Abstract—This work introduces the novel idea of using a bag of facial soft biometrics for person verification and identification. The novel tool inherits the non-intrusiveness and computational efficiency of soft biometrics, which allow for fast and enrollment-free biometric analysis, even in the absence of consent and cooperation of the surveillance subject. In conjunction with the proposed system design and detection algorithms, we also proceed to shed some light on the statistical properties of different parameters that are pertinent to the proposed system, as well as provide insight on general design aspects in soft biometric systems, and different aspects regarding efficient resource allocation.

I. INTRODUCTION

Soft-biometric considerations were first analytically introduced by Bertillon (see for example [1]) who brought to the fore the idea of using biometric, morphological and anthropometric determinations for person identification. A great majority of Bertillon’s early descriptors/trait currently fall under the category of soft biometrics, as defined by Jain et al. [3] to be a set of characteristics that provide some biometric information, but are not able to individually authenticate the person, mainly due to lack of distinctiveness and permanence. Currently state-of-the-art systems on intrusion detection and security mechanism, include by default at least one biometric trait. Incorporating soft biometrics can increase the system reliability, and can impart substantial advantages: soft biometric features reveal biometrical information of individuals, they can be partly derived from the main biometric identifier, they are easily captured, and their acquisition does not require enrollment since training can be performed in advance on individuals out of the specific authentication group. As further noted in [2], soft biometrics are not expensive to compute, can be sensed at a distance in a crowded environment, do not require the cooperation of the surveillance subjects, and are especially useful in narrowing down the search within a group of candidate individuals for face recognition. More recent work on soft biometrics has been performed predominantly with the aim of pre-processing for face recognition. Recent work on using soft-biometrics for verification can be found in [11] which though, unlike the current work, focuses instead on body soft biometrics. B. Novel bag of facial soft biometrics, and related applications Motivated by the above plethora of utilities, the present work develops a tool for detection of facial soft biometric traits that emphasizes on the most obvious facial identifiers, primarily mentioned by humans, when portraying an unknown individual. The constructed tool is streamlined to achieve reliability of authentication at reduced complexity, and hence focuses on a specific set of simple yet robust soft-biometric traits, including hair color, eye color and skin color, as well as the existence of beard, moustache and glasses. In conjunction with the proposed system, the current work also presents some aspects for soft-biometric system design, as well as statistical characterization of relevant parameters, and an analysis of capabilities and limitations of general soft biometric systems. The constructed tool has several diverse applications, including:

- Expediting face recognition by search pruning
- Re-identifying a described criminal
- Searching surveillance videos for suspects

Other applications relate to the ability to match people based on their biometric-trait preferences, acquiring statistical properties of biometric identifiers of groups, avatar modelling based on the instantaneous facial characteristics (glasses, beard or different hair color), statistical sampling of audiences, and many others.

II. SOFT BIOMETRICS AUTHENTICATION SYSTEM

A. Main parameters: authentication group, traits, trait-instances, and categories

The setting of interest corresponds to the general scenario where, out of a large population, an authentication group is randomly extracted as a random set of N people, out of which one person is picked for authentication (and differentiation from all the other members of the authentication group). We note that this general scenario is compliant with both, the case of person verification as well as of identification. A general soft-biometric system employs detection that relates to λ soft-biometric traits (hair color, skin color, etc), where each trait

\[ i, \ i = 1, 2, \ldots, \lambda, \] is subdivided into \( \mu_i \), trait-instances, i.e., each
trait \( i \) can take one of \( \mu_i \) values. We henceforth denote as category to be any \( \lambda \)-tuple of different trait-instances, and we let \( \Phi = \{ \phi_i \}_{i=1}^\rho \) define a set of all \( \rho \) categories, i.e., the set of all \( \rho \) combinations of soft-biometric trait-instances. The number of categories \( \rho \), that the system is endowed with, is given by

\[
\rho = \prod_{i=1}^\lambda \mu_i \tag{1}
\]

We slightly abuse notation and henceforth say that a subject belongs in category \( \phi \) if his or her trait-instances are the \( \lambda \)-tuple corresponding to category \( \phi \). We here note that to have conclusive authentication of a subject, and subsequent differentiation from the other subjects of the authentication group, it must be the case that the subject does not belong in the same category as other members of the authentication group. Given a specific authentication group, the maximum-likelihood (ML) optimizing rule for detecting the most probable category in which a chosen subject belongs, is given by:

\[
\hat{\phi} = \arg\max_{\varphi \in \Phi} P(\varphi) \cdot P(y/\varphi), \tag{2}
\]

where \( y \) is the observation vector, \( P(\varphi) \) is the pdf of the set of categories over the given population (note \( \sum_{\varphi=1}^\rho P(\varphi) = 1 \)), and \( P(y/\varphi) \) the probability that \( y \) is observed, given that the subject belongs in category \( \varphi \).

