

Résumé Détaillé : Systèmes Biométriques

Préparation à l'Examen

1 Concepts Fondamentaux et Métriques

1.1 Définitions et Modes de Fonctionnement

La biométrie utilise des caractéristiques biologiques (physiques) ou comportementales pour reconnaître des individus.

- **Vérification (1 :1)** - "Est-ce que je suis X ?" : L'utilisateur revendique une identité (ex : badge, login). Le système compare l'échantillon capturé à **un seul** modèle de référence. La réponse est binaire : Oui ou Non.
- **Identification (1 :N)** - "Qui suis-je ?" : L'utilisateur ne donne pas d'identité. Le système compare l'échantillon à **toute** la base de données (N personnes) pour trouver une correspondance. C'est plus coûteux en calcul.

1.2 Métriques de Performance (Critères d'évaluation)

L'équilibre d'un système se joue entre sécurité et confort :

1. **FAR (False Acceptance Rate)** - Taux de Fausse Acceptation :
 - Le système accepte un imposteur.
 - **Impact** : Risque de sécurité critique.
 2. **FRR (False Rejection Rate)** - Taux de Faux Rejet :
 - Le système rejette une personne légitime.
 - **Impact** : Problème de confort et de frustration utilisateur.
 3. **EER (Equal Error Rate)** :
 - Le point d'intersection où $FAR = FRR$.
 - C'est le meilleur indicateur global : **plus l'EER est bas, meilleur est le système**.
-

2 Reconnaissance Faciale

2.1 Processus Général

1. **Détection** : Localiser le visage dans l'image (Bounding box).
2. **Alignement** : Normaliser la pose (yeux à l'horizontale, redimensionnement).
3. **Extraction** : Transformer les pixels en un vecteur de caractéristiques.
4. **Matching** : Calculer la distance entre les vecteurs.

2.2 Méthodes Classiques (Avant Deep Learning)

Ces méthodes sont rapides mais **très sensibles** aux variations (lumière, pose, expression).

- **PCA (Eigenfaces)** - Analyse en Composantes Principales :
 - *Principe* : Réduit la dimension de l'image en ne gardant que les variations globales principales (les "Eigenfaces").

- *Limite* : Méthode non supervisée (ne connaît pas les classes/identités), mélange l'information d'identité avec l'éclairage.
- **LDA (Fisherfaces) - Analyse Discriminante Linéaire :**
 - *Principe* : Méthode supervisée. Cherche à **maximiser** la séparation entre les classes (personnes différentes) et **minimiser** la variation intra-classe (même personne).
 - *Avantage* : Meilleure que PCA pour la discrimination.
- **LBP (Local Binary Patterns) :**
 - *Principe* : Analyse la **texture locale** (le grain de peau) en comparant chaque pixel à ses voisins.
 - *Avantage* : Plus robuste aux changements de lumière, mais perd la structure globale du visage.

2.3 Méthodes Modernes (Deep Learning)

Les réseaux de neurones (CNN) apprennent une hiérarchie de caractéristiques (des bords simples jusqu'à l'identité).

- **DeepFace (Facebook)** : Utilise un alignement 3D complexe pour remettre le visage de face avant l'analyse. Performance proche de l'humain.
 - **FaceNet (Google) & Embeddings :**
 - *Concept clé* : L'**Embedding**. C'est un vecteur numérique compact (ex : 128 chiffres) qui représente le visage.
 - *Triplet Loss* : Méthode d'entraînement qui prend 3 images (Ancre, Positif, Négatif) et force le réseau à rapprocher l'Ancre du Positif et à éloigner le Négatif dans l'espace vectoriel.
-

3 Reconnaissance Vocale (Locuteur)

3.1 Concepts : La Physiologie de la Voix

Le modèle "Source-Filtre" distingue deux éléments :

1. **La Source (Cordes vocales)** : Produit le son brut et la fréquence fondamentale (F0 / Pitch). C'est une caractéristique **comportementale** (change avec l'émotion, la fatigue). Peu fiable pour la biométrie.
2. **Le Filtre (Conduit vocal)** : Anatomie de la gorge, bouche, nez. Il sculpte le son via des résonances (Formants). C'est la signature **physiologique et stable** (le timbre) recherchée en biométrie.

3.2 Modes d'Utilisation

- **Text-Dependent (Dépendant du texte)** : L'utilisateur doit prononcer une phrase précise (ex : mot de passe vocal). Très précis, mais contraignant.
- **Text-Independent (Indépendant du texte)** : Le système reconnaît le locuteur peu importe ce qu'il dit (conversation libre). Plus flexible, mais techniquement plus difficile.

3.3 Évolution Technologique

- **MFCC (Mel Frequency Cepstral Coefficients)** : La caractéristique "reine". Elle imite l'oreille humaine (échelle Mel logarithmique) pour extraire le contenu spectral pertinent.
- **GMM-UBM (Approche Classique)** :
 - **UBM (Universal Background Model)** : Un modèle géant de "voix moyenne" appris sur des milliers de personnes.

