

Viewpoint Invariant Subject Retrieval via Soft Clothing Biometrics

Emad Sami Jaha^{1,2}

¹Faculty of Computing and Information Technology
King Abdulaziz University
Jeddah, Saudi Arabia
ejaha@kau.edu.sa

Mark S. Nixon²

²School of Electronics and Computer Science
University of Southampton
Southampton, United Kingdom
msn@ecs.soton.ac.uk

Abstract

As much information as possible should be used when identifying subjects in surveillance video due to the poor quality and resolution. So far, little attention has been paid to exploiting clothing as it has been considered unlikely to be a potential cue to identity. Clothing analysis could not only potentially improve recognition, but could also aid in subject re-identification. Further, we show here how clothing can aid recognition when there is a large change in viewpoint. Our study offers some important insights into the capability of clothing information in more realistic scenarios. We show how recognition can benefit from clothing analysis when the viewpoint changes with partial occlusion, unlike other approaches addressing soft biometrics from single viewpoint data images. This research presents how soft clothing biometrics can be used to achieve viewpoint invariant subject retrieval, given a verbal query description of the subject observed from a different viewpoint. We investigate the influence of the most correlated clothing traits when extracted from multiple viewpoints, and how they can lead to increased performance.

1. Introduction

Traditional biometrics have been widely and effectively deployed for people identification or authentication. There are still several challenges and limitations to be confronted such as lower resolution and increased distance between the camera/sensor and the captured subject, which can substantially decrease the utility of desired biometric traits. Any surveillance environment can be a typical example where such problems can easily and frequently occur.

Recently, several soft biometric traits have been introduced offering lower sensitivity to the aforementioned problems, and providing many advantages over the traditional biometrics for human identification and retrieval. The basic approach uses human vision wherein labellers describe human body features using human understandable labels and measurements, which in turn allow for recognition and retrieval using only verbal descriptions as the sole query [1, 2]. The features also allow



Figure 1: *Front and side samples used to obtain Training and Query annotations respectively*

prediction of other measurements as they have been observed to be correlated [3]. Indeed, soft traits are not unique to an individual but a discriminative biometric signature can be designed from their aggregation. Verbal identification can be used to retrieve subjects already enrolled in a database [4] and could be extended, in a more challenging application, for retrieval from video footage [1]. The capability of verbal retrieval from images and videos can pave the way for applications that can search surveillance data of a crime scene to match people to potential suspects described verbally by eyewitnesses. Soft biometric databases based on categorical labels can be incorporated with other biometrics to enhance recognition, such as integrating soft body traits with a gait signature [4]. Soft comparative labels have been demonstrated to be more successful in representing the slight differences between people in bodily descriptions [1]. For surveillance purposes, different forms of soft biometrics take place in a variety of applications and scenarios [1, 2, 5-7].

Clothing has been rarely adopted to represent soft biometric traits for an individual and have been considered unlikely to be a clue to identity [8]. Defining and utilizing a list of clothing attributes for various purposes has been a concern of several research approaches [6, 8-12]. There are few research studies associated with using clothing for biometric purposes [2, 6-11, 13]. The majority of existing research employs computer vision algorithms and machine learning techniques to extract and use clothing descriptions in applications including: online person recognition [13]; semantic attributes for re-identification [10, 11]; and utilizing attributes like clothing colour and style to improve the observation and retrieval in surveillance [2, 7].

Attribute based approaches appear more suited to

analysis of surveillance imagery, in case of low resolution and poor quality. In such imagery, whilst obtaining identifiable faces may be impossible, clothing appearance may become the main cue to identity [10]. Clothing appears to be more effective for short-term identification and re-identification, since people change clothes. However, even with images captured on different days, there remains some valid information to compare and establish identity, since clothes are often re-worn or a particular individual may prefer a specific clothing style or colour [14].

