What Else Does Your Biometric Data Reveal? A Survey on Soft Biometrics

Antitza Dantcheva, Petros Elia, Arun Ross

Abstract—Recent research has explored the possibility of extracting ancillary information from primary biometric traits, viz., face, fingerprints, hand geometry and iris. This ancillary information includes personal attributes such as gender, age, ethnicity, hair color, height, weight, etc. Such attributes are known as soft biometrics and have applications in surveillance and indexing biometric databases. These attributes can be used in a fusion framework to improve the matching accuracy of a primary biometric system (e.g., fusing face with gender information), or can be used to generate qualitative descriptions of an individual (e.g., “young Asian female with dark eyes and brown hair”). The latter is particularly useful in bridging the semantic gap between human and machine descriptions of biometric data. In this paper, we provide an overview of soft biometrics and discuss some of the techniques that have been proposed to extract them from image and video data. We also introduce a taxonomy for organizing and classifying soft biometric attributes, and enumerate the strengths and limitations of these attributes in the context of an operational biometric system. Finally, we discuss open research problems in this field. This survey is intended for researchers and practitioners in the field of biometrics.

Index Terms—Soft biometrics, Biometrics, Computer Vision, Gender, Age, Ethnicity, Race, Cosmetics, Privacy, Semantics, Visual Attributes

I. INTRODUCTION

A. Biometrics

Biometrics is the science of recognizing individuals based on their physical, behavioral, and physiological attributes such as fingerprint, face, iris, gait and voice [111]. A classical biometric system acquires biometric data from an individual (e.g., a fingerprint image), extracts a set of features from the data, and compares this feature set with templates in the database in order to verify a claimed identity or to determine an identity.

While biometric data is typically used to recognize individuals, it is possible to deduce other types of attributes of an individual from the same data. For example, attributes such as age, gender, ethnicity, height, hair color and eye color can be deduced from data collected for biometric recognition purposes. Recent work [270] has established the possibility of computing the body mass index (BMI) from face images, thereby suggesting the possibility of assessing health from biometric data.

B. Soft Biometrics

These additionally deduced attributes, while not necessarily unique to an individual, can be used in a variety of applications. Further, they can be used in conjunction with primary biometric traits in order to improve or expedite recognition performance.

Fig. 1. Anthropometry card of Alphonse Bertillon, who originated the criminal identification system based on profile and full-face photos, and key body measurements (1892). These key measurements include body height, body weight, build, complexion, head length, head width, cheek width, measurements of right ear and left foot, as well as “peculiar marks” such as birthmarks, scars, and tattoos.

It is perhaps this latter application that has led to these attributes being referred to as soft biometrics [109], [110], [180] or light biometrics [4]. In this context, soft biometrics can be traced back to Bertillon [211] (see Figure 1), who brought to the fore the idea of using anatomical, morphological and anthropometrical characteristics for person identification. These attributes have also been referred to as semantics [223], [207], in reference to their semantic interpretation (e.g., describing a face as “young male”).

1) Scope and benefits: Various researchers have attempted to define the scope of soft biometrics. Jain et al. [109] defined soft biometrics to be the set of characteristics that provide some information for recognizing individuals, but that are not capable of distinguishing between individuals, mainly due to their lack of distinctiveness and permanence. Samangooei et al. [223], as well as Reid and Nixon [208], further associated soft
biometrics with labels which people use to describe each other: an association that nicely bridges the gap between human and machine descriptions of biometric data.

Combining the above with the ideas in Dantcheva et al. [42], and keeping in mind that such soft traits can go beyond person recognition, one could define soft biometrics as follows. Soft biometric traits are physical, behavioral, or material accessories, which are associated with an individual, and which can be useful for recognizing 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.

Fig. 2. Importance of soft biometrics. Typical video surveillance scenario: when faces are of low resolution, appear in different poses, and are either occluded or not visible, other attributes such as age, gender, hair color and style, height, body type, clothes and accessories can be used for identification and re-identification. Image obtained from PETS 2007 [63].

A) Benefits of soft biometrics: Soft biometrics are often descriptive and have a semantic representation. In addition - as noted by Jain et al. [110] - they can be inexpensive to compute, discerned at a distance in a crowded environment, and require less or no cooperation of the observed subjects. To elaborate, we note the following benefits.

Human understandable interpretation: Soft biometric attributes have a semantic interpretation, in the sense that that they can provide a description that can be readily understood by humans; for example the description “young, tall, female”. This makes them particularly useful in applications such as video surveillance, where they are directly compatible with how humans perceive their surroundings [210], [67], [43], [204], [47], [257], [45], [46]. In other words, when a human attempts to verbally describe a person, obvious characteristics regarding the person’s appearance such as gender, age, height and clothes color are often used (e.g., in police reports). This allows soft biometrics to be used in applications where traditional biometrics may be insufficient, as is argued, for example, by Klontz and Jain [125] in the case of the 2013 Boston bombings.

Robustness to low data quality: Some soft biometric attributes can be deduced from low-quality biometric data (see Figure 2. In this context, such attributes can be extracted, when primary biometric data is not conclusive, due to poor acquisition quality. For example, if the input iris image is of poor quality, one could utilize the surrounding periocular information to perform recognition, rather than relying on the iris itself.

Consent-free acquisition: Soft biometrics can often be captured without the consent and cooperation of the observed subject. For example, information about a person’s height or gender can be deduced from a distance.

Privacy: Since soft biometric traits are not distinctive, they only provide a partial description of a person (such as “female, tall, young”). This limitation has positive privacy ramifications when it comes to extracting and storing such soft biometric data.

2) Taxonomy: With the aforementioned scope and benefits in mind, it is worth identifying a taxonomy that can facilitate organization and categorization of soft-biometric traits. This taxonomy is based on utility, and it considers four groups of attributes: demographic, anthropometric, medical, and material and behavioral attributes. This categorization - and the more refined sub-categorization based on the modalities of face, iris, body, gait, fingerprint and hand (Figure 3) - will also help us structure the exposition of the state-of-art in the rest of this survey paper.

TABLE I
SOFT BIOMETRIC TAXONOMY WITH FOUR GROUPS: I) DEMOGRAPHIC, II) ANTHROPOMETRIC AND GEOMETRIC, III) MEDICAL, IV) MATERIAL AND BEHAVIORAL.

<table>
<thead>
<tr>
<th>Demographic attributes</th>
<th>age, gender, ethnicity, eye-, hair-, skin-color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthropometric and geometric attributes</td>
<td>body geometry, facial geometry</td>
</tr>
<tr>
<td>Medical attributes</td>
<td>health condition, BMI, body weight, wrinkles</td>
</tr>
<tr>
<td>Material and behavioral attributes</td>
<td>Hat, scarf, bag, clothes, lenses, glasses</td>
</tr>
</tbody>
</table>

The above taxonomy need not necessarily result in disjoint groups and it is certainly not a unique taxonomy. For example, taxonomies in Jain et al. [110] and Dantcheva et al. [42] (see also [39]) have different definitions that categorize traits based on their ability to distinguish between individuals as well as
their variability over time.

C. Domains of application

Automated soft biometric extraction has a number of applications: in the area of security where algorithms can locate a person-of-interest based on a specific set of soft biometric attributes; in image-tagging and video indexing where photo or video album management can be performed based, for example, on age, gender, and clothing; in human-computer interaction where data and personalized avatars can be automatically designed according to the user’s external appearance (e.g. hair- and skin-color, age and gender); in forensics where artists can amend sketches of the suspect or the victim based on old pictures; and in surveillance where suspects can be located based on semantic descriptions. Other applications include age-specific access control where, for example, children can be prevented from watching certain movies, accessing certain web sites, or entering bars or liquor-stores. There are industrial systems\(^1\) that extract demographic information of customers for customizing advertisements or for collecting aggregate data about consuming habits (e.g., based on age, gender, ethnicity). In addition, Electronic Customer Relationship Management (ECRM) can use soft biometrics-based categorization for effectively managing customers by offering customized products and services. For example, age or gender specific advertisement can be presented for consumer goods such as mobile phones, fashion, and food. In cosmetology, it is of interest to estimate the rejuvenating effect of decorative cosmetics and cosmetic surgery by computing the perceived age of an individual from their face image.

In video retrieval systems \([93], [252], [258], [190]\), soft biometric traits can be used to locate specific individuals in a video stream either by verbal descriptions (e.g., “individual with a red shirt”) or by automatically extracting soft biometric features from an input image and using these features to locate a matching individual in the video stream.

Finally, in health monitoring, soft biometrics are envisioned to play a major role in early diagnosis of illness, sickness prevention and health maintenance. Such traits include body weight / body mass index, skin abnormalities, and wrinkles. We will expand on this possibly later on in the paper.

Below, we describe the various contexts in which soft biometric traits can be used.

1) Uni-modal system: Often applications might require the extraction of a single soft biometric trait (e.g. gender in a gender-personalized advertising campaign), in a so called uni-modal soft biometric system. Such a system generally contains the “preprocessing”, “feature extraction” and “classification” modules, with the main focus being on the choice of representation (feature extraction).

2) Fusion with primary biometric traits: Here, the goal is to improve the recognition accuracy of a biometric system. Such an approach was proposed by Jain et al. \([110]\), who considered a hybrid system that combined fingerprint identification with soft biometric attributes such as age, gender and height, to improve the overall matching accuracy.

Let \( W = \{ w_1, w_2, ..., w_n \} \) be the set of \( n \) subjects enrolled in the database, and let \( x \) be the feature vector corresponding to the primary biometric system. The output of the primary biometric system is of the form \( P(w_i|x) \), \( i = 1, 2, ..., n \), where \( P(w_i|x) \) is the probability that the input data belongs to subject \( w_i \) given the feature vector \( x \). Let \( y = \{ y_1, y_2, ..., y_m \} \) be the soft biometric feature vector. Then the updated probability \( P(w_i|x,y) \) that the subject in question is \( w_i \), can be calculated using the Bayes rule to be

\[
P(w_i|x,y) = \frac{p(y|w_i)P(w_i|x)}{\sum_{j=1}^{m} p(y|w_j)P(w_j|x)}
\]

where \( p(y|w_i), i = 1, 2, ..., n \) represents the conditional probability of the random vector \( y \) given subject \( w_i \).

Other notable research on fusing soft biometrics and classical biometrics, include the works in \([229], [112], [109], [1], [289], [189]\).

3) Search space reduction: Soft biometrics can also be used to expedite the search in large biometric databases by filtering out subjects. A number of attributes such as age, gender, hair and skin color have been proposed for efficient filtering of face databases \([129], [130], [103]\). Furthermore, an analysis of the filtering-gain versus filtering-reliability tradeoff in using soft biometric traits to prune large databases was presented in \([41]\).

D. Visual attributes

The computer vision community refers to describable visual attributes as any visual and contextual information that is helpful in representing an image (cf. Scheirer et al. \([229]\)). In this approach, semantically meaningful labels are employed towards image retrieval and object categorization. In the context of human recognition, this semantic information can describe gender \([228]\), ethnicity \([229]\), accessories \([21]\), clothing style \([238]\), and facial-feature-shapes \([228]\). Related work include fusion of attributes by Scheirer et al. \([227]\), pruning of large-scale datasets by Russakovsky and Fei-Fei \([220]\), as well as studies on similarities between faces or objects based on relative attributes by Parikh and Grauman \([187]\) and Zhou et al. \([297]\). Other pertinent literature include \([62], [131], [52], [137], [156], [21]\). Of specific interest are the “zero-shot” learning approaches, where previously unseen objects are described using attributes of objects encountered in the training set (cf. Parikh and Grauman \([187]\)).

E. Structure of paper

The survey provides a review of salient techniques for extracting soft biometrics from modalities such as face, body, fingerprint, iris, and voice. While an exhaustive survey of all soft biometric traits is not possible due to the richness of the field (for example, we do not expand on traits relating to the ear, or to saccadic movements), we try to offer a holistic view of most of these traits. In this way, this survey paper is significantly different from other introductory overviews (see \([42], [210], [67], [213], [106] \) and Table II) that have focused on specific soft biometric traits such as gender, age or ethnicity.
The structure of this survey is based on the aforementioned taxonomy of soft biometrics. We discuss soft biometric traits that are heavily used as demographic attributes (Section II), as anthropometric (geometric) attributes (Section III), medical attributes (Section IV), and as miscellaneous material and behavioral attributes (Section V). Finally in Section VI we discuss open research problems that are currently being addressed in the field of soft biometrics.

