Soft Biometrics

Ruggero Donida Labati, Arun Ross, Antitza Dantcheva

Soft Biometrics

Ruggero Donida Labati*, Arun Ross, Antitza Dantcheva

Definition

Soft biometrics traits are physical, behavioral, or material accessories associated with an individual, which can be used for describing an individual. These attributes are typically gleaned from primary biometric data, are classifiable in pre-defined human understandable categories, and can be extracted in an automated manner. Examples include age, gender, hair color, eye color, and cosmetics.

Background

Soft biometrics consists of ancillary information extracted from primary biometric samples (such as face, fingerprints, hand, and iris). Examples of soft biometrics are age, gender, ethnicity, height, hair color, and eye color. As described in Dantcheva et al (2011), soft biometrics are typically fused with additional characteristics to increase the robustness of recognition systems in challenging conditions, or they are used to realize a list of quantitative descriptors of an individual (e.g., in forensic, medical, and ambient intelligence applications).

The first study on the discriminative capabilities of soft biometric attributes has been performed by Alphonse Bertillon (Rhodes (1956)) in the nineteenth century, who introduced the idea of a personal identification approach based on biometric, morphological and anthropometric features.

The term "soft biometrics" has been introduced in Jain et al (2004) to describe a set of characteristics that provide some information about the individual, but are not unique to that person, mainly due to lack of distinctiveness and permanence. The definition of soft biometric traits has evolved during time, by associating soft

* corresponding author

biometrics with “labels” which people use to describe each other.

Theory and Application

Taxonomy of soft biometrics

Studies in the literature analyzed a huge set of soft-biometric attributes. As an example, the literature survey presented in Hassan et al (2021) describes more than 170 characteristics extracted from biometric images and videos. The work presented in Nixon et al (2015) additionally considers attributes extracted from voice signals. Soft biometric attributes can be extracted manually or by using automated approaches. Soft biometric attributes can be expressed using categorical data representing a class (e.g., male and female), continuous numerical values (e.g., height expressed in centimeters), or can consist of comparative attributes (e.g., *A* is taller than *B*). However, only a limited number of studies have considered the use of comparative attributes in automatic systems.

The taxonomy presented in Dantcheva et al (2016) divides soft biometrics into four classes: demographic, anthropometric, medical, and material and behavioral attributes. Table 1 describes the types of data from which it is possible to extract different classes of attributes and provides examples of frequently used soft biometric attributes pertaining to each class.

Demographic attributes: Many studies in the literature regard the estimation of gender, age, and ethnicity. Most of these studies are based on face images and achieve better or comparable results with respect to human observers. A number of studies regard methods for gender estimation from body and speech characteristics. Furthermore, there are promising studies on gender estimation from fingerprint, iris, and hand. In particular, there are studies on age estimation from body and hand, which are especially useful for forensic anthropologists. Other studies on the extraction of demographic attributes mainly regard descriptors such as hair color, eye color, or skin tone, and are usually based on face images.

Anthropometric attributes: These attributes mainly consist of measurements of the face and body geometry. The studies on the analysis of face geometry are based on the automatic extraction of facial landmarks. Methods for measuring body geometry frequently use three-dimensional acquisition systems or calibrated acquisition setups to achieve accurate measurements of height and single body components (e.g., arms, legs, and head).

Medical attributes: Image-based automated self diagnostic methods can produce a wide variety of data from which it is possible to extract soft biometric attributes. As an example, there are methods for estimating the body fat index from face images and methods for estimating the weight of a person automated from body and face images. Skin lesions, moles, and wrinkles can also be used as distinctive patterns of an individual.

Material and behavioral attributes: Although material attributes can have limited time permanence, they can be used as distinctive patterns, for example, in surveillance applications. Examples of material attributes that can be inferred

are the colors of the clothes, type of shoes, eye glasses, hats, scarfs, and masks. Other attributes include tattoos, especially in forensic applications. Additionally, material attributes can be useful to tune primary biometric recognition systems. As an example, methods to estimate the presence of lenses or eye glasses can be used to improve the robustness of iris recognition systems. Cosmetic makeup can also be considered as a soft biometric attribute since it can modify the appearance of the face, thereby negatively affecting face recognition systems. Behavioral attributes can be estimated from frame sequences of walking individuals used for gait recognition (e.g., to estimate health cues) or from speech samples (e.g., to estimate the accent of the speaker).

Applications

Soft biometrics have a great number of possible applications such as: image tagging and video indexing, human-computer interaction, forensics, surveillance, age specific control, electronic customer relationship management (E-CRM), cosmetology, video retrieval, and health monitoring. This wide range of applications is due to some relevant benefits (Dantcheva et al (2016)), which are summarized below.
Human understandable interpretation: Soft biometrics have a semantic interpretation. They can easily be interpreted by humans and provide intuitive descriptions of the individual (e.g. tall, broad, and long hair). Therefore, soft biometrics can assist in application contexts involving human observers, as in surveillance, forensic, and medical scenarios.

Robustness to low data quality: Frequently, soft biometrics can be extracted from samples with quality not sufficient as far as the primary biometric trait is concerned.

Computational efficiency: The computation of soft biometric attributes is frequently more efficient with respect to the feature extraction step of primary traits.

Consent-free acquisition: Many soft biometric attributes can be computed from samples captured without subject consent. As an example, the height of a person can be inferred from surveillance footages acquired at long distances.

Privacy: Usually, a single soft biometric attribute is not sufficiently distinctive for recognizing an individual. Therefore, soft biometrics can be considered as less privacy invasive with respect to primary biometrics.
According to the use of soft biometric attributes, it is possible to distinguish three main application contexts.