Soft clothing traits are a new form of soft biometrics that can be associated with biometric signatures to achieve successful subject retrieval [9]. This motivates more interest in the latent ability of clothing information in subject retrieval and identification. In case of viewpoint change and partial occlusion, even some soft biometrics may likely to be more vulnerable especially to annotator subjectivity and missing of information [9]. Viewpoint invariance is a challenging problem has been considered in most biometric modalities. Subject retrieval is deemed as viewpoint invariant, if remains invariant to any angle, which a subject is likely to be seen from [15], such as front, side, and up-angled views.

The major contribution of this work concerns viewpoint invariance: we adopt a scenario that aims to correctly retrieve a subject from a database, using a verbal clothing description provided by a different user describing an unseen same-subject’s image, captured from different viewpoint. For images like Fig. 1, we study the correlation between features extracted from side- and front-view images, which can be used to best achieve viewpoint invariant subject retrieval.

2. Semantic Clothing Attributes

The advantages of using human description is invariance to lighting and slight occlusion, as well as to positioning such as rotation, scale and translation. The descriptions are available at low resolution or when a subject is at a distance, and the sensor is somewhat immune to ageing (compared with a camera’s performance). Table 1 outlines a set of adopted clothing attributes along with their descriptive *categorical* and *comparative* group of labels defined to be suited to analysis of surveillance data, and used for describing a subject’s clothing. We emphasise the difference between *Categorical* and *Comparative* such that, *Categorical* labels are nameable descriptions used to describe specific attributes of an individual’s clothing, and associated with multiple clothing categories or styles. *Comparative* labels are nameable descriptions used to describe only relative attributes of an individual’s clothing compared with another individual’s clothing. Each of the 21 clothing attributes is described by a specified group of suitable categorical labels. Seven of the 21 attributes can be either categorical or relative, whereas the remaining 14 are unsuited for comparison because they are binary or multi-

class attributes that can be described using only categorical (absolute) labels. The seven relative comparable attributes are shown in bold in Table 1. A previous analysis using bipolar scales to define traits for whole-body descriptions [16] was considered to define appropriate bipolar label-scales for these relative attributes. For all binary attributes, such as ‘Belt presence’, a label ‘Don’t know’ was included to reflect uncertainty.

Table 1: Clothing attributes with categorical/comparative labels

Body zone	Attribute	Categorical and Comparative (Cmp.) Labels
Head	1. Head clothing category	[None, Hat, Scarf, Mask, Cap]
	2. Head coverage	[None, Slight, Fair, Most, All] Cmp: [Much Less, Less, Same, More, Much more]
	3. Face covered	[Yes, No, Don't know] Cmp: [Much Less, Less, Same, More, Much more]
	4. Hat	[Yes, No, Don't know]
Upper body	5. Upper body clothing category	[Jacket, Jumper, T-shirt, Shirt, Blouse, Sweater, Coat, Other]
	6. Neckline shape	[Strapless, V-shape, Round, Shirt collar, Don't know]
	7. Neckline size	[Very small, Small, Medium, Large, Very large] Cmp: [Much smaller, Smaller, Same, Larger, Much larger]
	8. Sleeve length	[Very short, Short, Medium, Long, Very long] Cmp: [Much shorter, Shorter, Same, Longer, Much longer]
Lower body	9. Lower body clothing category	[Trousers, Skirt, Dress]
	10. Shape	[Straight, Skinny, Wide, Tight, Loose]
	11. Leg length (of lower clothing)	[Very short, Short, Medium, Long, Very long] Cmp: [Much shorter, Shorter, Same, Longer, Much longer]
	12. Belt presence	[Yes, No, Don't know]
Foot	13. Shoes category	[Heels, Flip flops, Boot, Trainer, Shoe]
	14. Heel level	[Flat/low, Medium, High, Very high] Cmp: [Much lower, Lower, Same, Higher, Much higher]
Attached to body	15. Attached object category	[None, Bag, Gun, Object in hand, gloves]
	16. Bag (size)	[None, Side-bag, Cross-bag, Handbag, Backpack, Satchel] Cmp: [Much smaller, Smaller, Same, Larger, Much larger]
	17. Gun	[Yes, No, Don't know]
	18. Object in hand	[Yes, No, Don't know]
	19. Gloves	[Yes, No, Don't know]
General style	20. Style category	[Well-dressed, Business, Sporty, Fashionable, Casual, Nerd, Bibes, Hippy, Religious, Gangsta, Tramp, Other]
Permanent	21. Tattoos	[Yes, No, Don't know]