TABLE II
EXISTING INTRODUCTORY OVERVIEWS ON GENDER AND AGE ESTIMATION TECHNIQUES.

<table>
<thead>
<tr>
<th>Modality</th>
<th>Scientific work</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Ng et al. [177]</td>
<td>2012</td>
</tr>
<tr>
<td>Gender</td>
<td>Khan et al. [120]</td>
<td>2011</td>
</tr>
<tr>
<td>Gender</td>
<td>Beksos-Calfa et al. [18]</td>
<td>2011</td>
</tr>
<tr>
<td>Gender</td>
<td>Ramanathan et al. [201]</td>
<td>2009</td>
</tr>
<tr>
<td>Gender</td>
<td>Mäkinen and Raisamo [157]</td>
<td>2008</td>
</tr>
<tr>
<td>Gender</td>
<td>Mäkinen and Raisamo [158]</td>
<td>2008</td>
</tr>
<tr>
<td>Age</td>
<td>Guo [90]</td>
<td>2012</td>
</tr>
<tr>
<td>Age</td>
<td>Fu et al. [68]</td>
<td>2010</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Fu et al. [65]</td>
<td>2014</td>
</tr>
</tbody>
</table>

II. DEMOGRAPHIC ATTRIBUTES

The term demographics, in addition to referring to the quantifiable statistics of a given population, refers to attributes such as age, gender, ethnicity and race that are widely used in common population statistics. Since the early publications in [164], [243], research on this class of soft biometrics has been embraced by the computer vision community.

A. Gender Estimation

The traditional definition of sex refers to the biological characteristics that differentiate men and women, as opposed to gender, which is related to the social and cultural distinctions between the sexes. However, very often, the terms “sex” and “gender” have been used interchangeably in the biometrics literature. Consequently, we do not make any explicit distinction between the two terms in this article.

Gender estimation remains a challenging task, which is inherently associated with different biometric modalities including fingerprint, face, iris, voice, body shape, gait, signature, DNA, as well as clothing, hair, jewelry and even body temperature (see [165]). The forensic literature [148] suggests that the skull, specifically the chin and the jawbone, as well as the pelvis, are the most significant indicators of the gender of a person; in juveniles, these shape-based features have been recorded to provide classification accuracy of 91% to 99%. It has been argued (see for example the work by Loth and Iscan [148]) that there is no single skeletal feature that definitely reveals the evidence of sexual dimorphism, and that there is in fact a cross-gender metric overlap of up to 85%, which can be attributed to environmental influences and pathologic conditions, such as diet and occupational stress. In spite of this, forensic experts argue [128] that near 100% gender determination accuracy can be attained by visual examination of the entirety of the skeleton.

Humans are generally quite good at gender recognition, as they have been programmed - from an evolutionary standpoint - to classify gender from early on in their lives [185]. As pointed out by Edelman et al. [54], humans perform face image-based gender classification with an error rate of about 11%, which is commensurate to that of a neural network algorithm performing the same task (at that point in time).

Despite this, automated gender recognition from biometric data remains to be a challenge and is impacted by other soft biometrics, for example, age and race; gender dimorphism is accentuated only in adults, and varies across different races.

1) Gender from face: In gender recognition from face, feature-based approaches extract and analyze a specific set of discriminative facial features (patches) in order to identify the gender of a person. This is a particularly challenging problem, as is implied from the fact that female and male average facial shapes are generally found to be very similar.

One of the primary challenges in face-based gender recognition is the step of feature selection, where one must judiciously select the type of considered features in order to improve gender recognition rates. Towards understanding this feature selection process, different types of strategies have been attempted, such as the work in Sun et al. [245] that employed genetic algorithms for eigen-feature selection. Other approaches focus on specific facial features, such as the approach by Zhang et al. [292] that focused on the eye brow and the jaw region.

Another challenge comes in unconstrained settings where the face image is affected by changes in illumination, pose, etc. While in more constrained settings face-based gender estimation has been reported to achieve classification rates of up to 99.3% (see Table III), this performance significantly decreases in more realistic and unconstrained settings.

The majority of gender classification methods contain two steps succeeding face detection, namely feature extraction and pattern classification.

Feature extraction: Notable efforts include the early work by Moghaddam et al. [169] and the work by Baluja et al. [15] who used raw pixel intensities as inputs to SVM and Adaboost classifiers, in order to achieve a 96% success rate on low resolution images. Interesting work can also be found in Cao et al. [22] who investigated facial metrology for pertinent gender traits, which resulted in error rates that were observed to be between 3.8% and 5.7% lower than that of appearance-based methods. Other feature extraction approaches are found in the work of Saatci and Town in [221], who presented an active appearance model (AAM) based geometric-approach for extracting gender and expression (using a SVM classifier with a radial basis kernel), as well as recent approaches that use SIFT [254], LBP [281], semi-supervised discriminant analysis (SDA) [19] or combinations of different features [83], [265].

Classification: A number of classification methods have been used for gender estimation, and a useful comparative guide of these classification methods can be found in Mäkinen and Raisamo [157]. One interesting conclusion of their work
was that image size did not greatly influence the classification rates. This same work also revealed that manual alignment affected the classification rates positively, and that the best classification rates were achieved by SVM.

The area of gender estimation has also received numerous other contributions such as those that go beyond using static 2D visible spectrum face-images. Related publications include the work of Han et al. [97] that explored the use of 3D face images and SVMs, achieving gender classification with an average error rate of 17.44% on the GavabDB database that contained multiple facial surface images of 45 males and 16 females. Chen and Ross [27] and Ross and Chen [219] used near-infrared (NIR) and thermal face images for gender classification. Their work demonstrated that the local binary pattern histogram (LBPH)-descriptor and SVM classifier offer the best accuracy results, which were reported to reach 93.59% for NIR images; an accuracy of 90.66% was achieved for thermal images using LBP+PCA+SVM. Satta et al. [226] used - in addition to facial features (represented by LBPH-Pyr (local) and LDA (global) features) - other contextual features such as hair (captured by HSV histograms, fuzzy color and texture histogram (FCTH), and edge histogram descriptor (EHD)) in an AdaBoost framework to obtain an accuracy of about 75%.

2) Gender from Fingerprint: Fingerprint-based gender classification has received attention in forensic anthropology as a pruning tool that can reduce the fingerprint search space by offering a likelihood that a specific fingerprint belongs to a male or a female. This approach exploits the fact that there appear to be differences in male and female fingerprints. Such differences include the ridge-thickness to valley-thickness ratio (RTVTR), the ridge count (the average ridge count is slightly higher in males than in females), and the count of white lines [182].

Gupta and Rao [91] used wavelet transformation and back propagation artificial neural networks to achieve an overall classification rate of 91.45% on a private database of 550 fingerprints (275 male, 275 female). Similar results were obtained by Badawi et al. [12], who employed Fuzzy Cognitive Maps (FCM) and neural networks to achieve a fingerprint-based gender classification rate of 88% (see Figure 4). Additionally, Tom et al. [255] used 2D wavelet transform and PCA to obtain 70% accuracy on a 547 subject-database, while Gnanasivam and Muttan [77] fused fingerprint features obtained by discrete wavelet transform (DWT) and singular value decomposition (SVD) to achieve an overall classification rate of 87.52%. Marasco et al. [160] used a combination of image quality and texture features for gender determination with overall classification rates of up to 88.7% on a dataset of 494 subjects. However, their experimental protocol does not indicate if subjects in the training and test sets were mutually exclusive. Recently, Rattani et al. [205] explored the use of classical texture descriptors - Local Binary Pattern (LBP), Local Phase Quantization (LPQ), Binarized Statistical Image Features (BSIF) and Local Ternary Pattern (LTP) - to estimate gender from fingerprint images and tested them on the WVU fingerprint dataset consisting of 237 users. In their experiments, the subjects used in the training and test sets were mutually exclusive thus resulting in a statistically unbiased evaluation. Fusion of descriptors and different fingers provided overall classification rates of up to 81.4%.

3) Gender from Iris: Thomas et al. [251] were the first to predict gender from near-infrared iris images using automated methods. They used a combination of texture features and seven geometric features, including the horizontal, vertical, and Euclidean distances between the pupil center and the iris center, to perform gender classification. Classification was performed by bagging 100 C4.5 decision trees, which resulted in an accuracy of 75% on a dataset of 28,000 iris-images, and an improved 80% accuracy when considering only Caucasian subjects.

Furthermore, Lagree and Bowyer [135] continued this line of research by exploring additional texture features (extracted by texture filters such as “spot detectors” and “line detectors”), and without any geometric features. The authors recorded an accuracy of 62% on a 600-image dataset of 60 different subjects. The reduced performance, compared to [251], is attributed to the use of a sequential minimal optimization (SMO) support vector classifier, as well as on the difference in the size of the datasets (a 50-fold difference) that resulted in a much smaller training set. Lagree and Bowyer [135] considered both gender and race classification. An interesting observation in their work was that gender classification from iris is a more challenging task than race prediction, and race prediction is harder on females than on males. A related work can be found in [16] where Bansal et al. combined statistical features and texture features using wavelets, and designed a gender prediction model using SVM that achieved an accuracy of 85.68%.

4) Gender from Body: Some biometric traits such as gait require the collection of full body images. Body-based classification of gender has received much attention since the human body has a number of cues for distinguishing gender, as well as the fact that body shape and gait have the potential to reveal the human gender from a longer distance. These cues include
body sway, waist-hip ratio, and shoulder-hip ratio (see [164]); for example, females have a distinct waist-to-hip ratio and swing their hips more, whereas males have broader shoulders and swing their shoulders more.

At the same time, extracting gender from body is negatively impacted by several influential factors which - in the case of side profile images - include shoes, background and clothes.

In addition, while the gait cycle contains plenty of information, it also introduces a high feature dimensionality, increasing computational complexity. To reduce dimensionality (a critical step for obtaining better gender estimation accuracy [177]), often, the human silhouette is extracted and a statistical feature - namely an average silhouette - is computed. This is referred to as the gait energy image (GEI). The GEI has been reported to be a good feature for gait and gender recognition because it is robust to silhouette errors and noise (see Figure 5 and also [286], [44]). To again simplify analysis, the body is divided into different regions, such as legs, waist, buttocks, back, chest and head, and these are analyzed separately for gender cues. Results have shown that the head and chest regions contribute significantly to gender cues.

From a computer vision point of view, human gait analysis can be appearance-based or it can be model-based, where characteristics such as height, step-frequency, or angle between two thighs are extracted. Given that extracting such individual characteristics can be a challenging task, most gait recognition algorithms are appearance-based. Table IV gives an overview of body-based gender classification algorithms with the best reported performance being 97.2% by Shan et al. [236], who fused GEI and facial features. The GEI-based SVM algorithm obtained a classification rate of 93%.

5) Gender from Hand: Extracting gender information from the hand dates back to forensics and archaeological efforts that deduced gender from skeletons with damaged skeletal regions that are traditionally good indicators of gender, viz., the pelvis, skull-region, and larger long bones (for single populations).

In the biometrics context, hand-based gender classification is useful as hand images are typically acquired under controlled position, orientation and illumination, resulting in lesser var-

TABLE III

Overview of face-based gender classification algorithms. Abbreviations used: principal component analysis (PCA), independent component analysis (ICA), support vector machines (SVM), Gaussian process classifiers (GPC), active appearance model (AAM), local binary pattern (LBP), active shape model (ASM), discrete cosine transform (DCT), semi-supervised discriminant analysis (SDA).

