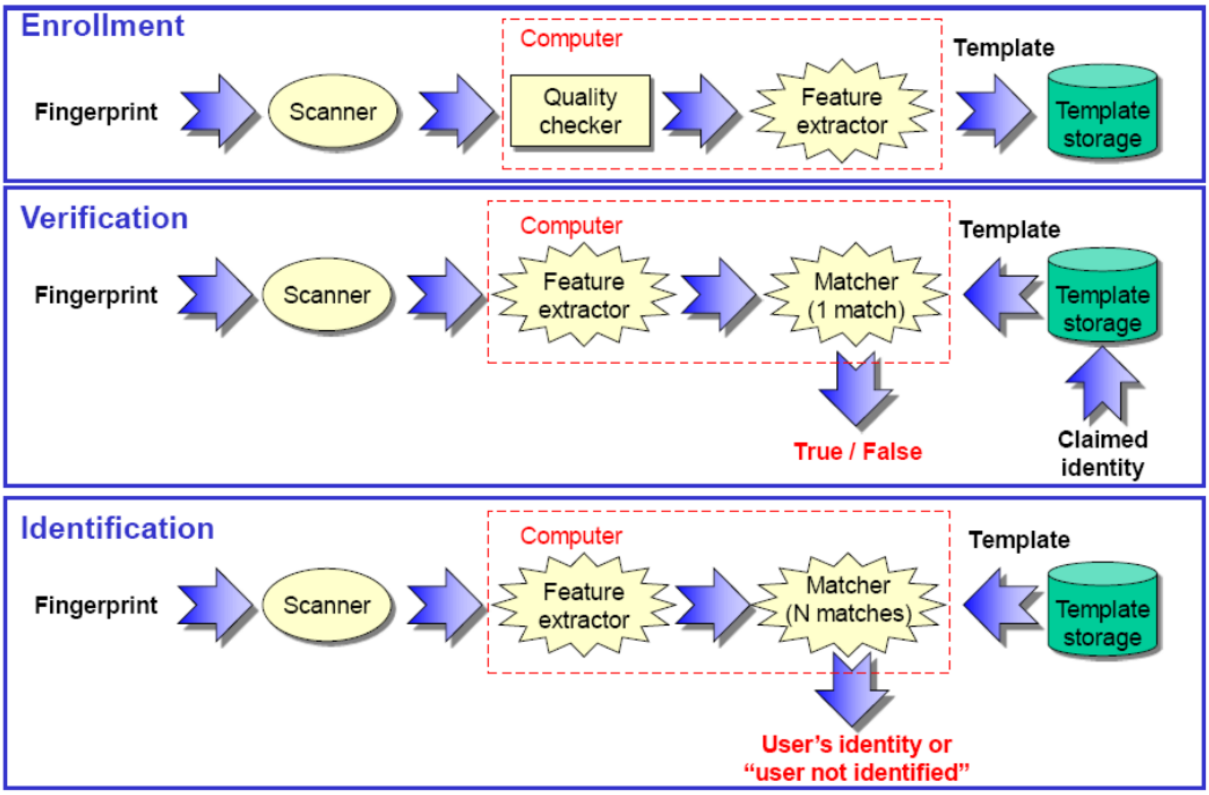


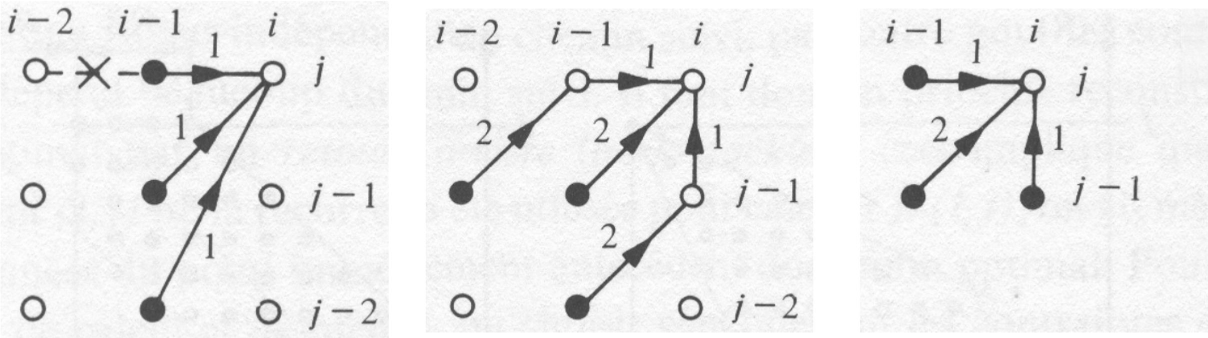
Figure 2: Enrolment, Verification and Identification

