

Soft Biometrics for Surveillance: An Overview

*D.A. Reid¹, S. Samangoei¹, C. Chen², M.S. Nixon¹,
and A. Ross²*

¹*School of Electronics and Computer Science,
University of Southampton, UK*

²*Lane Department of Computer Science and Electrical Engineering,
West Virginia University, USA*

Abstract

Biometrics is the science of automatically recognizing people based on physical or behavioral characteristics such as face, fingerprint, iris, hand, voice, gait, and signature. More recently, the use of soft biometric traits has been proposed to improve the performance of traditional biometric systems and allow identification based on human descriptions. Soft biometric traits include characteristics such as height, weight, body geometry, scars, marks, and tattoos (SMT), gender, etc. These traits offer several advantages over traditional biometric techniques. Soft biometric traits can be typically described using human understandable labels and measurements, allowing for retrieval and recognition solely based on verbal descriptions. Unlike many primary biometric traits, soft biometrics can be obtained at a distance without subject cooperation and from low quality video footage, making them ideal for use in surveillance applications. This chapter will introduce the current state of the art in the emerging field of soft biometrics.

sp005

Keywords: soft biometrics, surveillance, survey, fusion, identification

1. Introduction

s0005

Biometrics provides an automated method to identify people based on their physical or behavioral characteristics. Classical examples of biometric traits include fingerprints, irises, and faces which have been evaluated and demonstrated to be useful in several different applications ranging from laptop access to border control systems. Although these traits have been successfully incorporated in operational

p0005



sp010 Fig. 1. Surveillance frame displaying a few challenges in establishing identity of individuals in surveillance videos.¹

systems, there are several challenges that are yet to be addressed. For example, the utility of these traits significantly decreases when the input data is degraded or when the distance between the sensor and the subject increases. Thus, the use of alternate human attributes may be necessary to establish an individual's identity.

p0010 Soft biometric traits are physical or behavioral features which can be described by humans. Height, weight, hair color, and ethnicity are common examples of soft traits: they are not unique to the individual but can be aggregated to provide discriminative biometric signatures. Although these types of biometric traits have only been recently considered in biometrics, they have tremendous potential for human identification by enhancing the recognition performance of primary biometric traits.

p0015 Identification from a distance has become important due to the ever-increasing surveillance infrastructure that is being deployed in society. Primary biometric traits capable of identifying humans from a distance, viz., face and gait, are negatively impacted by the limited frame rate and low image resolution of most CCTV cameras. Figure 1 shows an example of a typical CCTV video frame. This frame shows a suspect in the murder of a Hamas commander in Dubai in 2010. The murder involved an 11-strong hit squad using fake European passports to enter the country. All the members of the hit squad used disguises including wigs and fake beards during the operation. From Fig. 1 it can be observed that although the image is at low resolution and the subjects' face and ocular features are occluded, a number of soft biometric features such as hair color, skin color, and body geometry can be deduced. Soft biometric traits can be extracted from very low quality data such as those generated by surveillance cameras. They also require limited cooperation from the subject and can be non-intrusively obtained, making them ideal in surveillance applications.

p0020 One of the main advantages of soft biometric traits is their relationship with conventional human descriptions (Samangoei et al., 2008); humans naturally use soft biometric traits to identify and describe each other. On the other hand,

¹Arabian Business <http://www.arabianbusiness.com/for-hamas-murder-suspects-40450.html>.

translating conventional primary biometric features into human descriptive forms may not always be possible. This is the semantic gap that exists between how machines and people recognize humans. Soft biometrics bridge this gap, allowing conversion between human descriptions and biometrics. Very often, in eyewitness reports, a physical description of a suspect may be available (e.g., “The perpetrator was a short male with brown hair”). An appropriate automation scheme can convert this description into a soft biometric feature set. Thus, by using soft biometrics, surveillance footage archives can be automatically searched based on a human description.

Biometric traits should exhibit limited variations across multiple observations of a subject and large variations across multiple subjects. The extent of these variations defines the discriminative ability of the trait and, hence, its identification potential. Soft biometrics, by definition, exhibit low variance across subjects and as such rely on statistical analysis to identify suitable combinations of traits and their application potential. This chapter examines the current state of the art in the emerging field of soft biometrics. Section 2 introduces the performance metrics used in biometrics. Section 3 discusses how soft traits can be used to improve the performance of classical biometric systems based on primary biometric traits. Using soft biometric traits to identify humans is reviewed in Section 4. Identifying gender from facial images is explored in Section 5. Finally, Section 6 explores some of the possible applications of soft biometrics. p0025

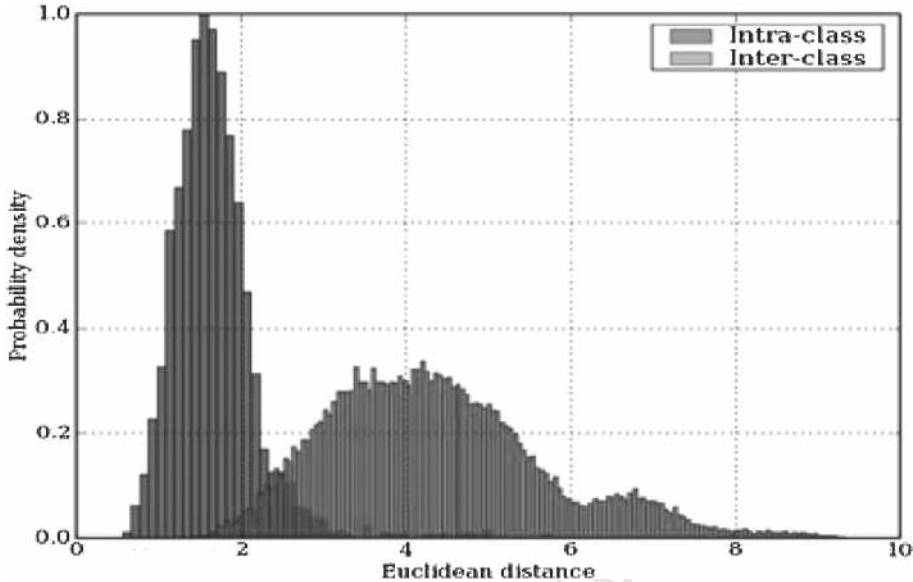
2. Performance metrics

s0010

The variation between multiple observations of an individual (intra-class variance) and the variation between subjects (inter-class variance) define the performance of a biometric system. If the intra-class variance of a biometric trait is low, then the trait is said to demonstrate permanence and repeatability. If the inter-class variance is high, then that biometric trait can be successfully used to distinguish between people. p0030

Once an unknown subject’s biometric signature (i.e., the feature set extracted from the biometric data) has been determined, the system must then identify the subject based on a database containing a labeled set of biometric signatures. The labels correspond to the identity of the subjects in the database. This is typically performed by calculating the similarity between the input biometric signature (known as the probe) and the signatures within the database (known as the gallery). Generally, each identity in the database will be ranked based on this similarity measure, producing an ordered list of identities. The rank-1 retrieval performance of the biometric system details the probability of the correct identity being first in this ordered list. p0035

While the above definition is for an *identification* system, it is possible for a biometric system to *verify* an individual’s identity. In such a scenario, the input biometric signature is labeled with an identity. Thus, the input signature is only compared against those biometric signatures in the database having the same identity label. If the similarity measure exceeds a threshold, then a “match” or an “accept” is said to have occurred; else, it is deemed to be a “non-match” or a “reject.” p0040



sp015 Fig. 2. Example intra/inter-class distributions.

p0045 When attempting to verify an identity, two errors can occur. A false accept (FA) occurs when the input signature is incorrectly “matched” against a different identity in the database. A false reject (FR) occurs when the input signature is incorrectly “rejected” when matched against the true identity in the database. These errors occur when (a) the biometric signatures of two subjects are very similar or (b) the biometric signature of a single subject varies with time. Typically, a threshold is chosen to define the similarity required for a “match.” This threshold determines the number of FA and FR errors. Figure 2 shows the intra-class and inter-class distances (i.e., a dissimilarity measure) of a sample biometric recognition algorithm. To achieve the best performance, the chosen threshold must minimize the FA and FR errors by separating the intra-class and inter-class distributions as optimally as possible. Receiver Operating Characteristic (ROC) curves show the trade-off between FA and FR errors at different threshold values; an example is shown in Fig. 3. The accuracy of two different ROC curves can be easily assessed by locating the point where the FA rate equals the FR rate: this is known as the equal error rate (EER). The EER provides a method of assessing the recognition performance of different biometric systems.

s0015 **3. Incorporating soft biometrics in a fusion framework**

p0050 Primary biometric traits such as face, fingerprints, and iris can suffer from noisy sensor data, non-universality, and lack of distinctiveness. Further, in certain applications, these traits may fail to achieve high recognition rates. Multimodal biometric systems (Ross et al., 2006) can solve these problems by combining multiple biometric traits, resulting in a biometric signature that is robust and more distinctive.