<table>
<thead>
<tr>
<th>Work</th>
<th>Features</th>
<th>Classifier</th>
<th>Datasets used for evaluation</th>
<th>Performance numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golomb et al. (1990) [79]</td>
<td>Raw pixels</td>
<td>Neural network</td>
<td>Private, 90 images</td>
<td>91.9%</td>
</tr>
<tr>
<td>Guta et al. (1998) [92]</td>
<td>Raw pixels</td>
<td>Hybrid classifier</td>
<td>FERET, 3006 images</td>
<td>96.0%</td>
</tr>
<tr>
<td>Sun et al. (2002) [245]</td>
<td>PCA</td>
<td>SVM</td>
<td>Private, 400 images</td>
<td>91.1%</td>
</tr>
<tr>
<td>Moghaddam and Yang (2002) [169]</td>
<td>Raw pixels</td>
<td>SVM</td>
<td>FERET, 1,776 images</td>
<td>96.6%</td>
</tr>
<tr>
<td>Jain and Huang (2004) [108]</td>
<td>ICA</td>
<td>LDA</td>
<td>FERET, 500 images</td>
<td>99.3%</td>
</tr>
<tr>
<td>Khan et al. (2005) [119]</td>
<td>PCA</td>
<td>Neural network</td>
<td>Private, 400 images</td>
<td>88.7%</td>
</tr>
<tr>
<td>Sun et al. (2006) [244]</td>
<td>LBP</td>
<td>Adaboost</td>
<td>FERET, 2,000 images</td>
<td>95.75%</td>
</tr>
<tr>
<td>Kim et al. (2006) [122]</td>
<td>Raw pixels</td>
<td>GPC</td>
<td>AR, 515 images</td>
<td>97.0%</td>
</tr>
<tr>
<td>Saatci and Town (2006) [221]</td>
<td>AAM</td>
<td>SVM</td>
<td>Private, 90 images</td>
<td>94.8%</td>
</tr>
<tr>
<td>Yang and Ai (2007) [281]</td>
<td>LBP</td>
<td>Adaboost</td>
<td>Private, 3,540 images</td>
<td>96.32%</td>
</tr>
<tr>
<td>Bekios-Calfa et al. (2007) [18]</td>
<td>PCA</td>
<td>SVM</td>
<td>UCN (nonpublic), 10,700 images</td>
<td>93.46% ± 1.65%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FERET, 994 images</td>
<td>93.57% ± 1.39%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PAL, 576 images</td>
<td>93.57% ± 1.39%</td>
</tr>
<tr>
<td>Xia et al. (2008) [275]</td>
<td>LBP, Gabor</td>
<td>SVM</td>
<td>CAS-PEAL, 10,784 images</td>
<td>93.74%</td>
</tr>
<tr>
<td>Mäkinen and Raisamo (2008) [157]</td>
<td>LBP</td>
<td>SVM</td>
<td>FERET, 411 images</td>
<td>86.54%</td>
</tr>
<tr>
<td>Gao and Ai (2009) [72]</td>
<td>ASM</td>
<td>Adaboost</td>
<td>Private, 1,300 images</td>
<td>92.89%</td>
</tr>
<tr>
<td>Toews and Arbel (2009) [254]</td>
<td>SIFT</td>
<td>Bayesian</td>
<td>FERET, 994 images</td>
<td>83.7%</td>
</tr>
<tr>
<td>Shan (2010) [234]</td>
<td>LBP</td>
<td>Adaboost</td>
<td>LFW, 7,443 images</td>
<td>94.44%</td>
</tr>
<tr>
<td>Guo et al. (2009) [83]</td>
<td>LBP, HOG, BIF</td>
<td>SVM</td>
<td>YGA, 8,000 images</td>
<td>89.28%</td>
</tr>
<tr>
<td>Wang et al. (2010) [265]</td>
<td>SIFT, context</td>
<td>Adaboost</td>
<td>FERET, 2,409 images</td>
<td>95.0%</td>
</tr>
<tr>
<td>Nazhir et al. (2010) [176]</td>
<td>DCT</td>
<td>KNN</td>
<td>SUMS, 400 images</td>
<td>99.3%</td>
</tr>
<tr>
<td>Ross and Chen (2011) [219]</td>
<td>LBP</td>
<td>SVM</td>
<td>CBSR NIR, 3,200 images</td>
<td>93.59%</td>
</tr>
<tr>
<td>Cao et al. (2011) [22]</td>
<td>Metrology</td>
<td>SVM</td>
<td>MUCT, 276 images</td>
<td>86.83%</td>
</tr>
<tr>
<td>Hu et al. (2011) [105]</td>
<td>Filter banks</td>
<td>SVM</td>
<td>Flickr, 26,700 images</td>
<td>90.1%</td>
</tr>
<tr>
<td>Bekios-Calfa et al. (2011) [19]</td>
<td>SDA</td>
<td>PCA</td>
<td>Multi-PIE, 337 images</td>
<td>88.04%</td>
</tr>
<tr>
<td>Shan (2012) [235]</td>
<td>Boosted LBP</td>
<td>SVM</td>
<td>LFW, 7,443</td>
<td>94.81%</td>
</tr>
<tr>
<td>Ramón-Balmaseda (2012) [203]</td>
<td>LBP</td>
<td>SVM</td>
<td>MORPH, LFW, Images of Groups, 17,814</td>
<td>75.10%</td>
</tr>
<tr>
<td>Jia and Cristianini (2015) [113]</td>
<td>Multi-scale LBP</td>
<td>C-Pegasos</td>
<td>Private, 4 million images</td>
<td>96.86%</td>
</tr>
</tbody>
</table>
Table IV

<table>
<thead>
<tr>
<th>Work</th>
<th>Feature extraction</th>
<th>Classification</th>
<th>Datasets</th>
<th>CCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cao et al. [23] (2008)</td>
<td>HOG</td>
<td>Adaboost, and Random forest algorithms</td>
<td>MIT pedestrian database (600M, 288F)</td>
<td>75%</td>
</tr>
<tr>
<td>Chen et al. [29] (2009)</td>
<td>8 WASS</td>
<td>Euclidean distance to WASS-templates</td>
<td>IRIP Gait (32M, 28F)</td>
<td>93.3%</td>
</tr>
<tr>
<td>Li et al. [144] (2008)</td>
<td>7 WASS</td>
<td>Euclidean distance to WASS-templates</td>
<td>CASIA Gait DB (31M, 31F)</td>
<td>93.28%</td>
</tr>
<tr>
<td>Yu et al. [286] (2009)</td>
<td>GEI</td>
<td>SVM</td>
<td>Human ID (84M, 16F)</td>
<td>96.7%</td>
</tr>
<tr>
<td>Yoo et al. [283] (2005)</td>
<td>figure sequential</td>
<td>SVM</td>
<td>CASIA Gait DB (31M, 31F)</td>
<td>95.97%</td>
</tr>
<tr>
<td>Shan et al. [236] (2008)</td>
<td>CCA-feature fusion of GEI and face images</td>
<td>SVM (RBF)</td>
<td>CASIA Gait DB (31M, 31F)</td>
<td>97.2%</td>
</tr>
</tbody>
</table>

Fig. 5. Gender from body: A human silhouette is extracted by subtracting the background and, subsequently, an average silhouette is computed, constituting the gait energy image, GEI. The image depicts male ((a)-(c)) and female ((d)-(f)) GEIs [44]. Specifically head and chest regions were observed to provide pertinent gender cues.

ability than, for example, facial images which are known to be affected by many factors such as facial expression changes.

In terms of suitability of analyzing hands for gender classification, it has been noted that hand breadth measurements as well as the index to ring finger ratio offer a strong insight on the gender of a person, but at the same time, these can vary between populations and are further influenced by genetics, environment and social conditions. Good studies on this can be found in Kanchan and Krishan [116] and Scheuer and Elkington [231]. Falsetti [58] and Lazenby [140] used measurements of metacarpals of the human hand and obtained gender classification rates of up to 92%. Another approach to determine the gender by Krishan et al. [127], employed the dimensions of hand and feet. Their study employed a test database containing 123 men and 123 women aged 17-20 years old from North India. By analyzing hand length, hand breadth, foot length and foot breadth to classify gender, the study concluded that left foot breadth provided the highest accuracy, which was recorded to be 86.9%.

Interestingly, while most studies on the specific topic are of anthropological and psychological nature, Amayeh et al. [8] adopted a computer-vision approach, in their effort to find gender specific traits in the shape of a hand. They segmented the hand into 6 parts (hand palm and fingers) (see Figure 6), and each part was represented by Zernike moments [121] and Fourier descriptors. Different fusion methods (feature-level, score-level and decision-level) were used to combine these descriptors, and in a very small database of 40 subjects, the method achieved an accuracy of 99%.

Fig. 6. Hand-based gender classification [8]. A female hand (left), a male hand (middle), and a segmented hand (right).

6) Gender from Speech: In the effort to classify gender, it is only natural to employ speech, since there are substantial perceived differences between typical male and female voices. Towards achieving audio based-automated gender recognition (AGR), speech signals can provide reliable traces, especially in cases of adults, while less reliable traces can be found in infants and elderly individuals (Madry et al. [155]). Audio-based gender traces can be of voice source-character (such as pitch frequency), or can be of vocal tract-character (such as the first four formants with their respective frequency, amplitude and bandwidth [111]). For voice-based approaches, the gender classification problem remains open. This is because, while the 170-275Hz female frequency range is disjoint and nicely differentiable from the male equivalent of 112-146Hz ([273]), in real world conditions consisting of background noise, low recording quality and variabilities in the physical and emotional state of a subject, human voices can be obfuscated, turning audio-based AGR into a significant challenge [197]. In clean speech data, Childers and Wu [31] analyzed ten vowels to predict the gender. Using a database of 52 talkers (27 male and 25 female), the method was recorded to have 100% gender classification accuracy. Sorokin and Makarov [239] found that features such as the instant of maximum glottis area, the maximum derivative of the area, the slope of the spectrum of harmonics of this spectrum, and the frequency pitch, can result in an accuracy of 94.7% for male voice recognition and 95.9% for female voice recognition. A comparative study of
AGR approaches can be found in the work of Walawalkar et al. [264], where different classifiers such as nearest neighbor classifier and SVMs were evaluated.

7) Hybrid approaches: In addition to the aforementioned purely-image or purely-audio based gender classification approaches, recent work has sought to further the gains of such approaches by combining them with video sources, resulting in audio-visual cues that offer a more comprehensive gender analysis, as well as higher resilience to degradation of any of these sources. Interesting work that employed this approach can be found in Pronobis and Magimai-Doss [196] which explored different audio and visual features, and proceeded to fuse both modalities at the classifier level. Experiments conducted on the BANCA corpus, verified the usefulness of this hybrid approach by showing that the integration of audio-visual cues indeed yielded higher resilience and better performance in noisy conditions. Similar to the aforementioned audio-visual approaches, other hybrid approaches seek to combine different biometric modalities in order to increase gender classification performance. Notable related work includes the work of Shan et al. [237] which combines gait and facial features. Specifically, this work employed Adaboost and SVMs to classify human GEIs along with facial features that were related by canonical correlation analysis (CCA) on the CASIA Gait Database (Dataset B) [285] consisting of 124 subjects (93 male, 31 female), and achieved an overall classification rate of up to 99.0% ± 1.3% for males and 92.0% ± 4.6% for females.

8) Databases: One distinct advantage of gender classification is that classical databases such as FERET, MORPH-II, CAS-PEAL and LFW already contain gender annotated information, while other face-based databases can be easily annotated with gender information by visual examination, as opposed to, for example, age information. Additionally, it is often the case that baseline-performances for gender classification are provided for databases such as FERET. In Table V we list different databases that have been used to evaluate gender classification algorithms and can thus serve as test-beds for developing novel gender-classification algorithms.

B. Age estimation

The ability to estimate age is considered to be an important trait among humans, for evolutionary reasons. Specifically humans learn early on to reliably estimate age of their peers based on holistic face features (e.g. outline of the face, face shape, skin texture), local face features (e.g. areas of the eyes, nose and forehead), as well as related configuration (e.g. symmetry) [212], [172]. From this category of facial features, age can be quantified from features such as the craniofacial index (see Ramanathan and Chellappa [200]), wrinkles, skin quality [17], facial hair and chin line (Alberta et al. [5]). Age can also be estimated using other body features such as hands, gait, clavicle (collarbone), the pubic bone, and the teeth [174]. A main challenge in estimating age - whether this estimation is performed by humans (human-estimated visually perceived appearance age) or by machines (algorithmically estimated age) - comes from the fact that the aforementioned features are a function of many unknowns, including genetics, health condition, living style, working environment, and sociality (Alley [7], Rhodes [212]), all of which vary across people, thus impacting age progression differently from person to person. This variability motivates research in algorithmic age estimation, which - as suggested by a recent study by Han et al. [95] - can often outperform humans at the same task.

Below, we will describe different methods of automated age estimation, focusing on the challenging module of age image representation. As in the case of gender classification, the task of automated age estimation can be separated into two modular subtasks; the aforementioned part of age image representation during which features are extracted and which will receive most of our focus here, and the modular part of automated classification or regression where the extracted features lead to either a classification decision (corresponding to say a binary decision on whether a person is ‘young’ or ‘old’), or where the extracted features are used in a regression framework to output a more refined decision (corresponding to, for example, an estimate of a person’s age). Table VI offers an overview of recent age estimation works and their reported performance. It is here worth mentioning the study by Fu et al. [68] that focuses on automated analysis of age progression, discussing automated age synthesis and also age estimation techniques.