Analysis, Modeling and Interpretation of Biometric Data

03. Mathematical Tools: Dynamic Time Warping (DTW) (02,03)

- In voice: Algorithm for measuring **dissimilarity** (distance) between two temporal sequences, which may vary in speed. Given a test word T and reference words R_1, ... R_N (all represented by sequences of feature vectors), we choose the reference word R_r with smallest distance $D(T, R_r)$.
- In voice: The recognition system is adapted to the single speaker who uttered the reference word. Limited vocabulary, without words too close phonetically. The words to be recognized are pronounced in an environment free of noise. They could be isolated in a perfect manner.
- **Non-linear time alignment**: The sequences are *warped* non-linearly in time dimension to determine a measure of their similarity independent of certain non-linear time variations.
- **Algorithm**. *Conditions*: Boundary; monotonicity; step size.
  - Recursively calculate a minimum accumulated distance for each point (i, j) taking into account some local heuristic constraints and weights.

– Each grid point (i, j) is associatied with a *local distance* $d(i, j)$ and an accumulated distance $D(i, j)$.



(a) Local constraints (A) **(b)** Local constraints (B) **(c)** Local constraints (C)

– Type C constraints: Allow a path of any shape satisfying monotonicity.

– Spectral distance:

esto?

04. Mathematical Tools: Gaussian Mixture Model (GMM) (02,03)

- A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities.
- Under the assumption that any arbitrary probability density function (PDF) can be approximated by a linear combination of uni-modal Gaussian densities, the Gaussian mixture models (GMMs) have been applied to model the distribution of a sequence of D-dimensional feature vectors.
- The sum of *mixture weights* equals 1.
- Model parameters: $\lambda = \{w_i, \mu_i, \Sigma_i\}$ (estimated with Expectation Maximization algorithm, although can also be estimated with Maximum A Posteriori estimation).
  - Expectation step: Compute a posteriori probability for component i
  - Maximization step: Maximize, guaranteeing a monotonic increase in model's likelihood.
- MAP estimation is used in speaker recognition to derive speaker model by adapting from a speaker independent universal background model (UBM), or to adapt a prior, general model.
- Decision is carried out using a likelihood test with H_0 (tested recording and speaker's model are from same source), H_1 (not H_0) and Bayes theorem (σ is the decision threshold):

$$\frac{P(T|\lambda_0)}{P(T|\lambda_1)} > \sigma$$

- Similarity domain normalization: $\log L(X) = \log p(X|S = S_c) - \log \sum_{S \in \text{Cohort}} p(X|S \neq S_c)$
- Normalization by a general/world model: A Gaussian mixture which models the parameter distribution for free-text utterances by many speakers.

05. Mathematical Tools: Hidden Markov Model (HMM) (02,03)

- HMM is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (i.e. hidden) states. Is doubly probabilistic finite-state machine.
- **Ergodicity**: There is transition from any state to any other state.
- **Three Basic HMM Problems**:

1. **Decoding** – Given the observation sequence $X = [x(1), x(2), ..., x(t), ..., x(L)]$ and the word model $W = (A, B)$, how do we choose a state sequence $Q = [q(1), q(2), ..., q(t), ..., q(L+1)]$ that is optimal in some meaningful sense sense (e.g. maximal probability)?
2. **Evaluation** – Given the observation sequence $X = [x(1), x(2), ..., x(t), ..., x(L)]$ and a word model $W = (A, B)$, how do we (efficiently) compute $P(X|W)$ (probability of the observation sequence)?
   - The Baum-Welch algorithm (**forward**): $\alpha_j(t) = \sum_i \alpha_i(t-1) \cdot \alpha_{ij} \cdot B_{ij}(X(t))$
   - The Baum-Welch algorithm (**backward**): $\beta_i(t) = \sum_j B_{ij}(X(t+1)) \cdot \alpha_{ij} \cdot \beta_j(t+1)$
   - **Total probability**: $P(X|W) = \alpha(L, q_F) = \beta(0, q|I)$
3. **Training** – How do we adjust the model parameters $W = (A, B, \pi)$ to maximize $P(X|W)$? Algorithm de Baum-Welch (forward-backward):
   - Calculate *all forward-backward* probabilities for all states q_i
   - Calculate posterior probability of transitions γ_{ij}, from state i to state j, conditioned on the observation sequence and the model.
   - Obtain a new estimate $a_{ij} = \gamma_{ij}(X)/\gamma_i$
   - If the value of the total probability has not improved compared to the previous iteration, the re-estimation has converged.