B. Design aspects in soft-biometric systems

In designing a soft-biometric system, the overall choice of the traits and trait-instances, must take into consideration aspects as traditional limitations on estimation reliability, which is commonly a function of the sensor resolution, and of the capabilities of the image-processing part of detection. In addition to this traditional aspect, new concerns come into the picture when designing a soft-biometric system as of the size and statistics of the authentication group (such as the possible similarities that might exist between different subjects), as well as the statistical relationship between the authentication group and \( \Phi \). The interrelated nature of the above aspects brings to the fore different tradeoffs. Such tradeoffs include for example the fact that an increasing \( \mu_i \), and thus also an increasing \( \rho \), generally introduce a reduction in the reliability of detection, but can potentially result in a welcomed increase in the maximum authentication group size \( (N) \) that the system can accommodate for. It then becomes apparent that designing and analysing soft-biometric systems requires a deviation from traditional design and analysis of classical multi-biometric systems, towards considering the role of the above parameters, and their effect on the trade-offs and the overall system performance. This approach motivates the proposed soft-biometric system design described in Section III, as well as the subsequent system analysis of Section IV which also includes simulation evaluation of the proposed system in the interference limited setting of very high sensor resolution.

III. THE PROPOSED SOFT-BIOMETRIC SYSTEM

In accordance with the above design aspects, and in an effort to find a good balance between authentication-reliability and complexity, we here propose a soft-biometric system that focuses on simple and robust detection from a bounded set of traits and their trait-instances. In what follows, we will describe these basic elements, as well as the employed detection algorithms.

A. Chosen features of the proposed soft-biometric system

In the presented bag of facial soft biometric traits for human authentication, we allocate \( \lambda = 6 \) traits, which we choose and label as:

1. skin color
2. hair color
3. eye color
4. presence of beard
5. presence of moustache
6. presence of glasses.

In this setting we clearly assign \( \mu_4 = \mu_5 = \mu_6 = 2 \), corresponding to the binary nature of traits \( i = 4, 5, 6 \). On the other hand, the first three traits are of a continuous character (see Table I) and had to be categorized in consideration to the tradeoff between reliability of detection and trait importance. Towards this we chose to subdivide trait 1 (skin color) into \( \mu_1 = 3 \) instances and label them (following a recommendation provided by the ethical partner of an ongoing EU project, ActiBio [16] to avoid any assumptions about race or ethnicity based on skin color) as:

- \{ skin color type 1, skin color type 2, skin color 3 \} using numbers that increase from light to dark,
- to subdivide trait 2 (hair color) into \( \mu_2 = 8 \) instances
  - \{ light-blond, dark-blond, brown-, black-, red-, grey-, white-haired, and bald \}
and to subdivide trait 3 (eye color) into \( \mu_3 = 6 \) instances:
- \{blue-, green-, brown-, grey-, green-, black-eyed \}

As a result, the proposed system is endowed with the ability to detect

\[
\rho = \prod_{i=1}^\lambda \mu_i = 1152 \tag{3}
\]

distinct categories. For the sake of clarification, we note two simple examples of such categories in \( \Phi \):

- skin type 1, brown hair, blue eyes, no beard, no moustache, no glasses \( \epsilon \Phi \)
- skin type 3, black hair, black eyes, beard present, moustache present, glasses present \( \epsilon \Phi \)

Having described the basic parameters of the system, we proceed to specify basic aspects of the detection algorithms that were used for trait-instance identification.

B. Detection algorithms

The basic detector consisted of an automatic frontal face and facial features detector, which was partially drawn and modified from the algorithms in [ ]. Implementation of the different detection algorithms (see Table I for an overview) was performed using OpenCV [ ].