1) Age from Face:

a) Geometric based approaches: Anthropometric (geometric based) approaches are mainly based on the cranio-facial development theory (see Alley [7]), where - in this case - the growth of the human head is modeled based on a number of anthropometric measurements over a span of time, from infancy to adulthood (see for example the case exhibited in Farkas [60] corresponding to 57 landmarks or fiducial points). Often, what provides more meaningful conclusions on age are ratios between different anthropometric measurements. This ratio-based approach was employed by Ramanathan and Chellapa [200] who used 8 such distance-measure ratios to model age progression. Their model was tested on a database partly adopted from FG-NET4, and was shown to provide 8-15% improvement in face recognition performance, on ages up to

---

2)https://www.idiap.ch/dataset/mobio
3)http://www.nist.gov/itl/iad/ig/colorferet.cfm
4)www.fgnet.rsunit.com
18 years. In terms of drawbacks, anthropometric-based models tend to be better for only differentiating between infants and adults. Furthermore, the corresponding measurements and their associated ratios can only be reliably determined from frontal or 3D images. An early 3D approach by Xia et al. [274], which was tested on the FRGCv2 dataset with 466 scans, was shown to achieve a mean absolute error (MAE) of 3.29 years for 4 age classes.

b) Approaches based on appearance models: Another image representation approach relates to appearance models, and considers texture and shape features (and thus indirectly anthropometry and wrinkles), for the purpose of determining the human age. Early approaches were presented by Kwon and Lobo [132], [133], analyzing wrinkles in facial regions such as the forehead and in the periorbital region. This analysis determined the classification of young and senior adults. A recent related study by Guo et al. [88] obtained best reported results (MAEs = 2.6 years, tested on the YGA database). These results were obtained by simultaneously considering age and gender automatic estimation, biologically-inspired features and manifold learning techniques.

Approaches based on active appearance models (AAMs) (see [33]) form a special class under the aforementioned appearance-based models. Interesting work can be found in [139] (Lanitis et al.) which extended the AAMs by proposing an aging function to describe the variation in age. In this aging function, the input is a vector containing 50 raw model parameters learned from the AAMs. The proposed function defines the relationship between the age of the individuals and the parametric description of the face images. Based on age-simulation (using a ‘Weighed Person Specific Age Simulation’ method), the age in all images was normalized to eliminate the effects of aging. This approach led to an improvement of face recognition performance from 63% to 71% and from 51% to 66%, on two different datasets. Other interesting AAM-based approaches have been employed by Guo et al. [84] and by Lakshmirprabha [136].

c) Approaches based on aging pattern subspaces: Aging pattern subspaces are formed from a sequence of images per person representing the chronological progress of aging, i.e., representing an aging pattern. Specifically, each point in the subspace corresponds to an aging pattern. The age of an unseen face image is determined by the projection in the subspace that can best reconstruct the face image. The position of the face in the subspace indicates the age of the associated face. In this approach, which was introduced by Geng et al. [74], [75], this sequence of images forms what is specifically known as the Aging pattern Subspace (AGES), which is then properly represented - for each person, at each given age - by AAMs (missing ages are synthesized using an EM-like iteration algorithm). In the testing phase, an image of a face is matched to the faces of the subspace (i.e., of AGES), and the degree of the matching is then translated into an estimate of the age. This approach was tested [74], [75] on the FG-NET face database, resulting in a reported MAE of 6.77 years. While promising, this approach has the drawback that it requires multiple images per subject in order to learn age progression. Additionally, wrinkles are - in general - not sufficiently represented by AAMs, and are thus downplayed in the process of estimation. A related approach can be found in the recent work by Guo et al. [88], [89] which draws from the feed-forward path theory of the cortex for visual processing. In this approach, after sequentially filtering an image with a Gabor filter and a standard deviation based filter, the obtained features are subjected to PCA resulting in a lower-dimensional bio-inspired-features (BIF). Recently, deep-learning aging pattern approaches were presented by Wang et al. [267] and Tian and Chen [253].

d) Approaches based on age manifolds: First presented by Fu et al. [71], the main idea behind this approach is that functions that estimate age can be of reduced dimensionality, i.e., will take as input only a reduced number of characteristic parameters. Such functions are used in conjunction with manifold learning algorithms which reveal these characteristic parameters and find a low-dimensional representation of the data (cf. Seung and Lee [232]). The resulting manifold can be linear (Yan et al. [280]) or non-linear (Tenenbaum et al. [250]). Furthermore Guo et al. [84] explored learning manifolds based on projections that preserve orthogonal locality. This approach was reported to provide an MAE of 5.3 years on FG-NET, if the size of the training data is large enough. The related linear manifold embedding method is referred to as orthogonal locality preserving projections (OLPP). Enhanced age estimation performance was recorded by Li et al. [143] who performed age-manifold learning under locality preserving requirements as well as under ordinal requirements.

---

TABLE V
EXAMPLES OF DATABASES THAT HAVE BEEN USED FOR EVALUATING AUTOMATED FACE AND BODY GENDER ESTIMATION METHODS.

<table>
<thead>
<tr>
<th>Name</th>
<th>Number of Subjects</th>
<th>Acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Georgia Tech Database</td>
<td>50</td>
<td>15 facial images per subject, different pose and lighting conditions</td>
</tr>
<tr>
<td>GavabDB</td>
<td>61</td>
<td>427 facial images of 3D meshes, 9 3D-images provided per subject</td>
</tr>
<tr>
<td>MORPFF [214] I and II</td>
<td>515 + 4000</td>
<td>multiple face shots per subject</td>
</tr>
<tr>
<td>LF2: Labeled Faces in the Wild</td>
<td>1680</td>
<td>13,233 images acquired under large variability in pose, expression and illumination</td>
</tr>
<tr>
<td>MIT AI Gait Data</td>
<td>24</td>
<td>indoor acquisition of 194 gait-sequences</td>
</tr>
<tr>
<td>CASIA Gait Dataset: (Dataset B) [285]</td>
<td>124</td>
<td>multiple sequences of outdoor walking subjects, extracted silhouettes available</td>
</tr>
<tr>
<td>CMU Mobo Database [81]</td>
<td>25</td>
<td>subjects walking on treadmill (slow, fast, inclined and carrying a ball)</td>
</tr>
<tr>
<td>Gait Challenge Database</td>
<td>122</td>
<td>1,870 sequences spanning 5 covariates</td>
</tr>
</tbody>
</table>

---

References:

http://marathon.csee.usf.edu/TrackB/results.php
http://www.ai.mit.edu/projects/gait/
http://www.gavab.etsii.urjc.es/recursos/
http://vis-www.cs.umass.edu/lfw/
http://www.anefian.com/research/face
http://www.galton.utoronto.ca/review/ptb/411/GaltonDB
d
http://www.galton.utoronto.ca/review/ptb/411/GaltonDB
http://faceaginggroup.com/
d
http://vcs-www.cs.umass.edu/lfw/
http://www.ai.mit.edu/projects/gait/
d
http://marathon.csee.usf.edu/GaitBaselines/
a) Limitations and challenges in race classification: Before addressing some of the progress in automated race classification, it is worth touching upon some points which suggest that ethnicity- or race-categorization is a rather challenging problem. The first important point - as indicated in the book of Mordini and Tzovaras [170] - is that intra-class variation can often be significantly larger than the inter-class variation between races. Similarly, recent findings by geneticists (cf. [161], [9]) show that there is often no clear relation between races on one hand and the frequency of particular genetic variants on the other. In the same vein, Rosenberg et al. [218] reported that within-population differences among individuals account for 93%-95% of genetic variations, whereas major ethnic groups differentiate only by 3%-5%. At the same time, however, the work in [218] identified six main genetic clusters, five of which corresponded to major geographic regions, and sub-clusters that often corresponded to individual populations.

Related to the challenging nature of race-categorization is the “other race effect”, where humans charged with a face recognition task are observed to perform significantly better when having to recognize the face of a person from their own race (cf. O’Toole et al. [184]). Interestingly, such an effect has also been witnessed in automated face recognition [192][184]. Specificaly, this “other race effect” was observed in an international face recognition competition where a fusion of algorithms from Western countries performed better in recognizing Caucasian faces, while a fusion of algorithms from East Asian countries performed better in recognizing East Asian faces.

b) Approaches based on chromaticity: Chromaticity, or skin tone, has long been employed as the primary feature for race classification; see for example Xie et al. [278], Roomi et al. [217], and Tariq et al. [249]. The latter work used human silhouette images, and achieved classification rates of 80.37% for ‘East and Southeast Asian’ subjects, 78.67% for ‘White Caucasian’ subjects, 58.21% for ‘Black’ subjects, and 33.33% for ‘South Asian’ subjects. We note here that chromaticity-based race classification has the limitation of being strongly sensitive to illumination changes.

c) Age from Body: In the context of age estimation, Lu and Tan [150] recently analyzed gait. In this setting - where, as the authors state, gait appearances vary between males and females even within the same age group - the approach was based on learning a multilabel-guided subspace to better characterize and correlate the age and gender information of a person for estimating human age. The authors extract Gabor features including both Gabor magnitude and Gabor phase information of a gait sequence, and perform multiple feature fusion to enhance the age estimation performance. The related experimental results present an MAE of 3.02 years on the USF database [225] consisting of 1870 gait sequences from 122 subjects (85 male and 37 female subjects of age between 19 and 59 years old).

3) Age from Hand: One characteristic work in this setting is that by Shamir [233] which presented a new approach of automatic age classification based on the features that appear on the back of the human hand. Estimating age from the hand, enjoys - as opposed to age estimation from the face - the advantage of being privacy preserving, and the advantage of being invariant to facial makeup and plastic surgery. Experimental results in the same work show that the method can differentiate between older and younger individuals with an accuracy of 88%. The hand photos were taken from Baltimore Longitudinal Study of Aging (BLSA). The dataset includes 106 subjects, and 212 images (two from each subject).

4) Databases: An increasing number of databases feature biometric data related to different ages. We report some of these in Table VII.

C. Ethnicity and race classification

The traditional definition of race is related to biological factors and often refers to a person’s physical appearance corresponding to traits such as skin color, eye color, hair color, bone/jaw structure, face and body shape, and other traits, while the traditional definition of ethnicity is more related to sociological factors and it relates primarily to cultural identifiers such as nationality, culture, ancestry, language as well as beliefs. The terminology overlaps and is often used interchangeably in the biometric literature, and hence we will also adopt the same loose convention here.

http://www.blsa.nih.gov/
TABLE VI
AGE ESTIMATION TECHNOLOGIES BASED ON CLASSIFICATION, ABBREVIATIONS USED; ACTIVE APPEARANCE MODEL (AAM), APPEARANCE MODEL (APM), AGE MANIFOLD (AMF), BIO INSPIRED FEATURES (BIF), DISTANCE METRIC LEARNING (DML), CUMULATIVE ATTRIBUTES (CA), CONVOLUTIONAL NEURAL NETWORK (CNN), DEEP LEARNED AGING PATTERN (DLA), GRAPH-BASED COMPOSITIONAL AND DYNAMIC MODEL (GB), ORDINAL DISCRIMINATIVE AGING (PLO), MALE (M), FEMALE (F).