3. Clothing Description Database

A standard database adopted for this research is the Soton Gait Database [17], which comprises a subset of full-body front- and side-view still images of each subject in the database with soft biometrics’ labels available [1, 4]. The front-view images are used to collect clothing *Training* descriptions, and the side-view images are used to collect clothing *Query* descriptions. Fig. 1 shows a sample of both front and side view images of the same subject. This subset consists of 115 individuals with a total of 128 front view samples. In all experiments, each sample is handled as an independent individual; multiple samples of a single individual are considered as different and independent entities by which each entity represents that individual only

if wearing exactly the same clothing. Otherwise, it is considered as another entity, even though it belongs to the same individual. Here, each entity (i.e. sample in this dataset) is referred to as a *subject*. Note that about 90% of the subjects in the database wear largely similar clothing (jeans, T-shirt, etc.), which appears sufficiently challenging for this initial study. We designed a website used to collect clothing labels and comparisons.

3.1. Clothing Data Annotation for Training

For training purposes, a set of categorical and comparative labels were collected from 27 users via a suitable web-form for each labelling type. All 128 samples were labelled by multiple users, with one or more separate user annotations per subject describing the 21 categorical attributes. All subjects were compared using the seven relative (attributes 2, 3, 7, 8, 11, 14, and 16) by multiple users. Then to enrich the comparison data from the available number of collected comparisons, additional comparisons were inferred when two subjects were both compared with another same subject. A summary of collected *Training* data and the inferred comparisons data is shown in Table 2.

3.2. Query Data Annotation for Retrieval

We designed a new web-form similar to the one used for *Training* data collection. This form shows subjects' side-view images instead of the front-view, and was used to collect a total of 107 *Query* annotations about six months later from 11 user annotators. Some of annotators were new to the system and had not participated in any annotation before, whereas the others had already participated in *Training* annotation. It was ensured that annotators were requested to annotate subjects different from those they had already observed before. This, in addition to the long time separation between *Training* and *Query* annotations, aimed to ensure that collected *Query* annotations were totally new descriptions and not affected by prior experience or annotator bias. Table 2 summaries collected *Query* data.

Table 2: The number of collected and inferred annotation data

<i>Training</i> data summary	Collected	Inferred	Total
Total user annotations	316	N/A	316
Total user comparisons	317	556	873
Total attribute annotations	6636	N/A	6636
Total attribute comparisons	2219	3892	6111
<i>Query</i> data summary	Collected	Inferred	Total
Total user annotations	107	N/A	107
Total attribute annotations	2247	N/A	2247

3.3. Soft Clothing Biometrics

3.3.1 Categorical Clothing Traits (Cat-N)

All collected categorical annotations in *Training* dataset are used to compose for each subject a single categorical feature vector of clothing descriptions. This feature vector comprises 21 soft clothing traits deduced from a set of labels provided by multiple users describing the same

subject. Each soft clothing trait is computed as a normalised average-label for a single attribute. Hence, the resulting feature vectors for all subjects are gathered to construct a categorical-based gallery of soft biometric signatures, we refer to as *Cat-21*.