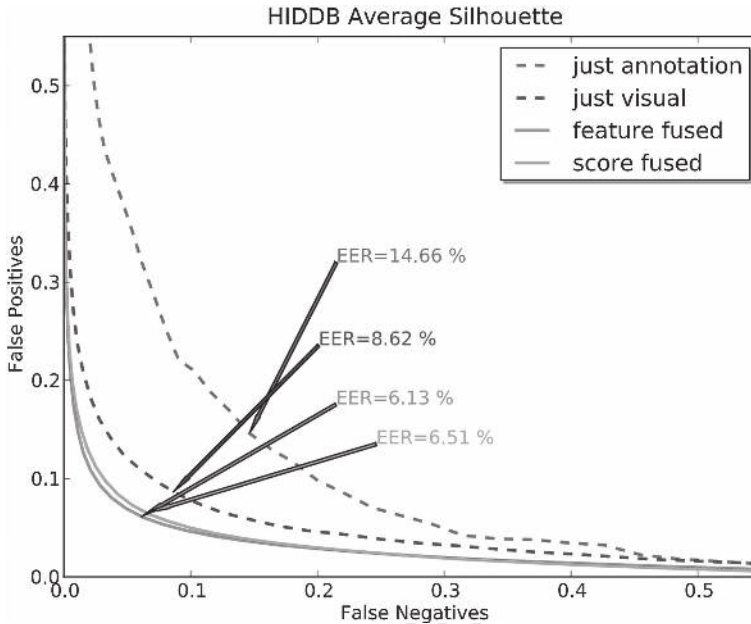


Fig. 3. Examples of ROC curves of four different techniques (Samangooci, 2010).

sp020

Multimodal systems offer improved performance, but the time taken to verify users can drastically increase thereby causing inconvenience to the subjects and reducing the throughput of the system. Soft biometric traits have been investigated to solve this problem (Jain et al., 2004b).

Jain et al. (2004a,b,c) experimented with the integration of soft biometrics in a biometric system. The primary biometric system compares the input biometric signature obtained from a user against each subject in the database. This determines the probability, $P(\omega_i|\mathbf{x}), i = 1, 2, \dots, n$, where n is the number of subjects within the database and $P(\omega_i|\mathbf{x})$ is the probability that the identity of the input primary feature vector \mathbf{x} is subject to ω_i . The secondary soft biometric system uses one or more soft traits to confirm the output of the primary biometric system. The authors used height, gender, and ethnicity for this purpose. Gender and ethnicity were automatically obtained from facial images using the technique discussed in (Lu and Jain, 2004). The height data was not available within the test data and, hence, a random height was assigned to each user. The soft biometric feature vector \mathbf{y} updates $P(\omega_i|\mathbf{x})$ resulting in $P(\omega_i|\mathbf{x}, \mathbf{y})$ that is calculated using Bayes' theorem:

p0055

$$P(\omega_i|\mathbf{x}, \mathbf{y}) = \frac{P(\mathbf{y}|\omega_i)P(\omega_i|\mathbf{x})}{\sum_{i=1}^n P(\mathbf{y}|\omega_i)P(\omega_i|\mathbf{x})}. \quad (1)$$

Experiments were performed on a 263-subject database using both multimodal and unimodal primary biometric systems. The authors first considered the fusion of a fingerprint-based unimodal biometric system with a single soft biometric trait (one of height, gender, and ethnicity). It was observed that fusion resulted

p0060

in improved accuracy compared to the fingerprint system. Height was seen to be more discriminative compared to gender and ethnicity, leading to a 2.5% increase in rank-1 retrieval accuracy—although this could be a result of the random generation of heights. Fusing all three soft biometric traits with fingerprints resulted in a 5% increase in rank-1 accuracy compared to using fingerprints alone. Finally, soft biometrics were used to improve a multimodal system featuring face and fingerprints. An improvement in rank-1 accuracy of 8% (over individual modalities) was observed when combining gender, height, and ethnicity information.

p0065 Jain et al.'s system showed the advantages of using soft biometric fusion in the context of both unimodal and multimodal biometric systems. Increasing the number of soft and primary biometric traits increases the uniqueness of a user's signature, leading to better discrimination between subjects. Ailisto et al. (2006) obtained similar success using body weight and fat measurements to improve fingerprint recognition, reducing the total error rate of 62 test subjects from 3.9% to 1.5%.

p0070 One concern, however, is the need for an automated technique to weight the soft biometric traits (Jain et al., 2004a). Marcialis et al. (2009) observed that certain soft biometric traits are only useful for a limited set of users. Their work only used soft biometric fusion when the user exhibited an uncommon soft trait thereby bypassing difficulties involved in weighting individual traits. It was assumed that the uncommon soft biometric feature could help in identifying a user from a set of possible candidate identities retrieved using primary biometric traits. An experiment fusing face with ethnicity and hair color was developed to verify this assumption. When fusing face and hair color (when uncommon), the EER on a database of 100 subjects fell from 6% to 4.5%. This paper clearly detailed the importance of uncommon traits and their ability to identify people. However, the use of hair color limits this technique to small databases and opens itself to spoof attacks.

p0075 The idea of utilizing uncommon traits was extended by Park and Jain (2010) and Jain and Park (2009) to identify people using facial marks. These marks include features such as scars, moles, freckles, acne, and wrinkles. Although not always permanent, most facial marks appear to be temporally invariant. The system proposed by the authors utilized facial marks, ethnicity, and gender to improve unimodal face recognition. One of the major advantages of facial marks is their utility (compared to automated facial matching) in courts of law since they are more descriptive and human understandable. In Park and Jain (2010), marks were characterized as salient localized regions on the face. Blob detectors based on the Laplacian of Gaussian are ideal for detecting such regions.

p0080 With respect to images, the second-order Laplace differential operator highlights changes in intensity. This is useful for detecting edges—which characterize the boundary between two regions of different intensity. The image is typically pre-smoothed to counter the operator's sensitivity to noise. Localized regions are identified as areas of either strong positive or negative responses for dark and light marks, respectively. Identified regions are highly dependent on the size of the mark and the size of the Gaussian kernel used for smoothing. To detect blobs of different sizes, the Laplacian operator is convolved at multiple scales by changing the size of the Gaussian kernel. In Park and Jain (2010), the resulting marks (blobs) were encoded into a 50-bin histogram and matched using the histogram intersection method. A commercial facial recognition system's EER was reduced from 3.85%

to 3.83% using facial marks. While this is a small reduction in EER, it demonstrates the utility of this soft biometric trait. Facial marks are especially beneficial when dealing with occluded or off-frontal face images. In their work, the authors artificially generated several examples of occluded face images, all of which were not recognized by the commercial facial recognition system. Upon using facial marks, the identities of subjects were correctly retrieved on average at rank 6. This demonstrates the benefit of utilizing uncommon traits and marks for human recognition in operational scenarios.

Thus, soft biometric fusion, when appropriately designed, can improve the accuracy of primary biometric systems with minimal inconvenience to the user. Soft traits can be used to either confirm results obtained from a classical biometric system or reduce the search space by filtering large databases. Soft biometric fusion is well suited for incorporation in security applications where speed and convenience are important. Further, in forensic applications, soft biometrics may help in confirming the identity of a subject. p0085

4. Human identification using soft biometrics

s0020

As stated in the introduction, soft biometric traits lack the distinctiveness and permanence to accurately identify a person. This definition remains true when dealing with single traits, but has been shown to be partially overcome when dealing with multiple soft biometric traits. Dantcheva et al. (2010) likens this to obtaining a single ridge of a fingerprint or a small section of the iris: these would not be unique enough to identify a subject. However, by agglomerating many such features a reasonably unique signature can be constructed. Soft traits have some advantages compared to classical unimodal and multimodal systems. p0090

One advantage of soft biometric systems is the bridging of the semantic gap between biometric traits and human descriptions. Soft biometric traits use human understandable descriptions (for example, height, hair color, and gender) and as a result can be naturally searched and understood. This also negates the requirement of obtaining biometric data of subjects before identification, allowing previously unencountered subjects to be identified using human descriptions. This presents exciting possibilities such as searching surveillance footage and databases based solely on an eyewitness's description. p0095

The two most popular traits for identification-at-a-distance are face (Zhao et al., 2003) and gait (Nixon and Carter, 2006). These can suffer from the poor sensor quality of most CCTV cameras. Low resolution can seriously impair facial recognition, and low frame rates (sometimes even time-lapse cameras) obscure the motion of the human body required for gait recognition. In contrast, several soft traits can often be obtained from very poor quality video or images. This has huge potential for immediate real-world use without upgrading the vast surveillance infrastructure. p0100

Ailisto et al. (2004) presented a soft biometric system aimed at addressing concerns of privacy, identity theft, and the obtrusive nature of previous biometric solutions. Their system used unobtrusive and privacy preserving soft traits, including height, weight, and body fat percentage. The system had applications in low- p0105

risk convenience scenarios where a relatively small number of people required identification, such as homes, small offices, and health clubs. Height, weight, and body fat were obtained from 62 subjects to mimic the target application environment. A monotonically decreasing penalty function was used to score the difference between stored measurements (templates) and the observed measurement (probe). A normal cumulative distribution was used to compute the scores:

$$\text{Score} = 2 \times (1 - P(Z \geq (|w_m - w_o|/c))). \quad (2)$$

Here, w_m is the model measurement, w_o is the observed measurement, c is a scaling factor based on an estimation of the variance of w_m , and $P(\cdot)$ is the standard normal cumulative distribution. Single modalities were shown to be very weak, with weight being the most distinctive resulting in a 11.4% total error rate (total false accepts and rejects). A combination of weight and height resulted in a 2.4% total error rate and the rank-5 retrieval accuracy was 100%. Using just three easy-to-obtain soft features allowed a database of 62 subjects to be sufficiently differentiated for the target application.

