Soft Biometrics (for Human Identification): Recognizing People from Human Description

Soft Biometrics

1. **Bertillonage 1890** - developed by Alphonse Bertillon in the late 19th century (around 1890), was a system of anthropometry used for the identification of individuals. It involved measuring various physical characteristics of a person's body, face, iris, ear, nose, and other features, such as the size and shape of the head, ears, nose, and limbs. Bertillonage was an early form of biometric identification before more advanced technologies like fingerprinting became widespread.
2. **Nandakumar and Jain 2004**
    - refers to a research paper or study authored by individuals named Nandakumar and Jain, published in 2004.
    - augmenting traditional biometrics

Types of biometric data

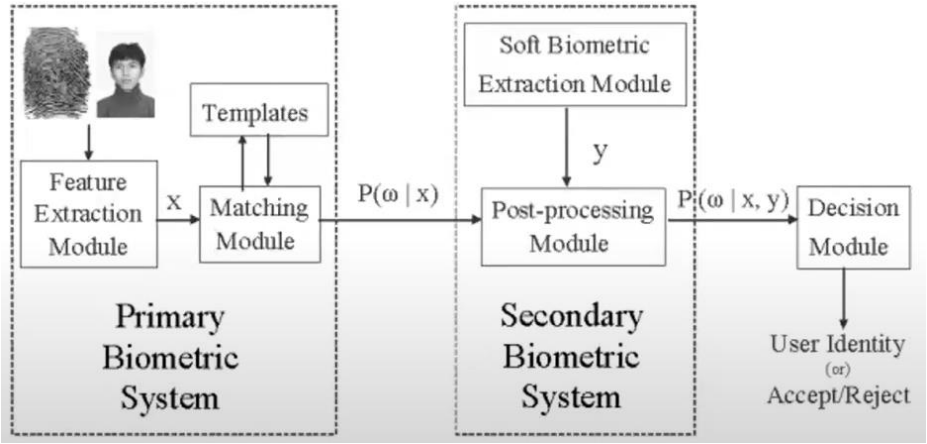
- Face Soft
- Body Soft
- Other Soft (such as tattoos, clothing, makeup, eyes & glasses, hair)

Advantages of Biometric

1. Human understandable description
    - rich in semantics, e.g., a face image described as a "young Asian male" bridges gap between human and machine descriptions.
2. Robustness to image quality
    - soft biometric attributes and low-quality data subject at a distance from the camera.
3. Privacy
    - lack of distinctiveness implies privacy friendly.
4. Performance improvement
    - use in conjunction with biometric cues such as face, fingerprint, and iris fusion to improve accuracy. ID invariance to viewpoint, illumination.

First mention of Soft Biometrics

The concept of integrating Soft Biometric Traits with a Fingerprint Biometric System marks a significant advancement in biometric identification technology. In this approach, the primary biometric system initially extracts the fingerprint data. Subsequently, the soft biometric extraction module incorporates additional identifying factors into the secondary biometric system. This integration holds promise for enhancing decision-making processes within the system. Notably, studies have shown that the recognition performance of a fingerprint system can improve by approximately 4% with the inclusion of soft biometrics such as ethnicity, gender, and height.