1) Uni-modal system: Many applications require the estimation of a single soft biometric attribute (e.g., parental control systems only need to estimate the age).

2) Identity recognition: Soft biometrics can be fused with other soft biometrics and / or primary biometrics to perform an identity verification or identification.

3) Search space reduction: Soft biometrics can be used to speed up identification for datasets composed of large numbers of individuals. As an example, age information can be used to reduce the search space of a biometric system.

Table 1 Main data from which it is possible to extract widely uesd soft biometric attributes.

Kind of data	Class of attributes	Examples of soft biometric attributes
Face	Demographic	Gender, age, ethnicity, hair/eye/skin color
	Anthropomorpic	Facial geometry
	Medical	Wrinkless, skin lesions, body mass index
	Material and behavioral	Glasses, hat, mask, scarf, accessories, makeup
Body	Demographic	Gender, age, ethnicity, hair/skin color
	Anthropomorpic	Body geometry
	Medical	Wrinkless, skin lesions, moles, body mass index, weight
	Material and behavioral	Clothes, shoes, accessories, tattoos
Iris	Demographic	Gender, age, ethnicity, eye color
	Material and behavioral	Glasses, lenses
Fingerprint	Demographic	Gender
Gait	Demographic	Gender, age
Hand	Demographic	Gender
Voice	Demographic	Gender, age, ethnicity

Open problems and Future directions

The research community is investigating different problems regarding soft biometrics.

A relevant problem consists of selecting the best set of soft biometric attributes for the considered application. In many cases, the division into categories of some features can be complex and not clear (e.g., the gender). Furthermore, the class labels provided by different human experts for the same sample can exhibit significant differences. There could also be limitations in terms of computational time and resources.

To increase the robustness and accuracy of current technology, the research community is investigating methods specifically designed to fuse soft biometric attributes with additional soft biometric information and primary biometric traits. Researchers are also studying quality assessment and enhancement methods for samples used to extract soft biometrics, with the aim of improving the robustness of current systems in challenging conditions. Additionally, there are some studies on the statistical modeling and mathematical analysis of systems based on soft biometrics, which seek to estimate theoretical bounds on the accuracy of these traits. Furthermore, the community is working on the standardization of some soft biometric attributes and their representation, with the goal of improving the interoperability between different applications and systems.

Another problem under investigation is the demographic bias. The study presented in Drozdowski et al (2020) proposes a detailed analysis of both technical and social matters.

A further relevant problem consists of the ethical implications of the technology. In many applications, soft biometric attributes can be extracted covertly from uncooperative users. As an example, in-

formation like ethnicity, age, and gender of an individual can be extracted from surveillance footages. Therefore, misuse of soft biometrics can infringe the privacy of the individuals.

The research community is also working on improving the security of systems based on soft biometrics. In this context, attacks to the sensor are a particularly relevant problem. The techniques used to perform presentation attacks are constantly evolving. Therefore, approaches based on soft biometrics that are robust to novel types of attacks should be studied.

Privacy protection is another relevant research topic. As an example, face images can reveal additional information such as ethnicity, gender, and age. As described in Othman and Ross (2015), there are studies to address this problem by utilizing deidentification methods that remove unnecessary information from the samples. As described in Mirjalili et al (2020) and Naresh Boddeti (2018), techniques based on semi-adversarial networks and homomorphic encryption are being developed to limit information leakage from biometric samples.

Summary

Soft biometrics are attributes that can be extracted from different kinds of biometric signals, images, and videos. These attributes are not unique to an individual, but can be used to augment the performance of a primary biometric system. Soft biometrics can be used in a wide set of application scenarios and are applied in three main contexts: uni-modal systems, identity recognition systems, and database filtering systems. However, the research community is still working on several open problems and on improving the current technology.

References

Dantcheva A, Velardo C, D'Angelo A, Dugelay JL (2011) Bag of soft biometrics for person identification: New trends and challenges. Multimedia Tools and Applications 51:739–777

Dantcheva A, Elia P, Ross A (2016) What else does your biometric data reveal? A survey on soft biometrics. IEEE Trans on Information Forensics and Security 11(3):441–467

Drozdowski P, Rathgeb C, Dantcheva A, Damer N, Busch C (2020) Demographic bias in biometrics: A survey on an emerging challenge. IEEE Trans on Technology and Society 1(2):89–103

Hassan B, Izquierdo E, Piatrik T (2021) Soft biometrics: a survey. Multimedia Tools and Applications

Jain AK, Dass SC, Nandakumar K (2004) Soft biometric traits for personal recognition systems. In: Zhang D, Jain AK (eds) Biometric Authentication, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 731–738

Mirjalili V, Raschka S, Ross A (2020) Privacynet: Semi-adversarial networks for multi-attribute face privacy. IEEE Trans on Image Processing 29:9400–9412

Naresh Boddeti V (2018) Secure face matching using fully homomorphic encryption. In: Proc. of the 9th IEEE Int. Conf. on Biometrics Theory, Applications and Systems (BTAS), pp 1–10

Nixon MS, Correia PL, Nasrollahi K, Moeslund TB, Hadid A, Tistarelli M (2015) On soft biometrics. Pattern Recognition Letters 68:218–230

Othman A, Ross A (2015) Privacy of facial soft biometrics: Suppressing gender but retaining identity. In: Agapito L, Bronstein MM, Rother C (eds) Computer Vision - ECCV 2014 Workshops, Springer International Publishing, Cham, pp 682–696

Rhodes H (1956) Alphonse Bertillon, Father of Scientific Detection. Abelard-Schuman