p0115 Samangoeei and Nixon (2010) developed a soft biometric system that focused on bridging the semantic gap by allowing conversions between a primary biometric trait and a set of 23 soft biometric traits (Table 1). This important step allowed subjects to be searched using either previously obtained biometric data or a human description of a previously unknown subject. Each soft biometric trait was described using multiple categorical labels, also referred to as semantic terms, allowing for easy human interpretation. A primary biometric signature was used to understand the subjective meaning behind each label. It is important to note that, in principle, any physical representation of the subject, ~~that is (a) a standard representation and (b) encompasses the person's soft attributes,~~ can be used. The use of gait signatures fulfills both these requirements, and permits the identification of subjects using gait signatures. By learning the relationships between the visual gait signature and the soft biometric features, the technique can be used to automatically label people based on their physical characteristics—thus converting gait signatures to human descriptions automatically.

p0120 Latent semantic analysis (LSA) (Deerwester et al., 1990), that is extensively used in document analysis, was employed to learn the structure between the two representations. The technique creates a vector space model which discovers a latent semantic structure based on the occurrences of features within documents. In this case, the documents are the subjects and the features are (a) the pixels within the biometric signature and (b) the semantic terms. By determining the co-occurrence between the features, we can detect the relationship between gait features and the semantic labels based on probability of occurrence.

p0125 The vector space is constructed using a set of subjects with semantically annotated gait signatures. Samangoeei et al. (2008) obtained soft biometric labels for each subject in the Soton gait database (Shutler et al., 2002). The feature vectors of the subjects in the training set were combined to create the term-document matrix \mathbf{O} . This occurrence matrix embodies the underlying structure between the gait signature and the soft biometric features. This structure is hidden under a majority of irrelevant occurrences between features. By removing the irrelevant relationships (noise) the

Table 1
 Semantic traits and corresponding terms

sp050

Trait	Terms
Arm length	[Very Short, Short, Average, Long, Very Long]
Arm thickness	[Very Thin, Thin, Average, Thick, Very Thick]
Chest	[Very Slim, Slim, Average, Large, Very Large]
Figure	[Very Small, Small, Average, Large, Very Large]
Height	[Very Short, Short, Average, Tall, Very Tall]
Hips	[Very Narrow, Narrow, Average, Broad, Very Broad]
Leg length	[Very Short, Short, Average, Long, Very Long]
Leg shape	[Very Straight, Straight, Average, Bow, Very Bowed]
Leg thickness	[Very Thin, Thin, Average, Thick, Very Thick]
Muscle build	[Very Lean, Lean, Average, Muscly, Very Muscly]
Proportions	[Average, Unusual]
Shoulder shape	[Very Square, Square, Average, Rounded, Very Rounded]
Weight	[Very Thin, Thin, Average, Fat, Very Fat]
Age	[Infant, Pre-Adolescence, Adolescence, Young Adult, Adult, Middle aged, Senior]
Ethnicity	[Other, European, Middle Eastern, Far Eastern, Black, Mixed]
Sex	[Female, Male]
Skin color	[White, Tanned, Oriental, Black]
Facial hair color	[None, Black, Brown, Blond, Red, Grey]
Facial hair length	[None, Stubble, Moustache, Goatee, Full Beard]
Hair color	[Black, Brown, Blond, Grey, Red, Dyed]
Hair length	[None, Shaven, Short, Medium, Long]
Neck length	[Very Short, Short, Average, Long, Very Long]
Neck thickness	[Very Thin, Thin, Average, Thick, Very Thick]

underlying semantic structure can be observed and an appropriate vector space can be constructed. Noise is removed by finding a rank-reduced approximation of the co-occurrence matrix. Singular Value Decomposition (SVD) is utilized to factorize the matrix allowing a rank reduced version to be determined. First, the matrix \mathbf{O} is factorized into three matrices such that:

$$\mathbf{O} = \mathbf{TSD}^T. \quad (3)$$

Here, \mathbf{T} and \mathbf{D} are orthogonal matrices and \mathbf{S} is a diagonal matrix. \mathbf{S} contains the singular values from \mathbf{O} and the matrices \mathbf{T} and \mathbf{D} contain the left and right singular vectors of \mathbf{O} . By reducing the rank of these matrices, the dimensionality of the problem is reduced, resulting in an approximation of \mathbf{O} . This approximation will ideally retain the most important information within \mathbf{O} ; this has been explored and proven in certain situations by Papadimitriou et al. (2000). The reduced rank k determines how many dimensions the data is condensed to and ultimately how much

information is removed. The diagonal matrix \mathbf{S} consists of r diagonal values that are ordered by size (and the corresponding row and column permutations applied to \mathbf{D} and \mathbf{T}). By removing the smallest singular values, the majority of the information is retained.

p0135 Once the noise has been removed and the semantic structure has been discovered, a k dimensional semantic vector space can be constructed. This space will consist of both features and subjects. The distance between these entities will correspond to their similarity or relevance. After the semantic space has been created, it can be queried to derive information about the structure. Comparisons can be made by finding the distance between the entities within the semantic space. Automatic soft biometric trait annotation is accomplished by inserting a test subject's gait signature into the semantic space. The closest and, hence, the most relevant soft biometric terms are then assigned to the subject resulting in a semantically annotated gait signature. The results from this approach were modestly successful showing an accuracy of 68% when determining semantic labels automatically from gait signatures (Reid and Nixon, 2010).

p0140 An interesting statistical analysis of these soft biometric traits was presented in (Samangoee, 2010). Each trait used to describe a person should be meaningful and provide additional information which differentiates the person from others. This property can be tested by determining the trait's ability to significantly separate the subjects within the database. If the subjects are not separated by a trait, then it could be said that the trait lacks any discriminative power and is not beneficial to the description (for the given set of subjects). To assess the discriminative power of each trait individually, the author used one-way ANOVA (analysis of variance).

$$\text{F-ratio} = \frac{\text{total between-group variance}}{\text{total within-group variance}}, \quad (4)$$

$$= \frac{\sum_i n_i (\bar{X}_i - \bar{X})^2 / (K - 1)}{\sum_{ij} (X_{ij} - \bar{X}_i)^2 / (N - K)}. \quad (5)$$

Here, X_{ij} represents the j th observation of the soft biometric of the i^{th} user and n_i denotes the number of observations of the i th subject. \bar{X}_i is the mean of the i th user's observations and \bar{X} is the mean across all subjects' observations. K represents the number of subjects while N represents the number of traits. Table 2 shows each trait's F-ratio, where a higher F-ratio indicates traits which are more successful at separating individuals.

p0150 It can be observed that "global" traits like gender, ethnicity, and skin color have more discriminative power than physical traits, like leg thickness. This is most likely due to the difficulty of labeling continuous physical traits compared to the categorical nature of the global traits. Traits like shoulder shape, proportions, and leg shape have been shown to be non-discriminative, thereby revealing their inability to distinguish between users. This important statistical analysis identifies the significance of each trait within a description and can be used to remove traits that do not contribute to additional information.

p0155 A statistical analysis of soft biometric systems utilizing categorical descriptions of physical traits was performed in Dantcheva et al. (2010, 2011) to determine the reliability of such a system in larger operational settings. When using categorical

Table 2
 F-ratio of several soft biometric traits based on the Soton gait database

sp055

Trait	F-ratio	Trait	F-ratio
Sex	383.70	Neck thickness	14.73
Skin color	149.44	Arm thickness	13.90
Ethnicity	96.10	Leg length	13.68
Hair length	79.05	Muscle build	12.85
Age	57.02	Leg thickness	11.61
Hair color	52.18	Hips	10.55
Facial hair length	25.72	Arm length	5.74
Height	25.14	Facial hair colour	5.61
Weight	20.75	Leg shape	3.25
Figure	20.69	Proportions	2.77
Chest	18.32	Shoulder shape	2.54
Neck length	15.57		

labels, it is important to consider the likelihood of a subject being indistinguishable from other subjects in the database: this is referred to as inter-subject interference (Dantcheva et al., 2010). Obviously the interference has a huge impact on the soft biometric system's performance and the number of traits recorded directly affects the probability of interference between subjects. The system developed within the project identified nine semantic traits, mainly focusing on facial soft biometrics. These include: the presence of a beard, mustache, and glasses, each containing two terms; the color of the skin, eye, and hair composed of three, six, and eight terms, respectively; body mass index consisting of four terms defined by population norms. Further, the color of clothing on the torso and legs were determined, each being labeled based on a set of eleven terms.