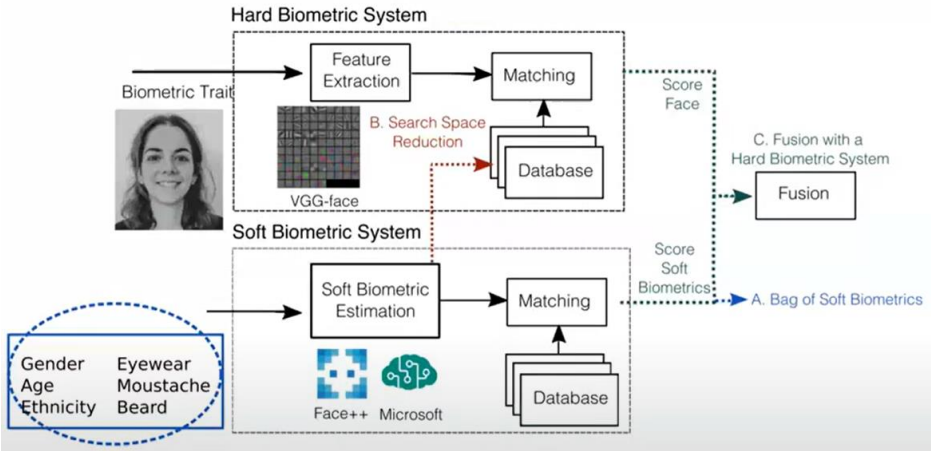


Facial Soft Biometrics for Recognition

a. Bag of Soft Biometrics

b. Search Space Reduction

c. Fusion with a Hard Biometric System



Descriptions and attributes for identification

a. Eyewitness statement
b. Database of images
c. Image of crime
d. Database of descriptions

Exploring Human Descriptions

- We explore semantic descriptions of:

a. physical traits
b. semantic terms
c. visible at a distance

On Semantic Descriptions

- Advantages
    1. No (feature/ sensor) ageing.
    2. Available at a distance/ low resolution/ poor quality.
    3. Fit with human (eyewitness) description/ forensics.
    4. Complement automatically perceived measures.
    5. Need for search mechanisms.
- Disadvantages
    1. Psychology/ perception
    2. Need for labelling.

Traits and Terms

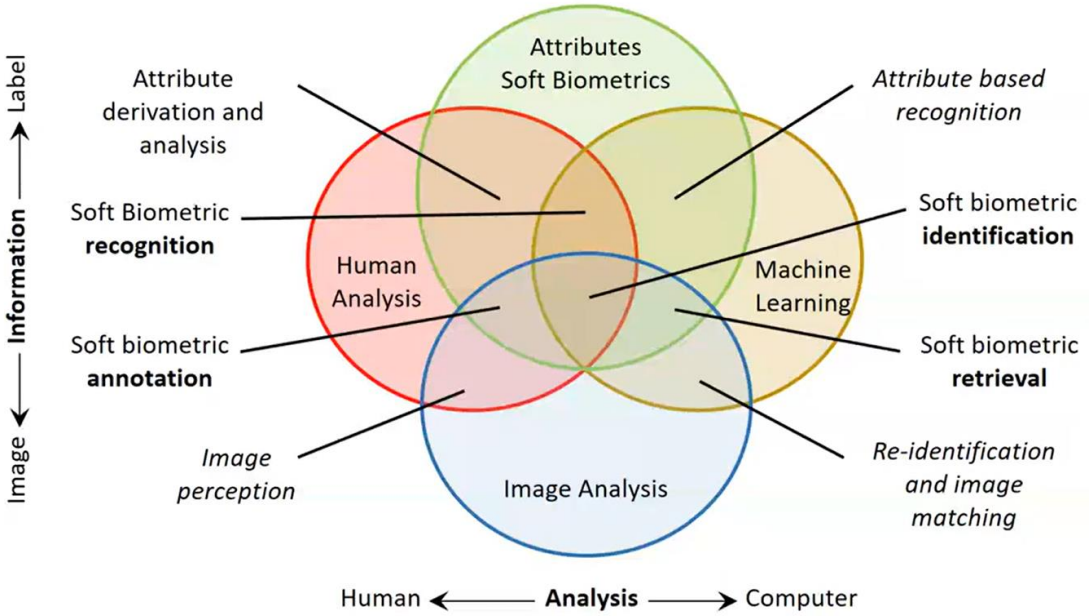
1. Global Features
    - Features mentioned most often in witness statements
    - Sex and age quite simple
    - Ethnicity
        a. Notoriously unstable
        b. There could be anywhere between 3 and 100 ethnic groups.
        c. 3 "main" subgroups plus 2 extra to match UK Police force groupings.
2. Body Features / Body Shape
    - Figure, Weight, Muscle Build, Height, Proportions, Shoulder Shape

Chest Size, Hip size, Leg/Arm Length, Leg/Arm Thickness

- Based on whole body description stability analysis by MacLeod et al.
- Features showing consistency by different viewers looking at the same subjects
- Mostly comprised of 5-point qualitative measures
- e.g. very fat, fat, average, thin, very thin
- Most likely candidate for fusion with gait

3. Head
   - Hair Colour, Hair Length, Facial Hair Colour/Length, Neck Length/Thickness

How does this fit with computer vision?