<table>
<thead>
<tr>
<th>Work</th>
<th>Features</th>
<th>Datasets used for evaluation</th>
<th>Number of subjects</th>
<th>Accuracy</th>
<th>Classification / Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guo et al. (2008)</td>
<td>AAM, AMF (OLPP)</td>
<td>FG-NET, YGA</td>
<td>82</td>
<td>88%</td>
<td>C</td>
</tr>
<tr>
<td>Kwon and Lobo (1999)</td>
<td>facial ratios, wrinkle index</td>
<td>Private</td>
<td>47</td>
<td>100%</td>
<td>C</td>
</tr>
<tr>
<td>Kanno et al. (2001)</td>
<td>APM</td>
<td>Private</td>
<td>110(M)</td>
<td>80%</td>
<td>C</td>
</tr>
<tr>
<td>Lanitis et al. (2004)</td>
<td>AAM</td>
<td>Private</td>
<td>40</td>
<td>MAE = 3.82 − 5.58</td>
<td>R</td>
</tr>
<tr>
<td>Zhou et al. (2005)</td>
<td>APM</td>
<td>FG-NET</td>
<td>82</td>
<td>MAE = 5.81</td>
<td>R</td>
</tr>
<tr>
<td>Ueki et al. (2006)</td>
<td>APM (raw image)</td>
<td>WIT-DB</td>
<td>3000(M), 2500(F)</td>
<td>50%(M), 43%(F)</td>
<td>C</td>
</tr>
<tr>
<td>Takimoto et al. (2006)</td>
<td>APM</td>
<td>HOIP</td>
<td>113(M), 139(F)</td>
<td>57.3%, 54.7%</td>
<td>C</td>
</tr>
<tr>
<td>Takimoto et al. (2007)</td>
<td>APM</td>
<td>HOIP</td>
<td>113(M), 139(F)</td>
<td>3.0(M), 4.1(F)</td>
<td>R</td>
</tr>
<tr>
<td>Geng et al. (2007)</td>
<td>AGES</td>
<td>MORPH</td>
<td>515</td>
<td>≈ 70%</td>
<td>C</td>
</tr>
<tr>
<td>Fu and Huang (2008)</td>
<td>AMF (OLPP)</td>
<td>YGA</td>
<td>800(M), 800(F)</td>
<td>MAE = 8.0(M), 7.8(F)</td>
<td>R</td>
</tr>
<tr>
<td>Zhang et al. (2008)</td>
<td>APM (patches)</td>
<td>YGA</td>
<td>800(M), 800(F)</td>
<td>MAE = 5.4(M), 6.33(F)</td>
<td>R</td>
</tr>
<tr>
<td>Guo et al. (2008)</td>
<td>quadr. regression</td>
<td>YGA</td>
<td>4000(M), 4000(F)</td>
<td>MAE = 6.0(M), 5.5(F)</td>
<td>R</td>
</tr>
<tr>
<td>Ni et al. (2009)</td>
<td>APM (patches)</td>
<td>Web data, FG-NET</td>
<td>78711 images</td>
<td>MAE = 7.42</td>
<td>R</td>
</tr>
<tr>
<td>Guo et al. (2009)</td>
<td>APM(BIF)+AMF</td>
<td>YGA</td>
<td>800(M), 800(F)</td>
<td>89.7%</td>
<td>C</td>
</tr>
<tr>
<td>Xiao et al. (2009)</td>
<td>DML</td>
<td>FG-NET</td>
<td>82</td>
<td>84%</td>
<td>C</td>
</tr>
<tr>
<td>Yan et al. (2009)</td>
<td>SSE</td>
<td>FG-NET</td>
<td>82</td>
<td>MAE=5.21</td>
<td>R</td>
</tr>
<tr>
<td>Ni et al. (2009)</td>
<td>RMIR</td>
<td>Web data, MORPH</td>
<td>77,021+55,608</td>
<td>MAE=8.6</td>
<td>R</td>
</tr>
<tr>
<td>Suo et al. (2010)</td>
<td>GB</td>
<td>FG-NET</td>
<td>82</td>
<td>82.7%</td>
<td>C</td>
</tr>
<tr>
<td>Lakshmirabha et al. (2011)</td>
<td>AAM</td>
<td>FG-NET</td>
<td>82</td>
<td>77%</td>
<td>C</td>
</tr>
<tr>
<td>Li et al. (2012)</td>
<td>PLO</td>
<td>FG-NET</td>
<td>82</td>
<td>88%</td>
<td>C</td>
</tr>
<tr>
<td>Chen et al. (2013)</td>
<td>CA - AAM</td>
<td>FG-NET, MORPH</td>
<td>82</td>
<td>MAE=4.67</td>
<td>R</td>
</tr>
<tr>
<td>Wang et al. (2015)</td>
<td>CNN - DLA</td>
<td>MORPH, FG-NET</td>
<td>515</td>
<td>MAE=4.77</td>
<td>R</td>
</tr>
</tbody>
</table>

TABLE VII
EXAMPLES OF DATABASES THAT HAVE BEEN USED FOR EVALUATING AUTOMATED AGE ESTIMATION METHODS.

<table>
<thead>
<tr>
<th>Name</th>
<th>Number of Subjects</th>
<th>Acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td>FG-NET</td>
<td>82</td>
<td>1002 color and gray scale images with variation in pose, illumination and expression. Age: 0–69 years.</td>
</tr>
<tr>
<td>MORPH I and II</td>
<td>515 + 4,000</td>
<td>Database annotated for age, gender, ethnicity, height, weight and ancestry. Album 1: 1,724 images corresponding to 515 subjects, captured between 1962 and 1998. Album 2: more than 20,000 images from more than 4,000 subjects. Age: 27–68 years.</td>
</tr>
<tr>
<td>WIT-DB</td>
<td>5,500 Japanese subjects</td>
<td>2,500 females and 3,000 males (12,008 + 14,214 images respectively). Age: 3–85 years.</td>
</tr>
<tr>
<td>HOIP</td>
<td>300</td>
<td>306,600 images (half males and half females). Age: annotated in 10 age categories from 15 to 64 with a range of five years per category, all with neutral expressions.</td>
</tr>
<tr>
<td>YGA</td>
<td>1,600 Asian subjects</td>
<td>8,000 high-resolution outdoor color images (half males and half females). Age: 0–93 years. The images include variations of illumination, facial expression and makeup.</td>
</tr>
<tr>
<td>3D Morphable Database</td>
<td>200 adult + 238 teenager</td>
<td>3D scans (half males and half females)</td>
</tr>
</tbody>
</table>

b) Approaches based on global features: Global feature representation, or holistic representation, has been found to preserve configurational face information (interrelations between facial regions), which is helpful towards race classification. Classical methods include the work of O’Toole et al. based on PCA of raw pixel intensities [175], [192], achieving accuracies of about 80% for distinguishing between Japanese and Caucasian subjects. The work by Lu and Jain [152] employed an LDA-based algorithm to classify 263 subjects into 2 categories (Asian and Non Asian) obtaining a success rate of about 96%.

c) Approaches based on local feature descriptor representation: These include methods proposed by Lin et al. [147], Klare and Jain [124], Heo and Savvides [102], and Hosoi et al. [104]. The latter employed Gabor Wavelet Transformation and retina sampling, along with a SVM classifier. On a
database containing 771 Asian, 774 European and 446 African subjects, an accuracy of about 94% was obtained. Zang and Yi [281] examined 11680 Asian and 1016 non-Asian subjects, and employed LBP in an AdaBoost classification to separately determine their ethnicity, gender, and age. For ethnicity classification, the resulting EER was 2.98%. Another approach was presented by Fu et al. [66] who used topographic independent component analysis (TICA) to form a hierarchical multi-level cortex-like mechanism, which achieved a classification rate of 82.5%.

Finally, one can also encounter early hybrid approaches that aim to combine local and global features. One such hybrid approach was presented by Ding et al. [50] that boosted local texture and global shape features, resulting in accuracies of up to 98.3%.

2) Race from Iris: Recent research has explored the possibility of deducing race from iris images (see Figure 8). For example, Qiu et al. [198] showed that geometric characteristics of the iris exhibit significant differences across races. This has been further confirmed in [199], [135], [294]. Building on this property, Qiu et al. [198] developed a model to predict ethnicity from iris texture. Their work considered a binary classification problem (Asian vs. non Asian), and employed the CASIA database with 2400 images for the Asian subjects, while for the non Asian subjects it employed the UPOL database with 384 images of 64 subjects and the UBIRIS database [195] with 1198 images of 240 subjects. Using AdaBoost, the work recorded an accuracy of 85.95%. The images in the UPOL and UBIRIS databases were acquired using visible light, while the images in CASIA were acquired in near-infrared illumination. The authors conducted an experiment to establish whether illumination affects the proposed method for ethnicity classification. The experiment consisted of selecting irises of Asian subjects from UPOL, UBIRIS and CASIA, and attempting to classify images into those captured under visible light versus those captured in the presence of near-infrared illumination. The related results showed a near-random classification (of illumination), suggesting that the presented ethnicity classification is due to inherent characteristics of iris texture, rather than illumination. An extension of this work by the same authors [199] computed texture features using 40 Gabor filters formed by eight orientations and five scales. The work used 200 images of Asian subjects from the CASIA database, and 200 images from the BioSecure database for the non-Asian subjects. The test dataset contained 2,400 additional images of 60 subjects (30 from the CASIA and 30 from the BioSecure database). A K-means algorithm was then used to form 64 clusters representing textons that are commonly occurring fundamental texture elements. An image was then analyzed in terms of 64-texton histograms, where each pixel was classified into one of 64 textons, so that an image is represented by a 64-element feature vector. With the above in place, an SVM classifier was then applied, achieving a correct classification rate of up to 88.3%. Lagree and Bowyer [135] focused on analyzing different features and texture filters. Using SMO support vector algorithms with Weka’s default parameter settings, the work reported an accuracy of 90.58% (the employed dataset consisted of 60 subjects and 600 images).

Fig. 8. Iris images corresponding to different ethnic groups, (a) a subject with Caucasian ethnicity, (b) a subject with Asian ethnicity [135].

3) Databases: In terms of classical databases - such as the FERET database - the advantage of having annotated ethnicity/race information is diminished by the fact that this information is often statistically unbalanced, simply because such databases were not created with race-classification in mind. This often forced researchers in multi-race classification to combine different databases to ensure a racially balanced and diverse information set. As a consequence of this, one faces the problem of employing different datasets that were created under different acquisition conditions. As a result, recent efforts have been made to create racially diverse databases. We list some of these in Table IX.

D. Relation between age, gender and race

It is often the case that a single facial feature can carry information about different soft biometric traits. That is why traits such as age, gender and ethnicity are often treated and categorized simultaneously (cf. [147], [87], [159], [94]) - an approach that is in line with the perceived correlation of these soft biometric traits. In addition to the correlation, exploring this intertwined nature of traits can carry substantial advantages; for example, from a genetic point of view, understanding the interaction of race with aging allows for conclusions on race-based differences in longevity and aging-associated diseases, as well as the possible role of genetic factors in such differences [73].

This perceived correlation between these traits has motivated additional work such as that in [59], which investigated if ethnicity-based gender classification can improve the accuracy of three gender estimation algorithms, namely pixel-based, HOG and LBP. The work concluded that joint treatment is not beneficial and that gender and ethnicity can be estimated separately, due to the fact that features used to estimate gender are shared by all ethnic groups and features used for ethnicity classification are present in both female and male faces.

III. ANTHROPOMETRIC ATTRIBUTES

We next focus on soft biometric traits that have been historically used to quantify the geometry and shape of the face, body and skeleton (i.e., of classical anthropometric measures).

http://phoenix.inf.upol.cz/iris/
http://biosecure.it-sudparis.eu/AB/
http://www.cs.waikato.ac.nz/ml/weka/
TABLE VIII
ETHNICITY CLASSIFICATION. ABBREVIATIONS USED: DECISION TREES (DT), MULTI-SCALE MULTI-RATIO LBP (MM-LBP), KERNEL CLASS-DEPENDENT FEATURE ANALYSIS (KCFA), ORIENTED GRADIENT MAPS (OGMs), GATE ENERGY IMAGE (GEI), SUPPORT VECTOR MACHINES (SVM), CATEGORIES (CAT).