To understand the likelihood of interference, first the number of distinct combinations of the traits, p , must be calculated

$$p = \prod_{i=1}^{\lambda} \mu_i. \quad (6)$$

Here, λ is the number of soft biometric traits and μ is the number of terms associated with the corresponding trait. Based on the aforementioned nine traits $p = 557,568$. It must be noted that many of these combinations will not be realized when dealing with small databases, leaving a set of non-empty combinations F ranging from 0 to $\min(p, N)$ where N is the number of people in the database. The authors assumed that the distribution of subjects across the p possible combinations is uniform and that $N < p$. Let C_F be the total number of feature vectors which result in the set of F combinations, where F ranges between 1 and N . Then

$$C_F = \frac{p!}{(p - F)!} \frac{N!}{(N - F)!} F^{N - F}. \quad (7)$$

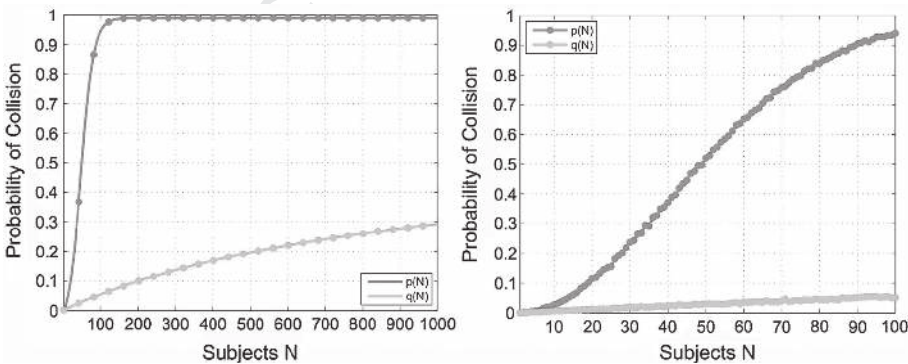
Here, the first term describes the total number of ways F combinations can be chosen to host the subjects. The second term describes the possible ways the F combinations can be filled with F subjects. The last term describes the number of ways the F combinations can be associated with the remaining $N - F$ subjects. Giving the distribution of $F, P(F)$, as

$$P(F) = \frac{C_F}{\sum_{i=1}^N C_i}, \quad (8)$$

where the mean of $F, E[F] = \sum_{F=1}^N FP(F)$. Finally the probability of error averaged over all subjects results as:

$$P(err) = 1 - \frac{E[F]}{N}. \quad (9)$$

p0175 Figure 4 shows the likelihood of interference occurring with N subjects where N ranges from 0 to 1000 subjects. The figure shows the probability of interference, $p(N)$, within a database of subjects and the probability of a randomly chosen subject from the database interfering with another subject (s), $q(N)$. Figure 4 clearly shows that with only 49 people a 50% chance of interference exists. This likelihood of interference can be reduced by increasing the uniqueness of each subject's trait signature. Increasing the amount of possible combinations of terms is one possible method for achieving this—only if the new term combinations further discriminate between the subjects. This can be achieved by either increasing the amount of traits or the detection of more terms per trait. In comparison, Samangooei et al. (2008)'s soft biometric system, featuring 23 traits, has 3.7×10^{15} possible combinations of semantic terms—decreasing the likelihood of interference. This important work clearly identified the need for maximizing the amount of term combinations and its effects on interference and ultimately the performance of the soft biometric system. Further statistical studies are required to identify the optimal number of term combinations for target application environments, taking into account the expected distributions across different soft traits.



sp025 Fig. 4. Interference probability in a N -sized population ranging from 0 to 1000, and a magnified version showing 0–100 (Dantcheva et al., 2010).

4.1. Imputation

s0025

Human physical traits and appearance inherently contain structure, features frequently co-occur or have fixed relationships with other features. This occurs either due to social aspects (long hair is common on females), genetics (black hair is common within people of Asian descent), or the morphology of the human body (taller people are more likely to have longer legs). Imputation techniques are a statistical approach used to predict missing variables. Using such techniques missing soft biometric features can be predicted utilizing the structure within human appearance. This structure offers a basis to improve the recognition of soft biometric traits and to make soft biometric systems more robust to missing traits or occluded visual features.

p0180

Reid and Nixon (2010) extended the soft biometric system presented in Samangoeei and Nixon (2010) and identified several scenarios where imputation would be critical. Occlusion is a major concern when applying soft biometric systems to unconstrained environments. Visual features can be concealed by scenery, the person's own body (self occlusion) or covariates such as bags, hats, and clothing. Structure can be utilized to predict the soft biometric terms normally obtained from occluded visual data, to provide a more complete description of the person. Similarly human descriptions obtained from witnesses, which could be used to identify people, are often unreliable and incomplete. Utilizing structure, these erroneous or missing

p0185

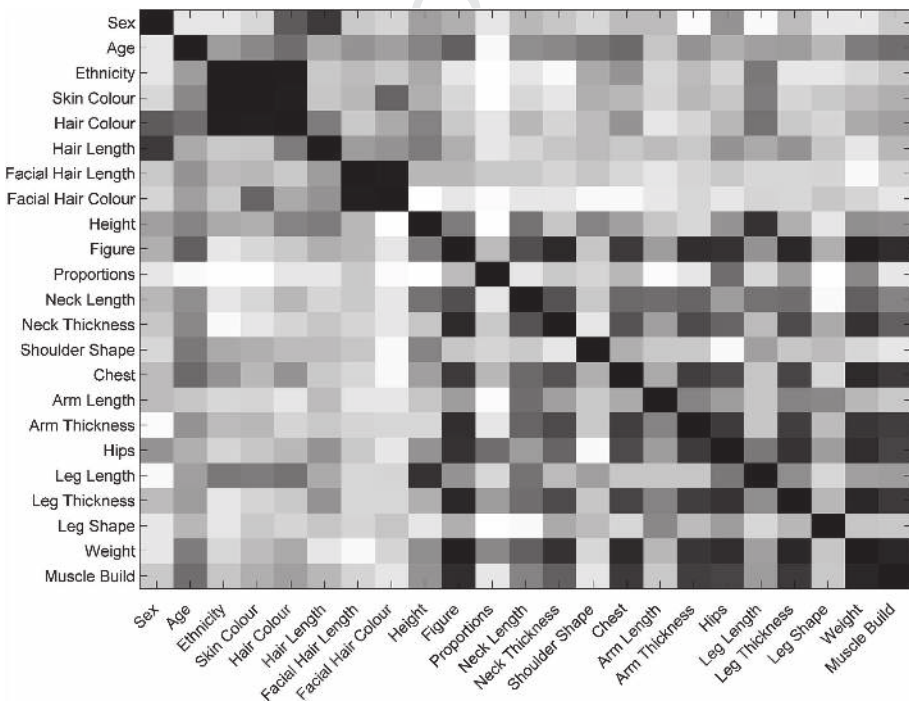


Fig. 5. Correlation (Pearson's r) between soft biometric features.

sp030

labels can be assessed and the most probable labels based on previous experience can be used to refine the description.

p0190 To verify that structure is present within human descriptions a correlation matrix was produced. This shows correlation between soft biometric traits based upon their occurrences within the Soton gait database (Shutler et al., 2002). Figure 5 shows the correlation matrix where darker cells represent more correlated features. It can be observed that large amounts of correlation occur within the 23 traits. The most prominent is that between skin color, hair color, and ethnicity, which can be seen in the top left corner. This relationship details the genetic likelihood that people from certain ethnic backgrounds are likely to have a certain skin color and hair color. Another interesting region within the figure is the lower right corner which details the relationship between physical attributes like weight, arm length, and leg thickness. Strong relationships exist between weight and thickness as well as height and length, producing high correlation. Some areas of the matrix contain little correlation, for example, the relation between weight and facial hair. The low correlation does not mean that there is no relationship between the features, only that it is not prevalent within the dataset.

p0195 It was first decided to utilize this known correlation to predict missing data. If the missing semantic trait is highly correlated with another trait, then it is beneficial to exploit this relationship to predict the missing term. The technique uses a similar method to the k nearest neighbor (k NN) classification technique. Each subject within the training set is compared to the subject containing the missing value. Typically this comparison involves finding the Euclidean distance between the two subjects' feature vectors. This has been modified to make use of the known correlation between traits. The similarity of a neighbor's trait is weighted by the correlation between that trait and the subject's missing trait. This favors neighbors with the same labels for traits with a strong relationship with the missing trait. The similarity between two subjects is determined as shown in Eq. (10) where X is the neighbor, Y is the subject with missing trait i . N is the total number of semantic traits and the matrix C contains the correlation between two semantic traits (values range from $[-1,1]$). The missing trait is predicted by taking the mode of the corresponding trait within the k nearest neighbors.

$$\text{Similarity}(X) = \frac{1}{N} \sum_{j=1}^N |C_{i,j}| (1 - |X_j - Y_j|). \quad (10)$$

p0200 Correlation cannot determine relationships between terms and traits, though it is adequate for determining linear relationships between traits. Table 3 shows the observations of skin color and hair color obtained from the Soton gait database. It can be observed that some terms, for example white skin, show more variance when compared to other terms from the same trait, for example oriental skin. By determining the correlation over all terms within a trait potentially strong ties between terms, for example oriental skin and black hair, are being lost. By observing a term's ability to predict a missing trait better accuracy can be achieved. It can be seen that ideal terms to predict hair color contain the least variance over their occurrences with hair color. This important property can be used to estimate the ability of a term to predict a missing trait and can be used to weight the similarity

Table 3
 Percentage of observations of hair colour and skin color

sp060

	Black	Blond	Brown	Grey	Red	Dyed
Black	1	0	0	0	0	0
Oriental	1	0	0	0	0	0
Tanned	0.86	0	0.14	0	0	0
White	0.01	0.22	0.7	0.03	0.01	0.03

when looking for the k nearest neighbors. Calculating the entropy of all the elements within a row provides a measure of certainty. This shows how successful the term is at predicting the missing trait. The entropy is used to weight neighbor's similarity. The similarity between two subjects is determined as shown in Eq. (12) where X is the neighbor, Y is the subject with missing trait i , which is composed of T_i terms. The matrix \mathbf{P} contains the percentages of observations (Table 3) between terms, such that $\mathbf{P}_{l,j}$ details the observations of term j with term l . N is the total amount of semantic terms within the semantic feature vector.