- **Adaptation** : On adapte l'UBM aux données spécifiques d'un locuteur pour créer son modèle GMM. On compare ensuite via le ratio de vraisemblance (LLR).
 - **i-vectors** : Modélisation factorielle qui sépare l'information du locuteur de l'information du canal (micro, bruit).
 - **Deep Learning (x-vectors)** : Utilisation de réseaux de neurones temporels (TDNN) pour extraire des embeddings robustes, insensibles à la durée de l'enregistrement.
-

4 Reconnaissance d'Empreintes Digitales

4.1 Structure

L'empreinte est formée de **crêtes** (lignes noires, ridges) et de **vallées** (espaces blancs, valleys).

- **Motifs globaux** : Boucle (Loop), Tourbillon (Whorl), Arc (Arch).
- **Minuties (Détails locaux)** : C'est ce qui est utilisé pour le matching.
 - **Terminaison** : Une crête qui s'arrête.
 - **Bifurcation** : Une crête qui se divise en deux.

4.2 Chaîne de Traitement (Pré-traitement Critique)

La qualité de l'image brute est souvent mauvaise. Le prétraitement est obligatoire pour éviter les fausses minuties.

1. **Amélioration (Enhancement)** : Utilisation de filtres de **Gabor** (orientés et fréquentiels) pour reconnecter les crêtes brisées et nettoyer le bruit.
2. **Binarislation** : Transformation de l'image grise en noir et blanc (seuillage, méthode d'Otsu).
3. **Amincissement (Thinning)** : Réduction de l'épaisseur des crêtes à 1 pixel pour obtenir un "squelette".
4. **Extraction des Minuties** : Repérage des croisements (Crossing Number) sur le squelette.

4.3 Matching

- **Basé sur les minuties** : On compare la position (x, y) , l'angle (θ) et le type de chaque minutie entre deux empreintes.
- **Deep Learning (CNN)** : Apprend directement des caractéristiques visuelles complexes, plus robuste aux doigts sales ou abîmés.

5 Biométrie Multimodale et Fusion

5.1 Pourquoi la multimodalité ?

Les systèmes unimodaux (une seule biométrie) ont des défauts majeurs :

- **Bruit** : Capteur sale, voix enrhumée, mauvaise lumière.
- **Non-universalité** : Certaines personnes n'ont pas d'empreintes lisibles (travailleurs manuels).
- **Spoofing** : Facilité de tromper un seul capteur (fausse photo, enregistrement vocal).
- **Précision** : La combinaison de plusieurs sources réduit drastiquement le taux d'erreur.

5.2 Types de Multimodalité

- **Multi-capteurs** : Même trait, capteurs différents (ex : Caméra 2D + Caméra Infrarouge).
- **Multi-algorithmes** : Même trait, traitements différents (ex : Empreinte traitée par Texture + Minuties).
- **Multi-instances** : Plusieurs unités du même trait (ex : Index gauche + Index droit, ou plusieurs photos du visage).
- **Multi-modal (Sens strict)** : Combinaison de traits différents (ex : Visage + Iris + Voix). C'est le plus robuste.

5.3 Architectures de Fusion

- **Série (Cascade)** : On teste d'abord une biométrie rapide/faible. Si le score est incertain, on déclenche la deuxième biométrie. (Bon pour le confort).
- **Parallèle** : Toutes les biométries sont capturées en même temps. (Bon pour la sécurité maximale).

5.4 Niveaux de Fusion (Où mélange-t-on l'info ?)

1. **Niveau Capteur (Sensor Level)** :
 - Fusion des données brutes (ex : fusionner pixels IR et RGB).
 - *Condition* : Nécessite des données compatibles/homogènes.
2. **Niveau Caractéristiques (Feature Level)** :
 - Concaténation des vecteurs de caractéristiques (ex : vecteur visage + vecteur voix).
 - *Problème* : "Malédiction de la dimensionnalité" (vecteurs trop grands, difficile à calculer).
3. **Niveau Score (Score Level) - LE PLUS COURANT** :
 - Chaque sous-système donne un score de similarité. On fusionne ces scores.
 - *Nécessité de Normalisation* : Comme les scores ont des échelles différentes (ex : 0-1 vs 0-1000), il faut normaliser (Min-Max, Z-score, Tanh).
 - *Règles de fusion* : Somme (très robuste), Produit (punitif si un score est nul), Max, Min.
4. **Niveau Décision (Decision Level)** :
 - Chaque système vote (Oui/Non).
 - *Règles* : Vote majoritaire, Règle ET (stricte), Règle OU (laxiste).
 - *Défaut* : Perte d'information (on ne sait pas si c'était un "Oui" limite ou un "Oui" franc).

5.5 Fusion Avancée : Filtre de Kalman

Le Filtre de Kalman permet une **pondération dynamique**. Contrairement à une moyenne simple, il estime la fiabilité (l'incertitude) de chaque capteur en temps réel.

- *Exemple* : Si la caméra détecte qu'il fait noir (bruit élevé sur l'image), le filtre de Kalman va automatiquement réduire le poids du visage et augmenter le poids de la voix pour la décision finale.