<table>
<thead>
<tr>
<th>Work</th>
<th>Features</th>
<th>Classifier</th>
<th>Datasets used for evaluation</th>
<th>Nr. of images / subjects</th>
<th>Cat.</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gutta et al. (98) [92]</td>
<td>RBF networks</td>
<td>inductive DT</td>
<td>FERET</td>
<td>1009 subj.</td>
<td>4</td>
<td>92.0%</td>
</tr>
<tr>
<td>Viola et al. (02) [262]</td>
<td>Haar-features</td>
<td>Adaboost</td>
<td>Private</td>
<td>4132 images</td>
<td>2</td>
<td>Total error &gt; 20.8%</td>
</tr>
<tr>
<td>Hosoi et al. (04) [104]</td>
<td>Gabor Wavelets, retina sampling</td>
<td>SVM</td>
<td>Private</td>
<td>1991 images</td>
<td>3</td>
<td>94%</td>
</tr>
<tr>
<td>Lu and Jain (04) [152]</td>
<td>LDA ensembles at multiple scales</td>
<td>LDA-based classifier</td>
<td>Private</td>
<td>263 subj.</td>
<td>2</td>
<td>96.3%</td>
</tr>
<tr>
<td>Lu et al. (06) [151]</td>
<td>Sampled range and intensity images</td>
<td>SVM</td>
<td>Private</td>
<td>1240 images, 376 subj.</td>
<td>2</td>
<td>Average error rates 2%</td>
</tr>
<tr>
<td>Yang and Ai (07) [281]</td>
<td>LBP, Chi square distance</td>
<td>AdaBoost</td>
<td>Private and FERET</td>
<td>12696 images</td>
<td>2</td>
<td>97%</td>
</tr>
<tr>
<td>Tariq et al. (09) [249]</td>
<td>Shape context and shape distance (silhouetted face profiles)</td>
<td>KNN</td>
<td>generated from 3D face models</td>
<td>441 images</td>
<td>4</td>
<td>33.33% - 80.37%</td>
</tr>
<tr>
<td>Zhang and Wang (09) [293]</td>
<td>2D and 3D MM-LBP</td>
<td>Adaboost</td>
<td>FRGC v2.0</td>
<td>180 subj.</td>
<td>2</td>
<td>99.5%</td>
</tr>
<tr>
<td>Zhang et al. (12) [291]</td>
<td>Uniform LBP face and gait</td>
<td>SVM</td>
<td>Private</td>
<td>36 subj.</td>
<td>2</td>
<td>93%</td>
</tr>
<tr>
<td>Xie et al. (12) [278]</td>
<td>KCFA + color based features</td>
<td>KNN, SVM</td>
<td>Private and MBGC</td>
<td>104,000 images</td>
<td>3</td>
<td>95%-98%</td>
</tr>
<tr>
<td>Zhang et al. (12)[290]</td>
<td>GEI: gait / Gabor: face</td>
<td>SVM / Adaboost</td>
<td>Private</td>
<td>22 subj.</td>
<td>2</td>
<td>&gt; 95%</td>
</tr>
<tr>
<td>Ding et al. (13) [50]</td>
<td>OGMs in 3D images</td>
<td>Adaboost</td>
<td>FRGC v2.0</td>
<td>466 subj.</td>
<td>2</td>
<td>98%</td>
</tr>
</tbody>
</table>

TABLE IX
EXAMPLES OF DATABASES THAT HAVE BEEN USED FOR EVALUATING AUTOMATED RACE OR ETHNICITY CLASSIFICATION METHODS.

<table>
<thead>
<tr>
<th>Name</th>
<th>Number of Subjects</th>
<th>Number of Classes</th>
<th>Ethnicity Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRGC 2.0</td>
<td>4007</td>
<td>5</td>
<td>Validation set: 22% Asian, 68% White, 10% Other</td>
</tr>
<tr>
<td>Color FERET</td>
<td>2,946</td>
<td>4</td>
<td>1,902 Caucasian, 352 Asian, 464 Oriental, 228 African [92]</td>
</tr>
<tr>
<td>Cohn-Kaadé</td>
<td>210</td>
<td>3</td>
<td>170 Euro-American, 27 Afro-American, 13 others [153]</td>
</tr>
<tr>
<td>JACFEE</td>
<td>56</td>
<td>2</td>
<td>28 Caucasian, 28 Japanese [60]</td>
</tr>
<tr>
<td>EGA</td>
<td>469</td>
<td>5</td>
<td>53 African-Americans, 111 Asians, 162 Caucasians, 75 Indian, 68 Latino [99]</td>
</tr>
<tr>
<td>MORPH-II</td>
<td>78,207</td>
<td>3</td>
<td>15,996 White, 58,326 Black, 3,885 Other [96]</td>
</tr>
<tr>
<td>PCSO</td>
<td>100,012</td>
<td>3</td>
<td>69,116 White, 26,457 Black, 4,439 Other [96]</td>
</tr>
<tr>
<td>LFWS</td>
<td>4,211</td>
<td>3</td>
<td>3,501 White, 352 Black, 358 Other [96]</td>
</tr>
</tbody>
</table>

These traits have become pivotal in the context of geometry extraction - i.e., the localization of facial landmarks related to eyes, mouth, nose, chin, etc. - which is a key step in a number of applications ranging from human identification to gender, ethnicity, and age estimation to emotion and expression recognition [288].

While a number of geometric estimation methods have been developed, these methods have to account for several different factors. Firstly, some of the geometric measurements can be correlated (Farkas [60]); thus, the model used for assessing correlation can impact the accuracy of estimation. Secondly, the estimation methods have to account for variations in sensor and calibration; this is necessary to ensure that the estimated 3D affine measurements are consistent across different systems [35][295]. Thirdly, if there are multiple cameras focusing on the same scene, then it may be necessary to judiciously combine the complementary information provided by these cameras during the estimation process.

“Forensic anthropometry” [126] studies somatometry (measurements of the body) [206], [191], [99], cephalometry (measurements of the head) [107], craniometry (measurements of the skull) [49] and osteometry (measurements of the bones of the skeleton) [242], and is used for person identification in the case of unknown human remains [10].
A. Anthropometric features from face

A number of efforts have been directed to suitably capture and represent different geometric anthropometric measures of the face. As mentioned earlier, the main reason that such facial geometric anthropometric measures are of importance, has to do with the fact that localization of facial landmarks (related to eyes, mouth, nose, chin), is often a key step towards precise geometry-extraction, which is in turn crucial towards precise geometry-extraction, which is in turn crucial for human identification and a class of other recognition systems [288] that generally employ these traits as trackers. In this context, algorithms such as AAM, active shape model (ASM), and deformable shape model have been used (cf. Ding and Martinez [51]) to obtain such facial landmarks.

Further motivation comes from the fact that - as noted by Nixon in [179] - a judicious combination of these facial features can result in a relatively high degree of distinctiveness for person recognition. In addition, face-based geometric human metrology is an important tool in forensic anthropology [116].

Another line of interesting work is the design of algorithms that robustly recognize people in the presence of occluded and disguised faces. In this setting, Ramanathan and Wechsler [202] combine appearance based approaches (PCA + LDA), and anthropometric / geometric measurements (19 manually extracted geometric measurements of the head and shoulders) via decision-level fusion (neural networks) and feature-level fusion (AdaBoost), to design an algorithm that is robust to occluded and disguised faces. Related work can also be found in Ghalleb et al. [76] where facial measurements - ultimately used for human recognition - are extracted based on geometry in the detected face (SMQT and split up SNOW classifier). This resulted in a soft biometric system, which was fused with different classical biometric systems, to achieve an improved matching accuracy.

B. Anthropometric features from the body

Among the many geometric traits, body height is the most prominent. In extracting this measure, different challenges remain, including that of the human pose which can serve as a primary biasing factor. Therefore, BenAbdelkader and Davis [20] resorted to averaging different body measurements over different poses in order to reduce the pose bias. Other methods that result in robust height estimation can be found in the work by Madden and Piccardi [154] where a person's height was estimated from surveillance video, thus allowing for session-based biometric matching using two disjoint camera views.

Body height can often be estimated together with other geometric measures. For example, the aforementioned work by BenAbdelkader and Davis [20] considered shoulder breadth (bideltoid breadth) in addition to body height, and the two measures were jointly averaged for the purpose of improving multi-target tracking across multiple cameras. In a similar spirit, but in the setting of person identification, Johnson and Bobick [115] extracted four static parameters (body height, torso length, leg length, and step length), achieving a typical error of about 30cm for anthropometric measure estimation. Furthermore, BenAbdelkader et al. [32] used height and stride (extracted from gait videos) for person recognition, achieving true recognition rates of 47% for fronto-parallel sequences of 41 people, and 65% for non-fronto-parallel sequences of 17 people.

Recently, body shape has been estimated in millimeter wave images (acquired at 94 GHz) [80]. Contour coordinates and dynamic time warping were used resulting in an EER of 1.33% on a dataset of 50 individuals (BIOGIGA [171]).

1) 3D techniques in geometric anthropometric measurements: Recently, 3D techniques have been used to obtain geometric anthropometric measurements. Some of these works include the study by Carnicky and Chorvat [24] that focused on the acquisition of 3D measurements with structured light illumination, and the work by Allen et al. [6] that developed a method for fitting high-resolution template meshes to 250 detailed human body range scans from the 3D CAESAR database with sparse 3D markers to model the variation of different bodies.

Additionally, in the work of Adjeroh et al. [2] spanning face and body metrology, the 1D-measurements in the 3D CAESAR database were used to show a high degree of correlation between certain face and body measurements. Utilizing the correlation structure between measurements, the researchers showed that gender could be estimated with a success rate of 100% (this corresponded to 44 measurements), and body weight could be estimated with a success rate of 92.5%. Based on the idea that some anthropometric measurements are highly correlated and that missing values can thus be predicted, an iterated closest point (ICP) registration-model and Laplacian mesh was presented by Hasler et al. [98], where a detailed 3D human body shape in the presence of clothes was modeled based on a space of human shapes, learned from a large database of registered body scans. Similarly, by using the visual hull, a method for estimating detailed 3D body shape under clothes was presented by Balan and Black [14]. This work used the HumanEva10 dataset. Furthermore, Criminisi et al. [35] computed 3D affine measurements from single perspective views, where affine scene structure could be determined from the image, without knowledge of the camera's internal calibration (e.g. focal length), nor of the explicit human pose. Interesting work can also be found in Godil and Ressler [78] who performed similarity-based retrieval in the 3D CAESAR anthropometric database (3D scans), where the technique was based on body and head shape representation. Finally geometric/anthropometric features have also been exploited to enhance gait recognition accuracy (cf. Moustakas et al. [173]).

C. Databases

While some datasets include the annotation of pertinent facial landmarks such as eyes, nose and mouth, these are mainly used for image alignment. Often, more detailed landmark annotation in common datasets is performed manually, in order to capture the necessary measurement ratios, which are often more valuable than the measurements themselves.

\[^{9}\text{http://www.hec.afrl.af.mil/cardlab/CAESAR/index.html}\]

\[^{10}\text{http://vision.cs.brown.edu/humaneva/}\]
Notable exceptions include the two paid-for datasets of the Civilian American and European Surface Anthropometry Resource Project (CAESAR).

CAESAR 3D Anthropometric Database (North American Edition) includes 40 anthropometric / geometric measurements per person, from a North American population sample of 2,400 male and female subjects, aged 18-65. It also includes demographic information and 3D model scans. There are several poses of each person.

CAESAR 3D Anthropometric Database (European Edition) was recorded in the same project as the North American Edition and also features 40 anthropometric / geometric measurements from a European population sample of 2,000 European male and female subjects, aged 18-65. Demographic information as well as 3D scans are provided under several poses.

IV. MEDICAL ATTRIBUTES

A. Image-based automated self diagnostic methods

is a novel and promising approach of enhancing classical medical care. Such methods, when applied properly, can provide a broad range of medical benefits, including early detection and prevention of illnesses, as well as efficient treatments when possible. Such methods are needed now more than ever, due to the ever increasing burden on traditional medical care systems. With this need in place, we are already seeing different medical sectors employ computer vision schemes, such as in the crucial area of monitoring the health of elderly people to improve their safety, autonomy and quality of life [216], [186], [266]. Such promising applications, and many others, have introduced the need for algorithms and systems that can reliably monitor the state of different human traits, which in turn accentuates the importance of being able to properly characterize soft biometric traits such as body weight, body mass index (BMI), or even skin color and quality.

B. Health and weight

Again in the context of health, body weight and BMI have received increased attention in recent years. This boost in attention is related to health concerns that have been expressed in several countries. For example, it has been stated that more than two-thirds of adults in the United States are now overweight or obese [282], and obesity - which is now recognized as a disease by the American Medical Association [12] - accounts for nearly one out of every 10 American deaths. Specifically, an increased body weight and thus BMI [13], is often associated with many health problems such as diabetes, heart disease, certain cancers, strokes, osteoarthritis and gallstones, all of which cost the United States 223 billion per year [14]. This has motivated the use of automated techniques for estimating body weight and BMI - that are derived from body weight and body height - which further has the advantage of being contactless, low-cost, unobtrusive, and unconstrained [15].