Table 3 presents an ordered list of Pearson's r value and p-value for correlated front-view and side-view categorical labels describing the same clothing attribute. A higher positive correlation value, closer or equal to 1, reflects a label's consistency and suggests that the attribute is invariant between different viewpoints. Our study on correlations between labels describing the same clothing attributes from the different viewpoints (front and side), can guide insight into which traits are more reliable and robust against viewpoint change. This in turn can lead to nominate a minimum number of effective categorical clothing traits to achieve enhanced invariant subject retrieval. As such, we reshape from *Cat-21* a second categorical gallery called *Cat-6*, comprising for each subject a feature vector composed of only a subset of the top six correlated traits (2, 4, 1, 9, 8, and 11), derived via ANOVA, shown in Table 3.

Table 3: Correlated *front* and *side* view labels per same-attribute

Soft clothing trait	P-value ($p \leq 0.05$)	Pearson's r
2. Head coverage	0.000	1.000
4. Hat	0.000	1.000
1. Head clothing category	0.000	1.000
9. Lower body clothing category	5.13E-45	0.922
8. Sleeve length	4.71E-27	0.819
11. Leg length (of lower clothing)	1.37E-17	0.709
5. Upper body clothing category	4.97E-16	0.684
13. Shoes category	1.13E-10	0.573
12. Belt presence	1.74E-06	0.443
7. Neckline size	0.001	0.331
20. Style category	0.030	0.210
6. Neckline shape	0.041	0.198
21. Tattoos	0.057	-0.184
10. Shape (of lower clothing)	0.107	0.157
15. Attached object category	0.396	-0.083
18. Object in hand	0.396	-0.083
14. Heel level	0.778	-0.028

3.3.2 Comparative Clothing Traits (Cmp)

All comparisons in *Training* dataset are used to derive for each subject a single feature vector of comparative clothing descriptions. Each value of the feature vector is supposed to be a relative measurement reflecting the degree of presence of a single comparable attribute. Thus, all collected comparative annotations need to be anchored, per attribute, to define these invariant relative measurements for each subject, and to arrange a list of ordered subjects with respect to a single attribute. To derive desired relative measurements and to achieve ordering for all subjects per attribute, a soft-margin Ranking SVM [18] is used, along with a supporting formulation of similarity constraints [19]. This is done to apply a pairwise technique based on learning a ranking function per attribute. Such learned ranking functions can be used not only to perceive the relative strength of attributes in a training sample, but also to predict the relative strength in a new test sample. Thus, for a set of attributes A , a ranking linear function r_a is learned for each attribute a such that:

$$r_a(x_i) = w_a^T x_i \quad (1)$$

where w_a is the coefficient of the ranking function r_a and x_i is a feature vector of attributes of i^{th} subject being ranked. A set of comparisons is rearranged into two groups to represent the pairwise relative constraints required to learn a ranking function. The first group consists of a set of dissimilarity comparisons D_a of ordered pairs so that $(i, j) \in D_a \Rightarrow i > j$, namely subject i has stronger presence of attribute a than j , whereas the second group comprises a set of similarity comparisons S_a of non-ordered pairs so that $(i, j) \in S_a \Rightarrow i \sim j$, namely i and j have similar presence of attribute a . D_a and S_a sets are then utilized to derive the w_a coefficients of r_a according to the following formulation:

$$\begin{aligned} & \text{minimise} \left(\frac{1}{2} \|w_a^T\|^2 + C \sum \xi_{ij}^2 \right) \\ & \text{subject to} \quad w_a^T(x_i - x_j) \geq 1 - \xi_{ij}; \quad \forall (i, j) \in D_a \\ & \quad |w_a^T(x_i - x_j)| \leq \xi_{ij}; \quad \forall (i, j) \in S_a \\ & \quad \xi_{ij} \geq 0 \end{aligned} \quad (2)$$

The degree of misclassification is measured by ξ_{ij} and the trade-off between maximising the margin and minimising the error (i.e. satisfying constraints) is denoted as C . The resulting optimal w_a function is able to enforce (explicitly) a desirable ordering for all training samples, in respect to a . A feature vector x_i is mapped using Eqn. (1) to a corresponding feature vector comprising a number of real-value relative measurements.