$$H(j) = 1 - \sum_{l=1}^{T_i} -\mathbf{P}_{j,l} \log \mathbf{P}_{j,l}, \quad (11)$$

$$\text{Similarity}(X) = \frac{1}{N} \sum_{j=1}^N H(j)(1 - |X_j - Y_j|). \quad (12)$$

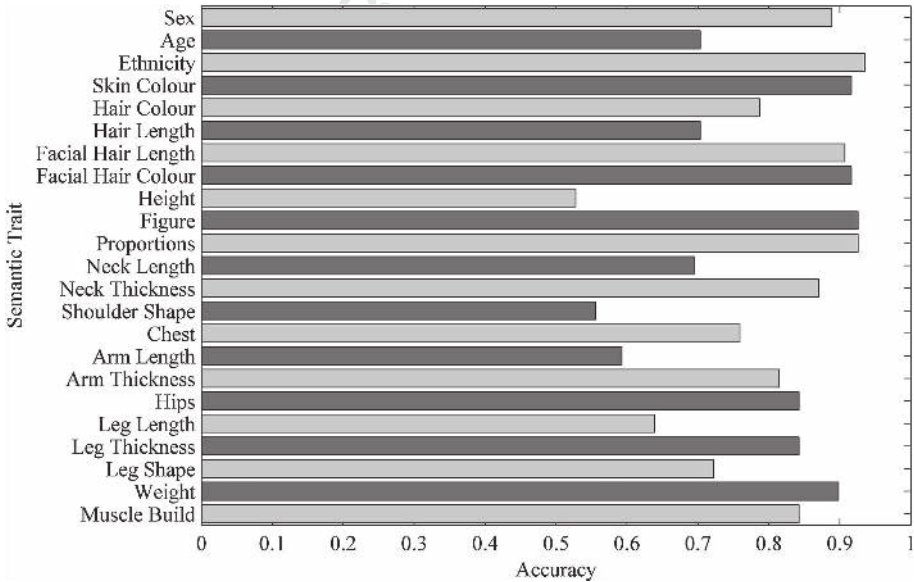


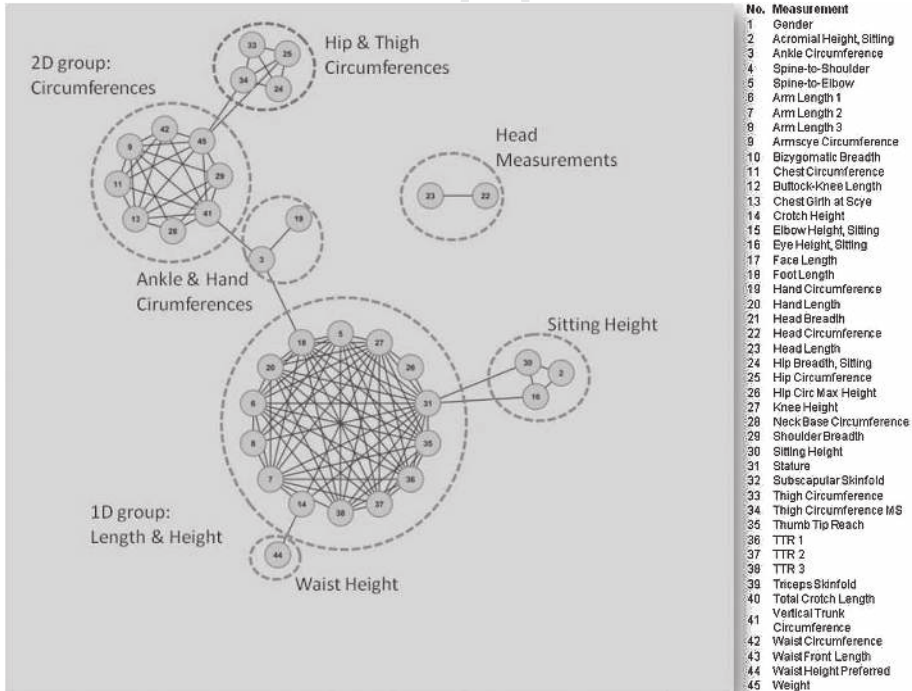
Fig. 6. Rebuilding soft data using entropy-based weighted k NN.

sp035

p0205 Figure 6 shows the accuracy of rebuilding each semantic trait using the entropy-based approach. It can be observed that the most successful traits are skin color and ethnicity, this is likely due to their strong correlation with other traits allowing accurate predictions of missing data.

p0210 **Adjeroh et al. (2010)** have also studied correlation and imputation in human appearance analysis. The most important difference with the previous study is the use of continuous data focusing on measurements of the human body. Data was gathered from the CAESAR anthropometric dataset which comprised of 45 human measurements or attributes for 2369 subjects. Like **Reid and Nixon (2010)**, the relationships between the human measurements were first assessed using the Pearson correlation coefficient. To visualize the correlation, a correlation graph was created—shown in Fig. 7. This graph shows connections between traits if the correlation was stronger than a threshold value. This clearly confirms the structure within human appearance and highlights clusters of traits with strong correlation. It can be observed that the measurements generally fall into two groups, both of which have physical meaning: the 2D group which contains circumferences of body parts and the 1D group containing lengths and heights. These clusters suggest that only a few measurements would have to be known to predict the majority of the other traits.

p0215 The metrology predictability network was developed to predict missing traits based on the most suitable subset of observed traits. Like the previous paper it



sp040 Fig. 7. The relations between measurements based on correlation (>0.81) (Adjeroh et al., 2010).

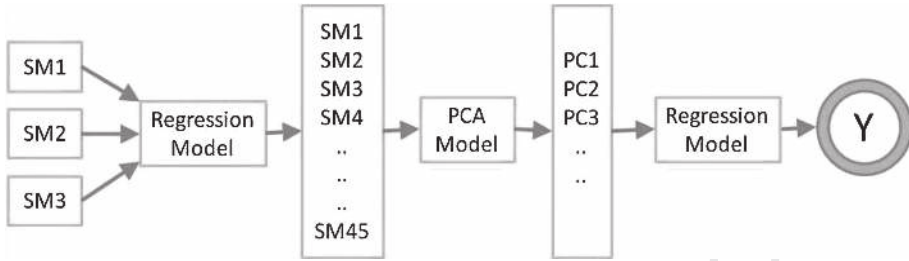


Fig. 8. Two-step prediction, using three seed measurements (Adjeroh et al., 2010).

sp045

appeared intuitive to use the most correlated traits to make the prediction. Using the correlation graph any nodes linked to the missing node are used in the prediction process. Traits which have been shown to accurately predict the missing trait are also considered. The expected error is assessed using multiple linear regression on training data from the CAESAR dataset

$$y = \mathbf{X}\beta + \epsilon, \quad (13)$$

where y is the response variable, \mathbf{X} is the input (predictor) variable matrix, β is the parameter vector, and ϵ is the error term. β is estimated using least squares such that:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T y. \quad (14)$$

Thirty one prediction models were constructed, each varying the order of the model, the number of variables, and the variable combinations. Using a training set of measurements, the error predicting a trait using a single prediction model is assessed. These errors are used to create a predictability graph (similar to Fig. 7) denoting the ability of a trait to predict another trait accurately, where edges denote errors which are below a threshold. p0175

The measurements obtained from the subject, called the seed measurements, are used to predict the missing traits. Some traits will be easy to predict due to their strong relationship with the seeds. These highly correlated traits are also used to predict traits with a weak relationship with the seeds. Principal component regression is used to predict the missing traits. An initial prediction is made using regression on the seed measurements, principal component analysis is used to reduce the measurements needed and then regression is applied to these features (Fig. 8). Experiments were conducted on the CAESAR dataset and 23 subjects from the CMU motion capture database. Four seed measurements were used—arm length, knee height, shoulder breadth, and standing height. Based on these seeds the remaining traits were predicted with an average mean absolute error of 0.041. Another experiment predicted all 41 measurements from just three seeds and used these measurements to predict the gender, resulting in a 88.9% correct gender classification rate. p0220

p0225 It has been shown that human appearance contains an inherent structure and just a few seed measurements are required to accurately predict the remaining features. This redundancy is vital when dealing with occlusion in visual data.

p0230 Structure is inherent within human appearance and is echoed in human descriptions and biometric representations. Imputation is crucial in operational settings where visual data is often occluded and human descriptions are often erroneous or incomplete. By utilizing structure within the soft traits, issues with view invariance and the subjective and unreliable nature of human descriptions can be addressed.

s0030 **5. Predicting gender from face images**

p0235 The problem of automated gender classification has attracted significant attention in many areas such as soft biometrics (Jain et al., 2004a; Mahalingam and Kambhamettu, 2011), demographic classification (Gutta et al., 1998; Yang and Ai, 2007), and human computer interaction. Gender classification is a fundamental task for human beings, as many social activities depend on the precise identification of gender. In the context of biometrics, gender is viewed as a soft biometric trait that can be used to index databases or enhance the recognition accuracy of primary traits such as face (Jain et al., 2004a; Mahalingam and Kambhamettu, 2011).