Pre-detection aspects: Before describing some basic aspects of the implemented trait detection algorithms, we note a
few pertinent issues that accompany detection. Regarding coordinate determination, we note that typical eye, skin and hair color detectors require knowledge of the eye coordinates, and similarly hair color detection requires knowledge of the coordinates for the upper head region. The precise computation and extraction of the characteristic regions of interest (ROI) (see Figure ?) for the eyes, mouth, nose and upper face coordinates, are essential for the subsequent detection. For higher accuracy, only in the training step, all coordinates were manually annotated. The considered ROIs for the selected soft biometric traits are illustrated in Figure ?. Identification of the ROI was generally followed by acquisition of the Hue, Saturation and Value (HSV) values. We note that the HSV color-space was chosen for being robust to illumination changes, as well as for the fact that it allows for a high degree of independence between the H, S, and V parameters, which renders the system capable to better handle light changes or shadows. Regarding outlier filtering, we used a simple threshold on the HSV values, based on the color standard-deviation σ. This was followed by HSV normalization. Regarding the statistical modelling, the probability density functions of skin, eye and hair color were computed using 3-component Gaussian mixture models whose parameters were estimated using the EM algorithm. Posterior probabilities over the observed HSV vectors for all trained trait instances were computed, followed by a majority vote decision on the detected trait instance.

1) **Eye Color Detection**: In this setting, careful and precise consideration of the ROI was particularly important, due to the regions inherently small size. The specific ROIs were retrieved using the circular Hough transformation, followed by pupil and reflection extraction, and then by acquisition of the HSV vectors. Regarding the training step, each eye color group was trained using images from the Caltech database [ ].

2) **Hair color detection**: The hair color ROI was chosen as a thin bar in the upper head region, as indicated in Figure ?. Training utilized 30 Feret images for each of the hair colors, where the classification was done manually.

3) **Skin color**: Detection of skin color was done in accordance to the eye coordinates which defined the ROI for the skin color detection to be the area underneath the ocular region. Training utilized 33 Feret images per skin color group, which were again annotated manually.

4) **Eye glasses detection**: Towards glasses detection, we considered that the areas around the eyes can be searched both for hints of glasses as well as for glass reflections. Challenges related to the fact that glass frames are either occasionally absent, or that they often resemble wrinkles, brows, shades and hair. Further challenge came from the fact that illumination variances hindered the appearance of reflections. These challenges were handled by placing emphasis on a ROI corresponding to the nose part of the glasses. The specific algorithm consisted of eye position detection, grey-level conversion, histogram equalization, extraction of region between the eyes, Laplacian edge detection and finally line detection.

5) **Beard and Moustache Detection**: In this case, face detection and feature-localization were followed by identification of the ROIs. These ROIs include the chin for the beard, and the area between the mouth and nose for the moustache. The color estimation was followed by outlier extraction and HSV normalization. The presence of beard and/or moustache was based on the Euclidean distance between the processed observation and skin- and hair-color information respectively. The presence of moustache was determined independently.

6) **Algorithmic dependencies**: As it is the case with general optimization problems, identification of algorithmic dependencies endows the system with increased reliability and computational efficiency. Towards this we refer to notable examples of such dependencies, such as that between skin color and glasses where, due to ROI overlapping, the presence of glasses has an impact on the perceived skin color. This information can be utilized and employed by modifying the ROI for skin color detection. Additionally we recall that skin color is employed in the detection of hair, beard and moustache, where furthermore the latter two traits are also contingent upon hair color. Figure ? sketches further dependencies of the mentioned facial soft biometric traits. Some of these dependencies were partly exploited in the process of detection.

Having described different aspects of the proposed soft-biometric system design, we proceed to shed some light on the statistical properties of the pertinent parameters and to provide some early results on the statistical characterization corresponding to these aspects of the proposed scheme, as well as simulations in the interference limited setting of very high sensor resolution.

We propose to utilize a set of facial soft biometric traits for human authentication, specifically \( \lambda = 6 \) traits: eye, hair and skin color, presence of beard, moustache and glasses. Those traits are subdivided in traits instances \( \mu \text{SkinColor}, \mu \text{HairColor}, \mu \text{EyeColor}, \mu \text{Beard}, \mu \text{Moustache}, \mu \text{Glasses} \). Furthermore we let \( \Phi = \{ \varphi \}_{i=1}^{\varphi} \) define a set of \( \rho \) combinations of soft biometric trait instances. Examples of different elements in \( \Phi \) include 'blue eyes, brown hair, skin 1, no glasses, no beard, no moustache' or 'black eyes, black hair, skin 3, glasses, beard, moustache'.