Table 4: Body and clothing galleries for retrieval experiments

Body-based biometrics	
<i>tradSoft</i>	4 categorical body soft biometrics (Age, Ethnicity, Sex, and Skin Colour)
<i>softBody</i>	17 categorical body soft biometrics including <i>tradSoft</i>
Combined clothing & body biometrics	
<i>tradCat-21</i>	21 categorical clothing traits combined with <i>tradSoft</i>
<i>softCat-21</i>	21 categorical clothing traits combined with <i>softBody</i>
<i>tradCat-6</i>	The best 6 categorical clothing traits with <i>tradSoft</i>
<i>softCat-6</i>	The best 6 categorical clothing traits with <i>softBody</i>
<i>tradCmp</i>	7 comparative clothing traits combined with <i>tradSoft</i>
<i>softCmp</i>	7 comparative clothing traits combined with <i>softBody</i>

Hence, from *Cat-21*, we form for each subject a feature vector using a subset of seven categorical traits, which are the only seven relative clothing attributes, shown in bold in Table 1 (attributes 2, 3, 7, 8, 11, 14, and 16). The resulting 7-value feature vectors of all subjects are used to learn seven optimal ranking functions for the seven relative attributes. The weighting of each function is derived using the formulation in Eqn. (2). The desirable per attribute ordering of all subjects is deduced from w . Then by Eqn. (1), each value of w is used to map each of 7-value feature vectors to a corresponding vector of seven relative measurements (i.e. comparative traits) describing a single subject. Eventually, a set of all mapped feature vectors is used to build a gallery of comparative clothing descriptions for all subjects, referred to as *Cmp*.

4. Viewpoint Invariant Retrieval by Clothing

In the context of this research, biometric based retrieval

is achieved such that, a query description of an unknown subject is used to probe a gallery by comparing and matching their biometric signature with all those signatures enrolled in the gallery. Subject retrieval can be distinguished from identification by which it is a task aims to identify an *unknown* subject using their biometric signature to find a match in those signatures enrolled in a database [1], in other words, it concerns the ability to generalise to unseen data. Therefore, retrieval tends to be more challenging and beneficial for extensive biometric applications. It is worth emphasising that a subject *Query* description is obtained from an annotator different from those annotators who provided the *Training* descriptions for the same subject. Moreover, a subject *Query* description is provided to describe an image different from the image of the same subject used for *Training* descriptions, and with different viewpoint (i.e. side view), where a number of clothing attributes are likely to be occluded or difficult to observe. It appears that such a context of subject retrieval provides extra challenges emulating some real-life cases in their conditions and complexity.

Soft body descriptions are available for Soton database, which were previously collected [4] such that, each of 115 individuals was labelled by multiple users' describing 23 soft bodily traits. Here, the performance of subject retrieval using these soft body biometrics alone is used as a baseline to be compared with their performance when supplemented by soft clothing biometrics. Soft clothing biometrics are examined in enhancing subject retrieval of two groups of soft body traits. The first group consist of only four (global) traditional soft traits (Age, Ethnicity, Sex, and Skin Colour), which is used to constrict the gallery (*tradSoft*) comprising all feature vectors of 4 normalised average-label. The second group comprises 17 soft body traits, observable from front and side views, including the former four traditional traits, which is used to constrict the gallery (*softBody*) comprising all feature vectors of 17 normalised average-labels.

Categorical and comparative soft clothing galleries (*Cat-21*, *Cat-6*, and *Cmp*) are used along with the two soft body galleries (*tradSoft*, and *softBody*) to compose six combined galleries enabling to examine and evaluate the capability and performance of soft clothing biometrics in subject retrieval. Table 4 lists and describes the produced soft biometric galleries of body traits and when combined with clothing traits. The combination between two galleries is performed by consistently concatenating every single feature vector from the first gallery with a corresponding feature vector from the second gallery, in which both feature vectors belong to the same subject. For body-based galleries, a set of 107 annotations are excluded from the obtained soft body data to be used as query-vectors. For each combined gallery, corresponding query-vectors are normalised then reshaped and concatenated according to the feature vectors in a tested gallery to enable matching