p0240 The study of automated gender classification from face images dates back to the early 1990s (Golomb et al., 1990) and has recently gained attention in the context of attribute-based face recognition (Kumar et al., 2009). The problem of gender classification from face images can be posed as a two-class problem in which the input face image is analyzed and assigned to one of the two classes: male or female. The problem of gender classification from face images ranges from the visible to infrared spectra (Ross and Chen, 2011; Chen and Ross, 2011) and from constrained (Moghaddam and Yang, 2002; Baluja and Rowley, 2007; Makinen and Raisamo, 2008) to unconstrained images (Prince and Aghajanian, 2009; Toews and Arbel, 2009; Shan, 2010; Hu et al., 2011).

p0245 Golomb et al. (1990) trained a back-propagation neural network (BPNN) to identify gender from human face images at a resolution of 30×30 pixels. An average classification rate of 91.9% on 90 exemplars was obtained compared to a human performance of 88.4%. Gutta et al. (1998) used hybrid classifiers consisting of an ensemble of radial basis function (RBF) networks and decision trees. The experiments were conducted on a collection of 3006 face images corresponding to 1009 subjects from the FERET database. The cross-validation results yielded an average accuracy of 96% for the gender classification task. Later on, Moghaddam and Yang (2002) utilized a support vector machine (SVM) for gender classification, based on low-resolution thumbnail face images of resolution 21×12 pixels. The average classification rate for fivefold cross-validation on 1755 FERET face images was 96.62% with the use of the Gaussian RBF kernel. Baluja and Rowley (2007) presented the use of an Adaboost classifier to identify the gender of a person from a low-resolution face image. The proposed system was extremely fast and yet comparable to the SVM-based classifier. They reported an accuracy over 93% on a dataset of 2409 FERET faces images with resolution 20×20 pixels.

Although the use of raw pixel values of face images in (Golomb et al., 1990; p0250
Gutta et al., 1998; Moghaddam and Yang, 2002; Baluja and Rowley, 2007) results
in very good gender classification performance, it is a simple feature representation
scheme that may not be robust enough in some complex scenarios involving pose
and illumination changes. Therefore, other types of feature extraction methods
have also been proposed. Yang and Ai (2007) used Local Binary Pattern (LBP)
histogram features (Ahonen et al., 2006) for gender feature representation, and
the real Adaboost algorithm to learn the best local features for classification.
Experiments were performed to predict the age, gender, and ethnicity information
from face images. Similar work was presented in Sun et al. (2006), where LBP features
and Adaboost classifier were combined to achieve better performance.

Other local-based descriptors have also been adopted in the work of gender p0255
classification. For example, Guo et al. (2009) evaluated gender classification results
based on LBP, histograms of oriented gradients (HOG), and Biologically Inspired
Features (BIF) with SVM as the classifier. It was demonstrated that gender
prediction was affected by age variations on a large database. Wang et al. (2010)
proposed a novel gender recognition method in terms of the Scale Invariant Feature
Transform (SIFT) descriptor and shape contexts. Again, Adaboost was used to select
features from face images to form the strong classifier. Other approaches utilized
gender-specific information, such as hair, to enhance gender prediction (Lian and
Lu, 2008), or genetic algorithms to select features encoding gender information
(Sun et al., 2002). Cao et al. (2011) proposed the use of facial metrology for gender
prediction.

All the aforementioned work mainly focus on datasets that were collected under p0260
well-constrained environments. Recently, gender classification on unconstrained
real-world face images has been attempted. Chen and Lin (2010) built a gender
classification system on real-world face images where the decision was based on the
region surrounding the detected face. Shan (2010) proposed to use the boosted LBP
features to represent face images and applied SVM to determine the gender on the
Labeled Faces in the Wild (LFW) dataset. An accuracy of 94.44% was obtained
on a dataset of 7,443 face images. Gallagher and Chen (2009) used social context
information in real-world group images to accomplish gender classification. They
argued that the structure information within the group provides meaningful context
for individuals. For example, men were more likely to stand in the corner of a
photograph than women. Gao and Ai (2009) targeted face gender classification on
consumer images in a multiethnic environment. To overcome the non-uniformity of
pose, expression, and illumination changes, they proposed a robust Active Shape
Model (ASM) to normalize the face texture. The consideration of ethnic factors can
help improve gender classification accuracy in a multiethnic environment. Recently,
Toews and Arbel (2009) extended gender classification to images involving arbitrary
viewpoints and occlusions. A viewpoint-invariant appearance model was learned for
the object class and a Bayesian classifier was trained to identify the model features
that indicate gender. In the work of (Prince and Aghajanian, 2009), images with
unconstrained pose, expression, and light conditions were considered for gender
classification based on additive logistic models.

With advancements in sensing technologies, thermal and near-infrared images p0265
are beginning to be used for face-related applications. For example, face recognition

in near-infrared (NIR) (Li et al., 2007; Zhang et al., 2010) and thermal (THM) (Socolinsky and Selinger, 2002) spectra has been motivated by the need to determine human identity in nighttime environments (Bourlai et al., 2011). Furthermore, changes in ambient illumination have lesser impact on face images acquired in these spectra than the visible spectrum. Current gender classification systems discussed in the literature have been designed for and evaluated on face images acquired in the visible spectrum. Little attention has been given to automatic gender classification from faces in the thermal or near-infrared spectrum. Ross and Chen (2011) attempted gender prediction on near-infrared face images. They later extended the work to gender prediction in the thermal spectrum (Chen and Ross, 2011). Compared to humans, machine learning-based approaches achieved better results on thermal face datasets.

p0270 In principle, a gender classification method can be divided into two components: (a) a feature extractor that extracts features from the face and (b) a feature classifier that assigns the extracted features into one of two classes—male or female. Feature extraction methods include the use of raw pixel face images (Moghaddam and Yang, 2002; Baluja and Rowley, 2007), Principle Component Analysis (PCA) (Balci and Atalay, 2002), Linear Discriminant Analysis (LDA) (Bekios Calfa et al., 2011), Independent Component Analysis (ICA) (Jain and Huang, 2004), LBP (Yang and Ai, 2007; Shan, 2010; Sun et al., 2006; Chen and Ross, 2011), and metrology (Cao et al., 2011). Some feature selection algorithms (Khan et al., 2005) have also been used to select gender specific features. Most gender classifiers are based on Neural Network (Golomb et al., 1990; Khan et al., 2005), Adaboost (Baluja and Rowley, 2007; Yang and Ai, 2007; Sun et al., 2006; Shan, 2010), Gaussian Process (GP) classifier (Kim et al., 2006), and SVM (Moghaddam and Yang, 2002; Guo et al., 2009; Chen and Ross, 2011). A systematic overview of methods for gender classification from face images in the visible and thermal spectra can be found in Makinen and Raisamo (2008) and Chen and Ross (2011), respectively.

p0275 An overview of gender classification methods and their accuracies is summarized in Table 4. It gives a brief insight into the different algorithms used in the past (Chen, 2011).

s0035 **6. Applications**

s0040 *6.1. Continuous authentication*

p0280 Most existing computers only authenticate users at the beginning of a session, leaving the system open to imposters until the user logs out. Continuous user authentication provides a method to continually confirm the identity of the user. Conventional biometric modalities such as face and fingerprint are either inconvenient for continuous operation or difficult to capture when the user is not explicitly interacting with the sensor. Soft biometrics offers a potential solution to this problem (Niinuma et al., 2010) by using features like the color of the user's clothes and facial skin.

p0285 When the user is initially authenticated using facial recognition and a password, soft biometric traits are obtained and recorded. Throughout the session the user is authenticated using these traits, without enforcing a strict posture or requiring

Table 4
 Overview of some studies on gender determination from a face image

sp065

Author & Year	Features	Classifier	Database, size	Accuracy (%)
Golomb et al. (1990)	Raw pixels	Neural network	Private, 90	91.9
Gutta et al. (1998)	Raw pixels	Hybrid classifier	FERET, 3006	96.0
Moghaddam and Yang (2002)	Raw pixels	SVM	FERET, 1755	96.62
Jain and Huang (2004)	ICA	LDA	FERET, 500	99.3
Khan et al. (2005)	PCA	Neural network	Private, 400	88.7
Sun et al. (2006)	LBP	Adaboost	FERET, 2000	95.75
Kim et al. (2006)	Raw pixels	GPC	AR, 515	97.0
Baluja and Rowley (2007)	Raw pixels	Adaboost	FERET, 2409	93.0
Yang and Ai (2007)	LBP	Adaboost	Private, 3540	96.32
Xia et al. (2008)	LBP, Gabor	SVM	CAS-PEAL, 10784	93.74
Gao and Ai (2009)	ASM	Adaboost	Private, 1300	92.89
Toews and Arbel (2009)	SIFT	Bayesian	FERET, 994	83.7
Shan (2010)	LBP	Adaboost	LFW, 7,443	94.44
Guo et al. (2009)	LBP, HOG, BIF	SVM	YGA, 8,000	89.28
Wang et al. (2010)	SIFT, context	Adaboost	FERET, 2409	95.0
Bekios Calfa et al. (2011)	PCA	LDA	FERET, 994	93.33
Chen and Ross (2011)	LBP	SVM	CBSR NIR, 3200	93.59
Cao et al. (2011)	Metrology	SVM	MUCT, 276	86.83
Hu et al. (2011)	Filter banks	SVM	Flickr images, 26700	90.1

constant verification. Facial recognition is also used periodically, when the biometric data is available, to guard against spoof attacks. Histograms of the various colors are gathered and the Bhattacharyya coefficient (Bhattacharyya, 1943) is used to calculate the similarity of two histograms, by measuring the amount of overlap.