Recently, BMI was estimated from face images by Wen and Guo [270]. In their work, the face was detected, normalized, and an active shape model was fitted, based on which, geometry and ratio features were extracted (cheekbone to jaw width, width to upper facial height ratio, perimeter to area ratio, eye size, lower face to face height ratio, face width to lower face height ratio and mean of eyebrow height), normalized and finally subjected to support vector regression. When training and testing using the MORPH-II database - which contains 55,000 images, with variations in age and ethnicity - the method obtained MAEs ranging from 2 kg/m² ('normal weight') to 9.5 kg/m² ('obese'), where the higher errors generally related to medically atypical cases such as 'under weight' and 'obese' subjects. Prior to that, Velardo and Dugelay [260] had devised a method to estimate weight from manually measured body-anthropometric data, by constructing a weight prediction metric using multiple regression analysis. When training and testing using the NHANES dataset [16] - which contains data of 28,000 subjects - the results showed that weight could be estimated in 93% of the test data with an error of less than 10%. Velardo et al. [261] presented a practical study with a weight estimation error of 4% on a self recorded Kinect RGBD dataset of 15 subjects, where body anthropometric measurements were extracted automatically. This work proceeded also to evaluate - based on the same measurements - the subjects' height and gender, using real-time reduced-complexity 3D model fitting. A related work for weight estimation by Labati et al. [134] studied frame sequences representing walking subjects, where different walking directions and lighting conditions were used to challenge the algorithms. Features such as body height, body volume (estimated by ellipses passing through the silhouettes), body shape and walking direction were extracted to train a neural network towards body weight estimation. The mean errors ranged from −2.35 kg to 0.02 kg on a small proprietary dataset of 20 subjects.

C. Health and skin lesions

In addition to geometric body characteristics, another class of soft biometrics that relates to medical conditions is the

11http://store.sae.org/caesar/3dna
12www.ama-assn.org/go/obesity
13The use of BMI as a sole measure of obesity has been challenged in the medical literature.
14http://www.health.harvard.edu/newsletters/harvard_mens_health_watch/2012/February
15Extracting BMI or body weight also offers other advantages, e.g., in forensics, as these measures can be sensed at a distance
16http://www.cdc.gov/nchs/nhanes.htm
human skin (see Figure 9). The main motivation here is clear; early detection of malignant skin lesions is very critical towards initiating early and effective cure methods. For cases like skin cancer - which is the most common cancer type, with 3.5 million cases of basal and squamous cell skin cancer diagnosed in the United States each year\textsuperscript{17} - early detection can allow for a suitable cure and even prevent death. While such automated systems cannot fully replace medical professionals, it is often noted that visual examination by an expert can be difficult (deeper subsurface inspection is required), while biopsies can be invasive and time consuming. As a result there have been substantial efforts to develop novel noninvasive optical imaging techniques as a way to detect and analyze skin lesions \cite{222, 100}, with recent advances involving smart hand-held devices \cite{36}, \cite{263}. In addition to skin cancer, image-based methods can apply to other common skin diseases like chronic inflammatory psoriasis \cite{149}, \cite{57}, pressure ulcers \cite{141} and hyper-pigmentation evolution \cite{194}.  

b) Health and wrinkles: In addition to skin lesions, wrinkles can also reveal stress states that relate to the medical condition of an individual. Naturally wrinkles also carry discriminative clues, as was analyzed by Batool et al. \cite{17}, where a set of facial wrinkles extracted from images were treated as curve patterns towards finding similarity between subjects. In their work, different metrics based on Hausdorff distance and curve-to-curve correspondences were presented in an effort to quantify this similarity, and bipartite graph matching was employed for finding correspondences between curves from two patterns. The resulting recognition rates exceeded 65% at rank 1 and 90% at rank 4, suggesting that a set of wrinkle patterns may be unique to an individual.

D. Databases

There is currently a limited number of publicly available biometric datasets with annotated health-information. This is partly due to the sensitive nature of medical data which raises several ethical and privacy concerns especially when the subject identity is divulged. Hence, studies on related topics often involve private medical datasets, accessible only by associates. The MORPH-II and the CAESAR datasets are exceptions, providing images which include weight information.

V. MATERIAL, BEHAVIORAL, AND OTHER SOFT BIOMETRIC ATTRIBUTES

We next consider objects that may be associated with a person (e.g. accessories such as hats and scarves), traits that have a strong behavioral utility, as well as other soft biometric traits that have received substantial attention lately. Some of these attributes - related to eye lenses, eye glasses, hats, scarfs, clothes, etc. - may not be immediately associated with biometrics in the classical sense; nonetheless they remain of high interest in identification and security applications, often due to their conspicuous nature. Some of these objects and attributes may naturally be associated with different utilities at once. Take for example the color of the clothes. This attribute, while heavily related to subject occlusion, can also assist in human recognition, as shown in the work by Denman et al. \cite{46} who presented a video surveillance and tracking system that is based on the color of clothes, where specifically color histograms of the head, torso and legs were processed in order to re-identify individuals. The same attribute was used also by D’Angelo and Dugelay in \cite{37} for analyzing the color of clothes in a sports-soccer scenario, with the aim of preventing fights between fans of opposing teams (shirt color can be a strong indicator of the team which a subject supports).

A. Material characteristics

Unlike most other attributes, this class consists of characteristics that can in fact hinder the recognition process. These include face occluding characteristics, such as scarfs, caps, hats and eye glasses. Such occluding accessories have received attention primarily for their ability to hinder person recognition or even human detection, from both machine and human perspectives. The consideration of such attributes - in the context of face recognition, occlusion and occlusion detection - can be found in papers by Yoon and Kee \cite{284} and Wen et al. \cite{269}. In their work, occlusion detection is performed for specific facial features such as the mouth region, while some other works such as that by Lin and Liu \cite{146}, observe the face holistically, classifying it as occluded or not, by employing PCA and SVM. Below we review some of these material characteristics.

Scarfs are of particular importance because they are sometimes used by individuals in a crime spree. Min et al. \cite{166} presented a scarf detection algorithm based on PCA and SVM, and reported a detection rate of about 99%, on the ARFD database \cite{163} which features 300 scarf-occluded and 300 not-occluded faces. This work was later used by Min et al. \cite{168} towards face recognition, where Gabor wavelets, PCA and SVM were employed for occluded faces, while non-occluded facial parts were computed by block-based LBP. This work used the AR face database\textsuperscript{18}, and reported a recognition rate of 94.83\% for non-occluded faces, and 92.08\% for occluded faces.

In contrast to scarfs, which are mainly used to cover the lower part of the face, headgear can occlude the upper part of the face. Du et al. \cite{53} studied the occlusion of hard hats in construction sites, and studied face detection using the Viola Jones algorithm, also analyzing the motion and color in the upper face region in a dataset of 5 video sequences. Similarly, cap detection was performed by Min and Dugelay \cite{167} who utilized dynamic time warping (DTW) and agglomerative hierarchical clustering, to achieve an accuracy of 87.5\% on a private dataset of 10 recorded subjects with and without caps in 40 videos.

Similarly, eye glasses and contact lenses can also have an occluding effect - albeit lesser than that of the previous two characteristics - and can interfere with the performance of iris and periocular-based recognition systems. In this context, the impact of contact lenses on iris recognition was examined by

\textsuperscript{17}http://www.cancer.org/cancer/skincancer/index
\textsuperscript{18}http://www-sipl.technion.ac.il/new/DataBases/Alexi\%20Face\%20Database.htm
Baker et al. [13]. Their experiments suggested that even non-cosmetic prescription contact lenses can impact the matching accuracy of iris recognition. Subsequent research by Erdogan and Ross [56] established the possibility of automatically detecting such lenses in iris images. Glass detection in face images was studied by Jiang et al. in [114], [277], [272], with best results obtained using thermal images [101].

**Makeup** can also be viewed as an occluding characteristic, and has the potential to substantially alter the appearance of a face. Furthermore, makeup and cosmetics can be used to successfully alter or camouflage wrinkles, birth moles, scars and tattoos. The impact of makeup on face recognition has been studied by Dantcheva et al. [38], while Chen et al. [25] performed makeup detection, achieving a detection accuracy of approximately 91%.

**B. Behavioral soft biometric traits**

Behavioral attributes are a relatively new class of soft biometric traits that can assist in detecting different human behaviors. For example, the *accent* of a person can be gleaned from human speech which can also convey ancillary information such as the gender of the speaker. The associated gender estimation and accent recognition can be fused with classical speech or speaker recognition methods in order to improve recognition performance. In this context, Deshpande et al. [48] showed promising results, where, focusing on American and Indian accents in the English language, they analyzed the second and third formant frequencies of specific accent markers and classified those based on Gaussian mixture models.

Some other behavioral traits that can be automatically gleaned from video and images include facial expressions which also carry supplementary biometric information. Interesting work can be found in related survey papers [288], [162], [224]. In the specific context of soft biometrics, Kashyap et al. [118] studied video segments of individuals encoded by a facial action coding system (FACS) for facial expression quantification. Related experiments studied the intensity of a number of facial units (each facial unit represents movement of a specific facial muscle), and confirmed the suitability of facial asymmetry, as well as of action unit combinations, towards person identification.

**C. Scars, marks, tattoos**

Scars, marks and tattoos are increasingly employed towards person recognition in forensics and law enforcement (cf. Spaun [240]). Such traits can be particularly useful in describing wanted or missing people, or even unidentified bodies, and the traits are often bundled together due to the high degree of distinctiveness they can jointly provide. Scars and marks are mainly defined by the location in which they occur, while tattoos carry information in their position, color, content (such as human forms, animals, flags, plants, words, and other objects), as well as the way in which they are ‘imprinted’ (chemical, branded, cut) into the skin (see ANSI/NIST standard report19). Despite the rich information that is embedded in tattoos and marks, there is substantial ambiguity associated with their classification. To overcome this ambiguity, Lee et al [142] proposed a content-based image retrieval system that uses SIFT features, reporting rank-20 retrieval accuracies of 98.6% on high resolution images (Web-DB), and 77.2% accuracy on an operational database (MI-DB). Furthermore, Jain and Park [112] [188] used AAM to locate and segment primary facial features (e.g., eyes, nose, and mouth) and employed Laplacian-of-Gaussian (LoG) and morphological operators in order to detect facial marks. Their algorithm was tested on a subset of the FERET database consisting of 426 images of 213 subjects, and showed a marginal improvement in rank-1 identification accuracy over the commercial FaceVACS face recognition system. A similar evaluation was conducted on the Mugshot-dataset having 1,225 images of 671 subjects, achieving an improvement in accuracy from 91.88% to 93.14%. Along the same lines, Lin and Tang [145] used SIFT-based extraction and fusion of skin irregularities - such as facial marks - to improve face matching performance by about 5% on the XM2VTS and HRDB database.

Additionally, Pierrard and Vetter [193], in the setting of face identification, proposed a method to extract moles using normalized cross correlation (NCC) matching and a morphable model. Experiments were conducted on the FERET database, and the authors claimed that their method was pose- and lighting- invariant (albeit computationally expensive) due to the underlying 3D morphable model. Srinivas et al. [241] used facial marks to differentiate between identical twins, while Nurhudatiana et al. [181] studied relatively permanent pigmented or vascular skin marks (RPPVSM) found on the backs of 269 male subjects, and employed this soft biometric trait in forensics.

1) **Databases: XM2VTS database**20. This is a multi-modal face database, corresponding to 295 subjects, where for each subject there are four recordings taken over a period of four months, and where each recording contains a speaking head shot and a rotating head shot. Sets of data taken from this database are available, including high quality color images, 32 KHz 16-bit sound files, video sequences and a 3d Model. XM2VTS has been used for skin irregularity detection (e.g. facial marks). Mugshot [268] database. This database contains


20http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/
face images of 1500 subjects. Each subject has two images: a frontal view and a profile view. The photographs are provided from archives of the FBI. The Mugshot dataset has been used in facial marks studies. Web-DB database 21. The Web-downloaded Tattoo Image Database contains 4,323 tattoo images downloaded from the Web. Each image is of size 90 × 90. MI-DB database. The Michigan State Police Tattoo Database contains tattoo images from suspects and convicts, recorded over a period of ten years, by the Michigan police department. Tattoos are photographed when a suspect / convict is booked, and assigned a keyword as defined in the ANSI/NIST standard and then stored in the database along with other demographic information. There are 69,507 operational tattoo images, each of size 640 × 480.