and retrieval. Based on the k nearest neighbour concept, the likelihood is estimated by the sum of Euclidean distance between each query-vector and all subject-vectors in a tested gallery, resulting in an ordered list of all subjects based on likelihood. Retrieval performance is evaluated and compared using several standard metrics: Cumulative Match Characteristic (CMC); Receiver Operator Characteristic (ROC) with its relevant measurements of the Area Under the Curve (AUC), and the Equal Error Rate (EER); and the Decidability Index (d'). With respect to all metrics, all approaches are ranked by overall performance.

4.1. Clothing with Traditional Soft Biometrics

Here, clothing descriptions are used to supplement the traditional soft traits (Age, Ethnicity, Sex, and Skin Colour) in subject retrieval, the experimental results detailed in Table 5 suggests that the retrieval performance is considerably and consistently enhanced by clothing traits in *tradCmp* and *tradCat-6* of clothing approaches. The retrieval accuracy of *tradSoft* at rank 1 is effectively enhanced by all clothing approaches up to 17% in the best case yielded by *tradCat-6*, and the average score up to rank 10 of *tradSoft* is also improved with a significant increase ranging from 9% to 28%. *tradCmp* is the best in terms of the decidability index and achieves the highest overall rank, indicating a better separation between genuine and imposter distributions and more successful discrimination between subjects. *tradCat-6* is the first to achieve 100% at rank 58. Fig. 2 illustrates that *tradCmp* followed by *tradCat-6* receive the best ROC scores with less errors outperforming *tradSoft* and *tradCat-21*. The performance of *tradCat-21* is low compared with the other clothing approaches, though it attains some recognition capability improving up to rank 34.

Fig. 3 shows two query examples of subject retrieval using *tradCmp* approach achieving the highest overall retrieval performance. In both examples, the top left corner image represents a side-view image used to derive the side-query descriptions. The remaining numbered front-view images represent the top k retrieved subjects from the test gallery, ordered based on their similarity to the query-description, where the query image and the correctly retrieved subject are bordered in yellow. In the first example on the left a query subject was correctly retrieved at rank 1, whereas in the second example on the right another subject was retrieved only at rank 7. In the second (right) example, it can be observed that all retrieved subjects are very similar in their clothing such as sleeve and leg length, and these similarities are correctly reflected by the match. It appears that in such a case, the strong similarities in comparative clothing traits and in the four traditional biometrics, may result in a confusion between such similar subjects. Despite of the desirable objective of retrieving the correct subject as the top match (rank 1), the retrieval of the correct subject within a small list (e.g. 10 subjects) appears reasonably successful and certainly will

be useful to narrow the search. So, retrieval may not always answer the question “is the top match correct?” but instead could answer “is the correct answer in the top k matches?”

Table 5: Performance of traditional soft traits and when add clothing

Approach	Top rank	AVG sum match scores up to rank		100% achieved at rank	EER	AUC	d'	overall rank
	=1	=10	=128					
<i>tradSoft</i>	0.176	0.347	0.872	73	0.183	0.127	1.882	3
<i>tradCat-21</i>	0.234	0.443	0.865	104	0.302	0.198	0.882	4
<i>tradCat-6</i>	0.346	0.617	0.944	58	0.144	0.090	1.223	2
<i>tradCmp</i>	0.308	0.607	0.946	69	0.113	0.077	2.006	1