### C. Number of soft biometric traits \( \lambda \)

The number of soft biometrics traits \( \lambda \) affects the over all categories number \( \rho \) and thus the matching performance and diversity of the soft biometric system. The added cost for a rising \( \lambda \) is higher processing time, which is not an issue.
since soft biometrics are per definition of low complexity computation. We worked on $\lambda = 6$ soft biometric traits: eye, skin and hair color, presence of beard, moustache and glasses. While the binary trait instances moustache, beard and glasses have $\mu_{\text{Beard}} = \mu_{\text{Moustache}} = \mu_{\text{Glasses}} = 2$, the choice of $\mu_{\text{SkinColor}}, \mu_{\text{HairColor}}$, and $\mu_{\text{EyeColor}}$ depends on one hand on the sensor resolution and the equal error rate (EER) of the detection algorithms and on the other hand on the soft biometric traits distribution within the observed group of subjects. Again, like with a higher $\lambda$, a higher $\mu_{\Lambda}$ number rises $\rho$. For the colour based traits the following trait instances were established:  
$\mu_{\text{SkinColor}} = 3$: Skin color 1, Skin color 2 and Skin color 3 (from light to dark, for ethical reasons denoted just by numbers)  
$\mu_{\text{Haircolor}} = 8$: light blond, dark blond, brown, black, red, grey, white and bald  
$\mu_{\text{EyeColor}} = 6$: blue, green, brown, grey, green and black

D. Number of effective categories $\Phi_e$

1) Total number of categories: We present in 4. detection algorithms for $\lambda = 6$ soft biometric traits, resulting in a $\Phi$ set of $\rho = \prod_{k=1}^{\Lambda} \mu_{\lambda} = 1152$ attained different combination categories (see table II).

<table>
<thead>
<tr>
<th>Skin Color</th>
<th>Hair Color</th>
<th>Eye Presence</th>
<th>Glasses Presence</th>
<th>Beard Presence</th>
<th>Moustache Presence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

TABLE II

FACIAL SOFT BIOMETRIC TRAITS AND THEIR INSTANCES

It is to be noted that $\rho$ increases polynomially with $\mu$ and exponentially with $\lambda$. In a symmetric setting the number of categories result in $\rho = \mu^\lambda$. This means that after a certain point the number of categories $\rho$ increases faster with $\lambda$ than with $\mu$ (see Table 2). The threshold can be computed:

$$\frac{d\rho}{d\mu} < \frac{d\rho}{d\lambda}$$  \hspace{1cm} (4)  

$$\lambda < \mu \cdot \ln \mu$$  \hspace{1cm} (5)

which means that for a small categories number ($\rho_{16}$) the number of trait instances $\mu$ is more progressive than the number of soft biometric traits $\lambda$, but for $\rho_{16}$ the number of soft biometric traits $\lambda$ affects $\rho$ stronger to the over all number of categories. A soft biometrics system with 7 binary categories has more than 2.5 times the categories’ number than a system with 2 categories with each 7 soft biometric trait instances.

2) Empty categories: To ascertain the set of effective categories $\Phi_e$, we need to consider correlations between traits and exclude categories with zero probability. We chose as observation group the 646 subjects of the color Feret database. We wanted to study the empty categories, the correlation of the six chosen traits and furthermore the distribution of the chosen soft biometric traits $P(\varphi)$. Figure 2 depicts the empty categories in dependence of the number of observed subjects $N$.

Out of this first analysis it is evident that there are 1028 empty categories when all 646 subjects are considered. The most populated category is $\varphi_{64}$="black hair, black eyes, skin color 2", which includes 80 subjects. A remedy towards a better distribution and higher distinctiveness is to include more trait instances inside the categories with a high amount of subjects. In the case of $\varphi_{64}$ it is thinkable to distinguish between black shades for eyes and hair color. The more effective approach is to include more soft biometric traits, which would not only separate the subjects from same categories, but also, according to 2.1, will increase rapidly the categories’ number. The variety of facial soft biometrics is depicted in table III. The assumed trait instances yield to $\rho_{max} = 80621568$ categories. Even if only the smallest fracture of this number for effective categories would form a powerful tool for people recognition.