- Continuous Density HMM (CD-HMM) – Parametric approach: Continues probability density functions (ergodic and left-right models).
- *Viterbi Algorithm*: By induction, find the path that leads to a *Max. Likelihood* considering the best likelihood at the previous step and the transitions from it.

06. Mathematical Tools: Principal Component Analysis (PCA)

07. Mathematical Tools: Linear Discriminant Analysis (LDA)

08. Enrollment and Template Creation

09. Verification and Identification System Errors

10. Evaluation of Biometric Systems (01,03)

Performance evaluation:

- **Evaluation factors**: Speech quality, Speech modality, Speech duration, Speaker population.
- **FMR** and **FNMR**:
  - False Match Rate **FMR** – proportion of impostor attempt samples falsely declared to match the compared nonself template (number of impostor FMs / number of impostor attempts)
  - False Non-Match Rate **FNMR** – proportion of genuine attempt samples falsely declared not to match the template of the same characteristic from the same user submitting the sample (number of genuine FNMs / number of genuine attempts)
  - **Calculation**: Create biometric templates using training data set. Define a test set with genuine and impostor trials. Run test and group genuine and impostor scores. Choose threshold value T and calculate FMR(T) and FNMR(T).
- **FAR** and **FRR**:
  - False Accept Rate **FAR** – Proportion of imposters accepted (security breaches)
    FAR=FMRx(1-FTA)
  - False Reject Rate **FRR** – Proportion of genuine users rejected (inconvenience)
    FRR=FTA+FNMRx(1-FTA)
- **DET** and **ROC** curves:

- Detection Error Tradeoff **DET** – can be computed from distributions of scores with a variable threshold. FAR vs FRR
- Receiver Operating Characteristic **ROC** – Correct Accept Rate as function of False Accept Rate (FAR)

- **CAR** and **CRR**:
  - Convenience – Correct Accept Rate **CAR**=1-FRR
  - Security – Correct Reject Rate **CRR**=1-FAR
- **FTA** and **FTE**:
  - Failure to Acquire Rate **FTA** – Proportion that cannot be verified (does not process a certain biometric)
  - Failure to Enroll Rate **FTE** – Proportion that cannot be enrolled (system fails to complete the enrolment process, due to bad quality)
- **Performance measures for identification**:
  - Correct Identification Rate (**CIR**) – proportion of identification transactions by users in the system s.t. the user's identifier is among the ones returned.
  - Cumulative Match Characteristic (**CMC**) – Identification rate as function of K.

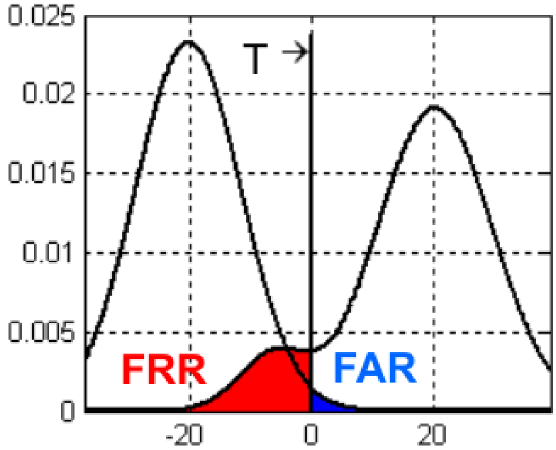


Figure 4: Threshold T, FRR and FAR

Miscellaneous

11. Forensic Automatic Speaker Recognition (03)

- **Forensic Biometrics** – Challenge is to automate forensic biometric methods. Applications of biometric principles and methods to the investigation of criminal activities: to demonstrate the existence of a crime and determine the author. *Forensic* means the use of science or technology in the investigation and establishment of facts or evidence in the court of law.
- **Existing systems and databases**:
  - Automatic Fingerprint Identification System (AFIS) and fingerprints databases
  - DNA sequencers and DNA databases
  - Challenge: Automatic Biometric Identification System (ABIS) and databases for voice, face...
- **Speaker Identification Integrated Project (SIIP)** – Aims to develop technology to rapidly identify suspects' voices and isolate conversations of interest in a wide range of cases (kidnapping, ransom or terrorist calls...)