In one experiment, a database of 20 subjects was constructed. Each subject was asked to perform six actions including turning their heads, leaning back in their chair, stretching arms, and walking away from the computer. The average false rejection and acceptance over all the recorded actions were 4.16% and 0%, respectively. Soft biometrics has been shown to provide secure continuous user authentication while being robust to the user's posture and not requiring manual registration of the soft biometric traits for each session.

p0290

6.2. Surveillance and re-identification

s0045

CCTV cameras have been widely introduced and accepted (Webster, 2009; Helten and Fischer, 2004). Their primary role within society is to assist in the fight against crime (Webster, 2009). This involves deterring and detecting crime, reducing the fear of crime, and to provide evidence when crime does occur. There has been considerable investment into the CCTV infrastructure (particularly in the UK) but currently these cameras (and the ensuing footage) are still generally monitored by humans. Due to the number of cameras within most cities, operators cannot monitor the data in intricate detail. This means looking for a single person can be time consuming and

p0295

prone to mistakes. Soft biometrics can potentially solve these problems by providing a method for searching surveillance footage using human descriptions.

p0300 Soft biometrics offer several benefits over other forms of identification-from-a-distance. Face recognition often requires good resolution images and gait recognition requires good frame rates. In comparison, certain soft biometric traits can be obtained from low-resolution and low-frame rate videos, and from an arbitrary viewpoint of the subject. Although not descriptive enough to identify people from large databases, they are adequate when dealing with smaller populations. The human compliant nature of soft biometric traits can also be exploited to allow searches based solely on a human description - possibly obtained from an eye witness. This allows for the use of soft biometrics when primary biometric identifiers cannot be obtained or when only a description of the person is available.

p0305 **Denman et al. (2009)** uses soft biometric traits to identify people using previous observations or human descriptions when traditional biometrics are unavailable. The height and color of the torso, legs, and head are used to model subjects. Identifying these three body components is done by first locating the person using background segmentation and then analyzing the color of moving pixels in each row. Large color differences are often found between the head, torso, and legs due to clothing that can be easily identified by examining color gradients. Average body proportions are used to identify the most likely color gradients representing the three desired regions. After the regions are located, a color histogram is recorded and the real-world height estimated. Heights are matched using average height and standard deviations, and color histograms are matched using the Bhattacharyya coefficient. The PETS 2006 surveillance database was used to test the system. This dataset features four cameras monitoring a train station: four recordings of 25 people were obtained. The system achieved an equal error rate of 6.1% when evaluated using the leave-one-out cross-validation scheme. These recordings included videos from two different viewpoints, demonstrating the view invariant nature of the selected soft traits. In comparison, primary biometric traits such as face, typically only work from one viewpoint.

p0310 Further work in soft biometrics has provided a technique to recognize subjects moving between multiple surveillance cameras in order to generate a rough framework for facial recognition (**Demirkus et al., 2010**). The technique uses gender, ethnicity, and session-based soft biometrics (skin color, upper and lower body clothing color, and hair color). Session-based soft biometrics are features which are reasonably constant for a short time period. These features, although not permanent, allow subjects to be identified when moving between different cameras. Once a person has been identified in the surveillance footage, the directional pose is determined. If the person is walking toward the camera, the face is analyzed to deduce ethnicity and gender, which is combined with the color-based traits that are extracted automatically. Gender and ethnicity are determined using support vector machines, and the color information is encoded using histograms. When a camera observes a new subject, their session-based features are compared to that of people previously observed by the camera network. If a match is found, the subject is given the same identity tag.

p0315 A custom low-resolution surveillance dataset was constructed featuring 100 subjects. An average correct classification rate of 60% and 83%, for gender and

ethnicity, respectively, was observed using just a resolution of 66×61 (pixels) facial images obtained from the video dataset. Gender and ethnicity were also used to partition the database of observed faces to speed up queries. The gender and ethnicity of the facial query were obtained and only faces featuring the same soft traits within the database were tested. The soft biometric partitioning reduced the time required for face recognition queries by almost a factor of 6 on a 600 subject database. Session-based soft biometrics are ideal for tracking people between cameras due to the speed in trait acquisition and their view invariant nature. Additional traits would allow for tracking in more crowded areas and would reduce the reliance on color, which is problematic if the cameras are not calibrated. Additional traits could also be used to partition the database further thereby reducing the time taken for primary biometric queries.

7. Conclusion

s0050

The use of soft biometric traits in automated human recognition systems has several benefits. It is, therefore, essential to carefully investigate issues related to its extraction and recognition capacity. p0320

Surveillance footage is generally of inferior quality and so traditional forms of identification-at-a-distance cannot be easily used. Soft biometrics offers a solution in this regard but lacks the distinctiveness that is expected of biometric traits. Applications with smaller populations are ideal, such as tracking people within a camera network or identifying people who are known to be located within a certain area. View invariance is a key aspect when working with surveillance footage—a successful technique must identify soft biometric features from any view of the subject. Emphasis must be placed on finding practical ways of obtaining view of invariant features (similar to [Denman et al. \(2009\)](#)) or developing methods to predict hidden features based on what can be observed. p0325

Bridging the semantic gap between human descriptions and biometrics will allow us to use natural “verbal” descriptions to search databases and surveillance footage. This is not possible using primary biometric modalities and, hence, opens up new applications and exciting prospects. Currently, biometric systems require an example of the subject before identifying them; this means unknown subjects cannot be automatically found. Human descriptions provide a way of finding unknown subjects using the description alone, which is critical in many crime-scene investigations. p0330

Human descriptions are often erroneous or incomplete due to the subjective nature of the descriptors and the psychological pressure experienced by eye witnesses. By utilizing the known pattern within human features and measurements, these descriptions could be corrected leading to a more informative and correct search criteria. The subjective nature of the descriptors could also be reduced by utilizing comparisons between people instead of absolute labels. Obtaining correct and detailed descriptions is vital when searching for suspects, and would allow soft biometric systems to function effectively. p0335

Soft biometrics typically require less computation and data quality compared to other forms of identification-from-a-distance, making them cheap and p0340

non-intrusive. The authors in Niinuma et al. (2010) clearly demonstrate the suitability of soft biometric traits for continuous user authentication. Determining applications which are suited to this form of biometric identification is essential for advancing the field. Statistical analysis underpins the performance of any biometric trait, and is vital for assessing the discriminability of feature measures as well as their recognition performance. By agglomerating multiple soft biometric features using fusion techniques, the recognition performance can be further enhanced leading to applicability in a wide variety of environments.

References


- Adjeroh, D., Cao, D., Piccirilli, M., Ross, A., 2010. Predictability and correlation in human metrology. In: Proceeding of IEEE International Workshop on Information Forensics and Security, pp. 1–6.
- Ahonen, T., Hadid, A., Pietikainen, M., 2006. Face description with local binary patterns: application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 28, 2037–2041.
- Ailisto, H., Lindholm, M., Makela, S.M., Vildjiounaite, E., 2004. Unobtrusive user identification with light biometrics. in: Proceeding of NordiCHI, pp. 327–330.
- Ailisto, H., Vildjiounaite, E., Lindholm, M., Makela, S., Peltola, J., 2006. Soft biometrics—combining body weight and fat measurements with fingerprint biometrics. *Pattern Recogn. Lett.* 27 (5), 325–334
- Balci, K., Atalay, V., 2002. PCA for gender estimation: which eigenvectors contribute? In: International Conference on Pattern Recognition, pp. 363–366.
- Baluja, S., Rowley, H.A., 2007. Boosting sex identification performance. *Int. J. Comput. Vis.* 71 (1), 111–119.
- Bekios Calfa, J., Buenaposada, J., Baumela, L., 2011. Revisiting linear discriminant techniques in gender recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (4), 858–864.
- Bhattacharyya, A., 1943. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.* 35, 99–109.
- Bourlai, T., Kalka, N.D., Cao, D., Decann, B., Jafri, Z., Nicolo, F., Whitelam, C., Zuo, J., Adjeroh, D.A., Cukic, B., Jeremy, D., Larry, H., Ross, A., Schmid, N., 2011. Ascertaining human identity in night environments. In: *Identified Video Sensor Networks*, pp. 471–478.
- Cao, D., Chen, C., Piccirilli, M., Adjeroh, D., Bourlai, T., Ross, A., 2011. Can facial metrology predict gender? In: *International Joint Conference on Biometrics*.
- Chen, C., 2011. Gender classification from facial images. West Virginia University, USA, Master's Thesis.
- Chen, D.-Y., Lin, K.-Y., 2010. Real-time gender recognition for uncontrolled environment of real-life images. In: *International Conference on Computer Vision Theory and Applications*, pp. 357–362.
- Chen, C., Ross, A., 2011. Evaluation of gender classification methods on thermal and near-infrared face images. In: *International Joint Conference on Biometrics*.
- Dantcheva, A., Dugelay, J., Elia, P., 2010. Soft biometrics systems: reliability and asymptotic bounds. In: *BTAS*, September, pp. 1–6.
- Dantcheva, A., Velardo, C., D'Angelo, A., Dugelay, J., 2011. Bag of soft biometrics for person identification: new trends and challenges. *Multimedia Tools Appl.* 739–777.
- Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R., 1990. Indexing by latent semantic analysis. *J. Am. Soc. Inform. Sci.* 41 (6), 391–407.
- Demirkus, M., Garg, K., Guler, S., 2010. Automated person categorization for video surveillance using soft biometrics. In: *Biometric Technology for Human Identification VII*.
- Denman, S., Fookes, C., Bialkowski, A., Sridharan, S., 2009. Soft-biometrics: unconstrained authentication in a surveillance environment. In: *Techniques and Applications, Digital Image Computing*, pp. 196–203.
- Gallagher, A.C., Chen, T., 2009. Understanding images of groups of people. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 256–263.
- Gao, W., Ai, H., 2009. Face gender classification on consumer images in a multiethnic environment. In: *IEEE International Conference on Biometrics*, pp. 169–178.