TABLE III

FACIAL SOFT BIOMETRIC TRAITS AND THEIR TRAITS

<table>
<thead>
<tr>
<th>Skin Color</th>
<th>Hair Color</th>
<th>Eye Color</th>
<th>Glasses Presence</th>
<th>Beard Presence</th>
<th>Moustache Presence</th>
<th>Age</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

We note that the correlations between soft biometric categories $\varphi$ result in a reduced importance of the single categories, namely in a lowered distinctiveness and thus have an impact on the detection. To compute the correlation between the remained categories we used Pearson’s product-moment coefficient used for indication of linear-correlation.

$$r_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$  \hspace{1cm} (6)

$$r_{\text{EyeColor, HairColor}} = -0.1964$$  \hspace{1cm} (7)  

$$r_{\text{HairColor, SkinColor}} = -0.1375$$  \hspace{1cm} (8)  

$$r_{\text{EyeColor, SkinColor}} = 0.370$$  \hspace{1cm} (9)  

$$r_{\text{Moustache, Beard}} = 0.6359$$  \hspace{1cm} (10)

The highest correlation exhibits the presence of moustache and beard, where the presence of moustache, when beard is present was asserted in 97.8%. The opposite statement of present beard, when moustache is given, does not hold. Figure 3 (a) depicts the eye and hair color distribution.

The study of the Feret database bore in addition disclosure about the $\rho$ probabilities $P(\varphi)$ of the N=646 subjects, with other words the distribution of the set of subjects within the all over categories.

In case of many empty and unbalanced categories due to the given distribution of the subjects in the observed group, the consideration of new soft biometric traits suggests itself. In our case we consider adding the soft biometric traits gender and age to the already implemented traits.

3) Number of subjects in an observed group: Here we want to find out, how to compute the number of subjects in an observed group, so that the soft biometrics authentication system is optimally employed. More subjects would result in
a higher EER, less the underexploitation of the system. Here again about the scenario In the case of identification the only way to unambiguously recognize a person is an exclusive category membership, each category can be only assigned to one subject. For verification or reidentification we have to consider the probability of collision inside the group. A collision is in this case the belonging of two or more subjects to the same category \( \phi_x \). While in no means conclusive an analysis relating to the birthday paradox can shed light on this problem. Under the assumption of an uniform distribution the probability \( p(N) \) for a collision in an observed group with \( N \) subjects is:

\[
p(N) = 1 - p(N) = 1 - \left(1 - \frac{1}{\rho}\right) \cdot \left(1 - \frac{2}{\rho}\right) \cdots \left(1 - \frac{N - 1}{\rho}\right)
\]

(11)

\[
p(N) = 1 - \frac{\rho!}{\rho^N (\rho - N)!}
\]

(12)

(13)

This is the best case scenario, with the assumption of uniformly distributed categories. To generalize the birthday paradox it can be written following:

\[
p(N; \rho) = \begin{cases} 
1 - \prod_{k=1}^{N-1} \left(1 - \frac{k}{\rho}\right) & N \leq \rho \\
1 & N > \rho
\end{cases}
\]

(14)

To compute the reverse result, of what is the size of an observed group \( N \) to obtain a collision probability \( p(N) \) of 50% following approximation is given:

\[
p(N; \rho) \approx 1 - e^{\frac{N(N-1)}{2\rho}} = 1 - \left(\frac{\rho - 1}{\rho}\right)^{\frac{N(N-1)}{2\rho}}
\]

(15)

\[
N(p; \rho) \approx \sqrt{2\rho \cdot \ln \left(\frac{1}{1-p}\right)}
\]

(16)

which results in 39 subjects in the case of 1152 uniformly distributed categories.

As expected the probability for a collision in the case of the assumed distribution is higher than in the uniform case, but as depicted in Figure 3 only to a certain extent.

With 1152 equally distributed categories the observed group could contain more than 700 subjects, so that 50% probability of a particular person to collide with another is obtained. It is to be noted that the birthday paradox gives us an idea, as to how often there might be interference "with respect to categories" between different subjects, but it doesn’t define the probability of interference. It describes the probability that interference exists without any guarantee that this interference will cause any trouble with the authentication of the chosen subject.

For the more realistic scenario of non uniformly distributed categories \( \varphi \) in Figure 3 we considered the example distribution of an online survey for 5142 subjects from Central Germany asked for their hair and eye color[]. This distribution allows \( \rho=35 \) categories, constructed by the trait instances \( \mu_{\text{Haircolor}} = 7 \) and \( \mu_{\text{Eyecolor}} = 5 \). We recall that the probability of being in category \( P(\phi_i) \), \( i=1,2,\ldots,\rho \). The probability that all \( N \) subjects are in different categories is the sum of all possible combinations of not colliding categories. A combination consists of the product of the product of noncolliding categories:

\[
p_{\text{non-collision}}(N) = \sum_{i\neq j \neq \cdots \neq z} P(\phi_i)P(\phi_j)\cdots P(\phi_z)
\]

(17)