- **Forensic Speaker Recognition (FSR)**:
  - **Aural-perceptual methods**: earwitnesses, line-ups
  - **Visual methods and *voiceprint?***: visual comparison of spectrograms of linguistically identical utterances (utterly misleading!)
  - **Aural-instrumental methods**: analytical acoustic approach combined with an auditory phonetic analysis
  - **Automatic methods**:
    - *Speaker verification* – not adequate
    - *Speaker identification* – not adequate
    - *Voice as biometric evidence* (How to measure biometric evidence?)
- **Automatic Speaker Recognition (ASR)**:
  - FASR $\neq$ Speaker Verification
  - H_0 (H_1) – speaker's model λ_0 and the tested recording T have same (different) source.

$$\frac{P(H_0)}{P(H_1)} \cdot \frac{P(T|H_0)}{P(T|H_1)} = \frac{P(H_0|T)}{P(H_1|T)} \qquad \frac{P(T|\lambda_0)}{P(T|\lambda_1)} > \sigma \qquad (\sigma \text{ – Decision threshold})$$

- See Table 1.
- **Forensic Automatic Speaker Recognition (FASR)**:
  - **Forensic specifity** – Role of *forensic science* is to testify to the worth of the evidence *quantitatively* (if possible). Forensic science provides *opinion* to help investigators and courts of law answer important questions. Up to the *judge/jury* to use this information in deliberations and decision.
  - **Evaluative forensic science opinion** – opinion of evidential weight, based upon case specific propositions (hypotheses) and clear conditioning information (framework of circumstances) that is provided for use as evidence in court. Is based upon the estimation of a likelihood ratio (in relation with Bayesian interpretation of evidence).
- **Bayesian Interpretation of Biometric Evidence** ($H_0 \equiv$ suspected speaker is source of the recording). Via Bayes Rule, we use the data to update prior beliefs about unknowns. See Figure 5. Freedom of: choosing evidence evaluation and its value; formulating propositions; choosing automatic speaker recognition method.

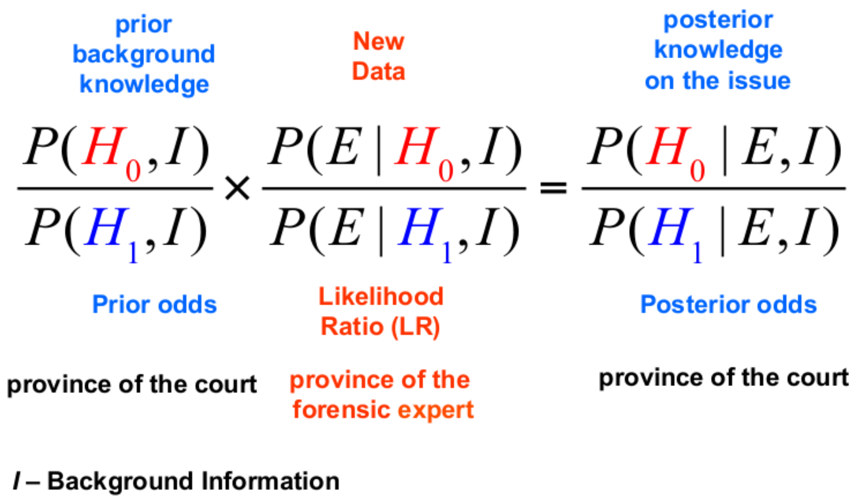


Figure 5: Odds form of Bayes' Theorem: Bayesian Interpretation of Forensic Evidence

- **Measures**:

1. **Biometric Evidence** – Quantified degree of *similarity* between the speaker dependent features extracted from the trace and the extracted from recorded speech of a suspect (model).
   - **FASR – Univariate (Scoring) Method** – See Figure 1. The *score* is used together with the distributions of *between-sources variability* and the *within-source variability* to reach a decision.
2. **Strength of Evidence – Likelihood Ratio** – A likelihood ratio $LR = P(E|H_0)/P(E|H_1)$ of 9.16 means that it is 9.16 times more likely to observe the score (E) given H_0 than given H_1. If $LR > 1$ then H_0, else H_1.
   - **FASR – Multivariate (direct) Method** – E is the multivariate feature representation of trace evidence.
3. **Evaluation of the Strength of Evidence** – (similar idea to Figure 4) *Principle*: Estimation and comparison of likelihood ratios that can be obtained from same speaker and different speaker trials. H_0 true →suspected person recording and questioned recording are from same speaker.
   - Tippett plots I to obtain *Probability of Misleading Evidence (accuracy)* PMEH_0 and PMEH_1.
   - Tippett plots II to obtain EPP (Equal Proportion Probability), PD (Probabilistic Distance) of case LR_{case} to PMEH_0.
   - Empirical Cross-Entropy (ECE) and Log-Likelihood Cost (CLLR)

     these two?

12. Forensic Biometrics (Fingerprints, Face, DNA, Ear, Gait)

13. Biometric Standards

14. Securing Biometric Data and Biometric Encryption

15. Biometrics in Identity Documents

16. Privacy and Legal Issues