VI. OPEN RESEARCH PROBLEMS

There are a number of open research problems in this expanding field.

a) Correlation between soft biometric traits: The issue of correlation is not well studied in the soft biometric literature. While Adjeroh et al. [2] investigated the correlation between geometric attributes of the human body, such a study has not been undertaken for other traits. Understanding the correlation structure between these traits will allow for the design of effective fusion schemes for combining them.

b) Defining the number of categories for a soft biometric trait: A main issue also relates to finding an efficient and robust way to create the corresponding categories for a particular trait. While for some traits, this categorization is straightforward, for other traits the division into categories can be complex and not immediately clear. For example gender recognition is easily viewed as a binary classification problem (male vs. female), while other traits such as age or race can be continuous and ambiguous in nature. The challenge is the fact that this classification must (at least partly) adhere to the semantic interpretation of traits (recall human compliance), as well as must account for the significant differences between actual traits and the human perceived version of these same traits.

In addition to the aforementioned challenges relating to the perceptual idiosyncrasies of humans, there are also challenges relating to the limitations of algorithmic and hardware resources. In short, a soft biometric system and its classification must take into consideration the resources of the system. For instance, from an image processing point of view, trait categorization can be limited by the employed estimation algorithms as well as the quality of the input images. As a result, while a certain camera-and-algorithm system may be able to robustly differentiate between many categories, another may come short of that, hence introducing the need for careful consideration of resources on the degree of refinement of categories.

c) Automatic quality assessment of extracted soft biometric traits: On a similar note, another challenge is to reliably extract these traits. For instance, from an image-processing point of view, the challenge of reliable feature extraction is often related to ambient factors such as lighting. Illumination variation is a significant problem which, in many cases - especially in surveillance videos - can result in massive variations in color. At the same time, soft biometrics offer substantial flexibility as far as acquisition is concerned. For example, traits such as hair color, skin color, height, weight, gait and different facial measurements, can be sensed (and estimated) from both frontal as well as side views of a subject. So the challenge is to balance the flexibility in acquisition afforded by soft biometric traits (e.g., under different pose and expression changes) with variations due to sensing (e.g., illumination) by optimizing the number of categories that can be used for a particular soft biometric trait (e.g., number of discrete colors for clothing). Automatic assessment of data quality may be necessary for this reason.

d) Subjective perception by humans: One wide open challenge is to design soft biometric matching systems that account for human variability in describing different traits. The human compliance attributed to soft biometrics, while useful in many ways, also introduces problems relating to the broad variability in syntax and language when it comes to describing different soft biometric traits (e.g., hair can be described as “red”, “auburn” to “brown” - see [30] for an interesting experiment on this). Furthermore, humans can be inaccurate when describing measurements [287], where descriptions can depend on the person’s own attributes and their own perception of population averages and variation [209]. Early literature to overcome this include methods that apply reference examples, or that apply comparative / pairwise ranking (cf. [209]). For example, determining whether a person is taller than someone else is easier and more accurate than an absolute estimation of the person’s height.

e) Fusion of soft biometrics with primary traits: Another topic is the incorporation of soft biometric traits into primary biometric systems. One could imagine a soft biometric system that quickly classifies and filters a database, and where this pruning is followed by a more robust but slower classical biometric search. In this setting, the challenge is to design and fuse the component systems in a way that satisfies the specific speed and reliability requirements of the overall system.

f) Statistical modeling and mathematical analysis of soft biometric systems: It is necessary to provide a mathematical analysis of the limits and capabilities of soft biometric systems. This need is accentuated by the fact that such systems are bound to get bigger and more complex (by considering numerous biometric and soft biometric traits), as well as more demanding. Systems may be required to process an abundance of information pertaining to large populations (e.g., over 1.2 Billion subjects in the UIDAI biometrics project22). Analysis of complex soft biometric systems is necessary; while simulations and empirical studies can often provide some insight into the limits of such systems, empirical approaches may not be able to accurately characterize the performance of large-scale systems with increased complexity.

One related problem that has remained largely unexplored is that of estimating the reliability of person recognition using

---

21http://www.tattoodesign.com/gallery/

22http://uidai.gov.in/
soft biometrics. Such analysis would have to consider the effect of the population statistics, as well as have access to mathematical models for the sensing and algorithmic capabilities of the system. This can help in automatically establishing the number of categories that are necessary for a soft biometric trait in order to guarantee a certain degree of performance. Further, it may be instructive to derive theoretical bounds on the discriminative and classification capabilities of a soft biometric system. Very preliminary work on this can be found in [40].

Along the same lines, it would be worthwhile to predict the expected improvement in matching accuracy and reduction in computational complexity when utilizing soft biometric traits as filters on large databases. Such analysis would be beneficial in applications such as time-constrained human identification in video surveillance systems. The analysis should take into account the implicit errors in soft biometric classification, since these errors would affect the overall identification performance. A very early analysis can be found in the work by Dantcheva et al. [41].

g) Spoofing of soft biometrics: The potential of using soft biometrics in several sensitive commercial and security applications, has to be balanced with understanding and minimizing the vulnerability of such traits to obfuscation (where an individual camouflages their soft biometric traits) and spoofing (where an individual attempts to look like another person in terms of their soft biometric traits). This is particularly important because the very nature of soft biometrics allows for easier spoofing than classical biometric traits; indeed one can easily imagine how gender, age and ethnicity can be disguised simply by styling and cosmetics. Similarly, eye color can be spoofed with color lenses, hair can be dyed, and skin color - and the overall facial appearance - can be modified by makeup. In this regard, the use of cosmetic products towards soft biometric spoofing would be a realistic threat. Early work on this can be found in [26], where the impact of makeup on gender- and age-estimation algorithms has been studied. Specifically the authors in this work discuss gender spoofing (see Figure 11) where male subjects attempt to look like females and vice versa, as well as discuss age alteration where female subjects attempt to look younger or older than they actually are. Experimental results with several gender- and age-estimation algorithms, suggested that gender and age estimation systems can be impacted by the application of facial makeup. Additional proof of the impact of makeup on age has been reported in [61], thus further accentuating the open challenge of designing algorithms that are less susceptible to spoofing and obfuscation techniques.

h) Ethical issues: In addition to the mainly algorithmic challenges mentioned above, there exist other challenges in the area of soft biometrics, including ethical challenges related to the use of skin color, ethnicity and gender in recognition systems (see [170]). Ethical issues might also rise as a result of the fact that soft biometrics can be covertly extracted from data collected for other purposes. At the same time, soft biometrics do a better job at preserving privacy, given that the stored signatures are often less descriptive than, say, a high resolution image of an iris. The challenge is to properly traverse this thin line between privacy enhancement and privacy infringement.

i) Privacy implications: In the context of biometrics, privacy refers to the assurance that the biometric data collected from an individual is not used to deduce any type of information about the individual [123], i.e., it should be used only for matching purposes. However, as discussed in this article, biometric data offer additional information about an individual which can be automatically deduced. This can be viewed as privacy leakage since an entity can learn additional information about a person (or population) from the stored data, without receiving authorization from the person for such a disclosure. Therefore, it is necessary to ensure that biometric data stored in a system are used only for the intended purpose and not for purposes that may result in a ‘function creep’. This has heightened the need to develop differential privacy constructs, where the biometric data can only reveal certain attributes (e.g., gender) while other attributes (e.g., ethnicity) are suppressed. The term “de-identification” has often been used in this context.

De-identification has become an essential keyword, since a large amount of video surveillance systems have been employed in public spaces. Such surveillance unintentionally invades the privacy of individuals captured in videos. To circumvent such a privacy invasion, subjects in the video can be de-identified. De-identification refers to the obfuscation of the identity of an individual without obscuring the performed action, see [3]. For maintaining the purpose of video surveillance, limited information related to subject and action has to be kept. Gender, ethnicity, body height, color of clothes and color of hair may adequately serve as remaining evidence. In [259] the authors present a privacy preserving video surveillance system, where subjects are de-identified and the remaining information includes the subject height, weight and color of clothes. Specifically, this approach uses body height, weight and clothes color to track a person inside a building under surveillance, and to display the position of this person on a monitor, in the form of a square placed within the map of the building. The square displayed in the monitor is properly colored; the upper part of the square has the color of the clothing of the upper body of the subject, while the lower
part of the square is colored to match that of the clothes on the lower part of the body.

On the other hand, de-identification of soft biometrics, while retaining the facial identity [183] can be useful for extending different levels of privacy to a face image in a central database.

Fig. 12. Perturbing gender information in face images by Othman and Ross [183]. The soft biometric attribute of a face (e.g., gender) is suppressed, while preserving the ability of the face matcher to recognize the individual. This ensures that the stored biometric data is not used for purposes beyond what was expressed during the time of data collection.

j) Standardization: ANSI/NIST-ITL 1-2011 (Data Format for the Interchange of Fingerprint Facial, & Other Biometric Information - Part 1) suggests “additional descriptive information” such as eye patch, clear glasses, dark glasses, head covering, scarf, moustache, beard, eye color, hair color. Further, the standard “American National Standard for Information Systems- Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo (SMT) Information” widely covers the use and classification of scars, marks and tattoos. Further standardization will be essential to balance privacy and utility.

VII. CONCLUSIONS

A review of the biometric literature suggests that research in soft biometrics is on the rise. This expansion is due to the large number of applications that can benefit from the extraction of soft biometric traits. Examples of such applications include security, surveillance, retrieval, and health care. In this article, we reviewed some of the methods that have been developed for extracting soft biometric attributes from image, video and audio data. We also introduced a taxonomy to organize the various soft biometric traits that have been presented in the biometric literature. Further, we discussed the benefits and limitations of different soft biometric approaches. Finally, we discussed some of the open research problems in soft biometrics research.

We believe that the main advantage of soft biometric traits lies in their ability to describe people from a human perspective. Therefore, it bridges the gap between machine and human descriptions of a person. We also believe that the performance of soft biometric systems needs to be considered carefully, as it can be affected significantly by several factors such as sensing and feature extraction. By carefully investigating the accuracy, reliability and distribution of different soft biometric traits, it is possible to judiciously use them in large-scale biometric systems. However, it is necessary to be cognizant of the privacy implications of using soft biometric traits. By balancing privacy with performance, it is likely that soft biometric traits will have a critical role to play in next generation identification systems.

REFERENCES


Antiza Dantcheva is a Marie Curie Post-Doctoral fellow at the STARS team, INRIA, France. Previously, she was a Post-Doctoral fellow at the Michigan State University and the West Virginia University, USA. She received her PhD in Signal and Image Processing in 2011 from Eurocom / Telecom ParisTech in France. She was the recipient of the Best Poster Award in ICB 2013 as well as the Tabula Rasa Spoofing Award in ICB 2013. Her research interests are in soft biometrics for security and commercial applications, where she has worked on retrieval of soft biometrics from images, as well as their corresponding analysis.

Petros Elia is an Assistant Professor with the Department of Mobile Communications at EURERCOM in Sophia Antipolis, France. He received the B.Sc. degree from the Illinois Institute of Technology, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Southern California (USC), Los Angeles, in 2001 and 2006 respectively. Since 2008 he has been at EURERCOM.

His latest research deals with the role of feedback and complexity in multiser communications, MIMO, cooperative and multiple access protocols and transceivers, complexity and isolation and connectivity in dense networks, queueing theory and cross-layer design, coding theory, information theoretic limits in cooperative communications, and surveillance networks. He is a Fullbright scholar, the co-recipient of the SPAWC-2011 best student paper award on the topic of reduced complexity bidirectional communication with limited feedback, and of the NEWCOM+ distinguished achievement award 2008-2011 for a sequence of publications on the topic of reduced complexity multimode communications in the presence of little or no feedback.

Arun Ross is an Associate Professor in the Department of Computer Science and Engineering at Michigan State University. He received the B.E. (Hons.) degree in Computer Science from the Birla Institute of Technology and Science, Pilani, India, in 1996, and the M.S. and Ph.D. degrees in Computer Science and Engineering from Michigan State University, East Lansing, in 1999 and 2003, respectively. Between 1996 and 1997, he was with the Design and Development Group of Tata Elxsi (India) Ltd., Bangalore, India. He also spent three summers (2000 - 2002) with the Imaging and Visualization Group of Siemens Corporate Research, Inc., Princeton, NJ, working on fingerprint recognition algorithms. His research interests include pattern recognition, classifier fusion, computer vision, and biometrics. He is the coauthor of the books “Introduction to Biometrics: A Textbook” and “Handbook of Multibiometrics”, and the co-editor of “Handbook of Biometrics”. Arun is a recipient of NSF's CAREER Award and was designated a Kavli Frontier Fellow by the National Academy of Sciences in 2006. He was an Associate Editor of the IEEE Transactions on Image Processing and the IEEE Transactions on Information Forensics and Security.