AU:1

- Golomb, B.A., Lawrence, D.T., Sejnowski, T.J., 1990. Sexnet: a neural network identifies sex from human faces. In: *Advances in Neural Information Processing Systems*, pp. 572–577.
- Guo, G., Dyer, C., Fu, Y., Huang, T., 2009. Is gender recognition affected by age? In: *International Conference on Computer Vision Workshops*, pp. 2032–2039.
- Gutta, S., Wechsler, H., Phillips, P.J., 1998. Gender and ethnic classification of face images. In: *International Conference on Face and Gesture Recognition*, pp. 194–199.
- Helten, F., Fischer, B., 2004. What do people think about CCTV? Findings from a Berlin survey. *Urban Eye* 13, 1–52.
- Hu, S.D., Jou, B., Jaech, A., Savvides, M., 2011. Fusion of region-based representations for gender identification. In: *International Joint Conference on Biometrics*.
- Jain, A., Huang, J., 2004. Integrating independent components and linear discriminant analysis for gender classification. In: *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 159–163.
- Jain, A.K., Park, U., 2009. Facial marks: soft biometric for face recognition. In: *IEEE International Conference on Image Processing*, November, pp. 37–40.
- Jain, A.K., Dass, S.C., Nandakumar, K., 2004a. Can soft biometric traits assist user recognition? In: *Proceedings of SPIE*, vol. 5404, pp. 561–572.
- Jain, A.K., Dass, S.C., Nandakumar, K., 2004b. Soft biometric traits for personal recognition systems. In: *International Conference on Biometric Authentication*, pp. 731–738.
- Jain, A.K., Nandakumar, K., Lu, X., Park, U., 2004c. Integrating faces, fingerprints, and soft biometric traits for user recognition. In: *BioAW*, vol. LNCS 3087, pp. 259–269.
- Khan, A., Majid, A., Mirza, A.M., 2005. Combination and optimization of classifiers in gender classification using genetic programming. *Int. J. Knowl.-Based Intell. Eng. Syst.* 9, 1–11.
- Kim, H.-C., Kim, D., Ghahramani, Z., Bang, S.Y., 2006. Appearance-based gender classification with gaussian processes. *Pattern Recogn. Lett.* 27, 618–626.
- Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K., 2009. Attribute and simile classifiers for face verification. In: *International Conference on Computer Vision*, pp. 365–372.
- Lian, X.-C., Lu, B.-L., 2008. Gender classification by combining facial and hair information. In: *Advances in Neuro-Information Processing*. Springer-Verlag, pp. 647–654.
- Li, S.Z., Chu, R., Liao, S., Zhang, L., 2007. Illumination invariant face recognition using near-infrared images. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 627–639.
- Lu, X., Jain, A.K., 2004. Ethnicity identification from face images. In: *Proceedings of SPIE*, vol. 5404, pp. 114–123.
- Mahalingam, G., Kambhmettu, C., 2011. Can discriminative cues aid face recognition across age? In: *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 206–212.
- Makinen, E., Raisamo, R., 2008. Evaluation of gender classification methods with automatically detected and aligned faces. *IEEE Trans. Pattern Anal. Mach. Intell.* 30 (3), 541–547.
- Marcialis, G.L., Roli, F., Muntoni, D., 2009. Group-specific face verification using soft biometrics. *J. Vis. Lang. Comput.* 20 (2), 101–109.
- Moghaddam, B., Yang, M.-H., 2002. Learning gender with support faces. *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (5), 707–711.
- Niinuma, K., Park, U., Jain, A.K., 2010. Soft biometric traits for continuous user authentication. *IEEE Trans. Inform. Forensics Secur.* 5 (4), 771–780.
- Nixon, M.S., Carter, J.N., 2006. Automatic Recognition by Gait. *Proc. IEEE* 94 (11), 2013–2024.
- Papadimitriou, Raghavan, P., Tamaki, H., Vempala, S., 2000. Latent Semantic indexing: a probabilistic analysis. *J. Comput. Syst. Sci.* 61 (2), 217–235.
- Park, U., Jain, A.K., 2010. Face Matching and Retrieval Using Soft Biometrics. *IEEE Trans. Inform. Forensics Secur.* 5 (3), 406–415.
- Prince, S.J.D., Aghajanian, J., 2009. Gender classification in uncontrolled settings using additive logistic models. In: *International Conference on Image Processing*, pp. 2557–2560.
- Reid, D.A., Nixon, M.S., 2010. Imputing human descriptions in semantic biometrics. In: *Proceedings of the 2nd ACM Workshop on Multimedia in Forensics, Security and Intelligence*, pp. 25–30.
- Ross, A., Chen, C., 2011. Can gender be predicted from near-infrared face images? In: *International Conference on Image Analysis and Recognition*, pp. 120–129.
- Ross, A.A., Nandakumar, K., Jain, A.K., 2006. *Handbook of Multibiometrics*. Springer.
- Samangooei, S., 2010. *Semantic biometrics*. Ph.D. Thesis, University of Southampton.

- Samangooei, S., Nixon, M.S., 2010. Performing content-based retrieval of humans using gait biometrics. *Multimedia Tools Appl.* 49 (1), 195–212.
- Samangooei, S., Guo, B., Nixon, M.S., 2008. The use of semantic human description as a soft biometric. In: 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, September, pp. 1–7.
- Shan, C., 2010. Gender classification on real-life faces. In: *Advanced Concepts for Intelligent Vision Systems*, pp. 323–331.
- Shutler, J., Grant, M., Nixon, M.S., Carter, J.N., 2002. On a large sequence-based human gait database. In: *Proceeding of RASC*, Springer Verlag, pp. 66–72.
- Socolinsky, D.A., Selinger, A., 2002. A comparative analysis of face recognition performance with visible and thermal infrared imagery. In: *International Conference on Pattern Recognition*, vol. 4, pp. 217–222.
- Sun, Z., Bebis, G., Yuan, X., Louis, S.J., 2002. Genetic feature subset selection for gender classification: a comparison study. In: *Proceedings of Applications of Computer Vision*, pp. 165–170.
- Sun, N., Zheng, W., Sun, C., Zou, C., Zhao, L., 2006. Gender classification based on boosting local binary pattern. In: *Advances in Neural Networks*, pp. 194–201.
- Toews, M., Arbel, T., 2009. Detection, localization, and sex classification of faces from arbitrary viewpoints and under occlusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 31 (9), 1567–1581.
- Wang, J., Li, J., Yau, W., Sung, E., 2010. Boosting dense SIFT descriptors and shape contexts of face images for gender recognition. In: *Computer Vision and Pattern Recognition Workshops*, pp. 96–102.
- Webster, C.W.W., 2009. CCTV policy in the UK: reconsidering the evidence base. *Surveill. Soc.* 6 (1), 10.
- Xia, B., Sun, H., Lu, B.-L., 2008. Multi-view gender classification based on local gabor binary mapping pattern and support vector machines. In: *International Joint Conference on Neural Networks*, pp. 3388–3395.
- Yang, Z., Ai, H., 2007. Demographic classification with local binary patterns. In: *International Conference on Biometrics*, pp. 464–473.
- Zhang, B., Zhang, L., Zhang, D., Shen, L., 2010. Directional binary code with application to PolyU near-infrared face database. *Pattern Recogn. Lett.* 31, 2337–2344.
- Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A., 2003. Face recognition: a literature survey. *ACM Comput. Surv. (CSUR)* 35 (4), 399–458.

AUTHOR QUERY FORM


	<p>Book: <i>Handbook of Statistics</i> Chapter: 13</p>	<p>Please e-mail your responses and any corrections to: E-mail: xxx@elsevier.com</p>
---	---	---

Dear Author,

Any queries or remarks that have arisen during the processing of your manuscript are listed below and are highlighted by flags in the proof. (AU indicates author queries; ED indicates editor queries; and TS/TY indicates typesetter queries.) Please check your proof carefully and answer all AU queries. Mark all corrections

and query answers at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file http://www.elsevier.com/framework_authors/tutorials/ePDF_voice_skin.swf) or compile them in a separate list, and tick off below to indicate that you have answered the query.

Please return your input as instructed by the project manager.

<p>Uncited references: References that occur in the reference list but are not cited in the text. Please position each reference in the text or delete it from the reference list.</p>	
<p>Missing references: References listed below were noted in the text but are missing from the reference list. Please make the reference list complete or remove the references from the text.</p>	
Location in Chapter	Query / remark
<p>AU:1, page 24 </p>	<p>Please provide complete details for reference Dantcheva et al. (2011).</p>