Labelling via CrowdFlower

Labelling via CrowdFlower offers a professional environment for data annotation. It provides the capability to continuously evaluate labelers, ensuring consistent quality throughout the labelling process. CrowdFlower also facilitates access to a diverse pool of labelers, representing a wide population for more comprehensive data analysis. Moreover, it offers cost-effective solutions compared to alternative platforms. Notably, CrowdFlower remains accessible in regions where other platforms, such as Amazon Mechanical Turk, are not available, making it a viable choice for users in the UK.

Addressing issues with absolute descriptors requires innovative approaches:

1. **Comparative Human Descriptions:** By comparing attributes between subjects, we can infer continuous relative measurements rather than relying on absolute values. This approach enables a more nuanced understanding of differences and similarities.
2. **Developing Relative Attributes:** Utilizing ranking Support Vector Machines (SVM), we can establish relative attributes, enabling the comparison of attributes across different subjects. This method facilitates the extraction of meaningful insights from data.
3. **Ranking Comparative Descriptions:** Leveraging the ELO rating system, originally designed for chess, we can rank comparative descriptions. This system transforms comparative labels into a ranked list, providing a structured framework for interpreting and analyzing relative attributes effectively.

Components

- Data
- Labels (categorical or comparative)
- Ranking algorithm (for comparative labels
- Feature selection
- Computer vision (feature extraction, color mapping)
- Classifier (e.g. kNN, SVM, DBN)

Recognition by crowdsourced body labels

1. **Lower Recognition Accuracy with Absolute Labels:** Using absolute labels tends to result in lower recognition accuracy. Absolute descriptors may not capture the nuanced variations in body attributes effectively, leading to reduced performance in recognition tasks.
2. **Increased Accuracy with More Labels and Comparisons:** However, increasing the number of labels and comparisons can enhance accuracy. By incorporating a larger and more diverse set of labels, along with comparative descriptions, the system gains a richer understanding of body attributes, thereby improving recognition performance.

Superfine Labels

- Most ‘fine’ are actually coarse.
- Our comparative attributes are superfine.
- Comparison / ranking gives many advantages.

Recognition by clothing

- Clothing generally unique
- Short term biometric
- Has strong invariance
- Links with computer vision and automatic clothing analysis / reidentification
- Clothing has the ability to handle 90-degree change.

Conclusion

In conclusion, soft biometrics emerge as fundamental metrics for identification, offering a diverse range of capabilities and applications. They extend beyond mere performance enhancement, enabling the exploration of new application scenarios that were previously unattainable. Particularly suited to surveillance contexts characterized by poor lighting, distance, or low resolution, soft biometrics exhibit distinct advantages. However, their potential remains largely untapped, necessitating further investigation into various aspects such as covariates and anti-spoofing measures to fully realize their performance benefits. Moreover, the integration of soft biometrics motivates the exploration of new insights into automated identification systems compared to traditional human identification methods. As research in this field continues to evolve, soft biometrics holds promise for revolutionizing identification processes and enhancing security in various domains.

Situation in which Soft Biometric can be Applied in this World

One situation where soft biometrics can be applied is in airport security screening processes. In this scenario, soft biometrics can be used to enhance the identification of individuals in crowded and busy environments, where traditional biometric methods like fingerprint or iris scanning may be impractical due to logistical constraints.

Attributes related to this situation could include:

1. **Facial Recognition**: The shape of the face, eyes, nose, and mouth.
2. **Gait Recognition**: The unique way individuals walk (gait).
3. **Clothing and Accessories**: Clothing patterns, accessories, and luggage characteristics.
4. **Behavioral Biometrics**: Body language, gestures, and stress levels detected through physiological signals.
5. **Voice Recognition**