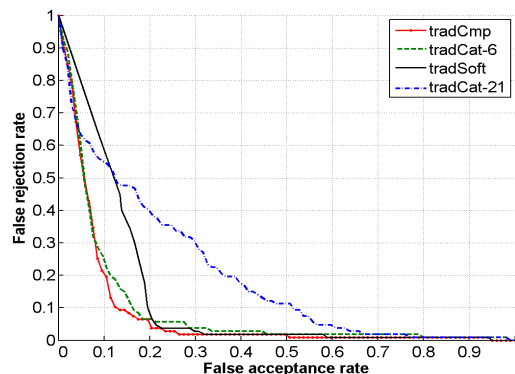


Figure 2: ROC of traditional soft biometrics and when add clothing

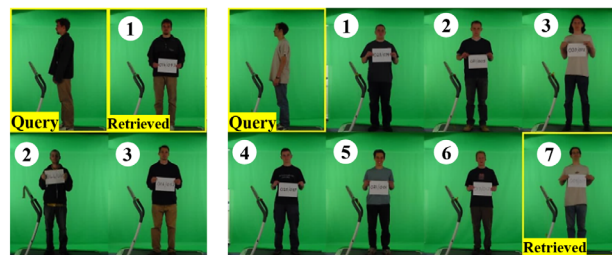


Figure 3: Retrieval query example using *tradCmp*, left: a subject correctly retrieved at rank 1, right: a subject retrieved but at rank 7

4.2. Clothing with Bodily Soft Biometrics

Clothing descriptions are reused here to supplement the front/side observable 17 soft body descriptions including the four traditional traits (see table 4). Table 6 reports metric scores of all approaches, and Fig. 4 compares the CMC performance (up to rank 40) of these approaches. The clothing approaches *softCat-6* and *softCmp* respectively provide the highest performance, while *softCat-6* gains the best scores in all evaluation metrics but d' and EER, and outperforms all approaches. Further performance analysis of *softCat-6* is shown in Fig. 5-(a) which demonstrates estimated distributions of match scores in terms of intra-class (genuine) and inter-class (imposter), and 5-(b) represents the EER and trade-off between two errors FAR and FRR along a set of thresholds varying from 0 to 1. The rank 1 retrieval of *softBody* is obviously enhanced when adding clothing from about 86% to 92% by *softCmp*, and to 94% by *softCat-6*, given in Table 6. As such, the clothing analysis can effectively augment soft body descriptions. This performance is reflected in the class distributions in

Fig. 5-(a) and in the ROC analysis in Fig. 5-(b), which confirm the potency *softCat-6* labels. Since *Query* descriptions have been acquired using side-view images, some items are difficult to observe or occluded such as neckline shape or size and belt presence. As *softCat-21* consists of all clothing traits including those affected traits, this can produce, for affected traits, inconsistent descriptions and undesirable increase in within-class variance. That appears to be a reason for the low performance compared with the other approaches. Another reason could be the noise caused by adding a large number of clothing traits to another large number of body traits.

Table 6: Performance of body soft biometrics and when add clothing

Approach	Top rank =1	AVG sum match scores up to rank		100% achieved at rank	EER	AUC	d'	overall rank
		=10	=128					
<i>softBody</i>	0.868	0.962	0.996	30	0.064	0.015	3.442	3
<i>softCat-21</i>	0.729	0.899	0.987	39	0.108	0.042	1.802	4
<i>softCat-6</i>	0.943	0.985	0.998	27	0.070	0.012	3.117	1
<i>softCmp</i>	0.916	0.981	0.997	30	0.070	0.014	3.406	2

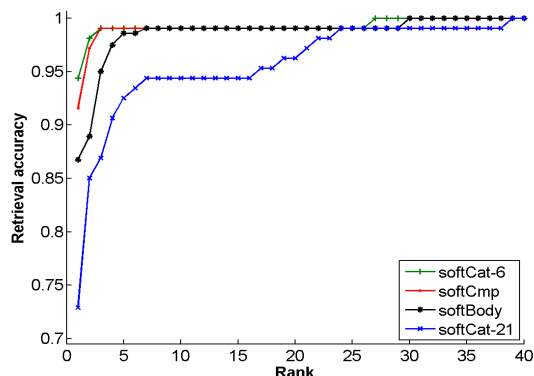


Figure 4: CMC (rank 40) of body soft biometrics and when add clothing

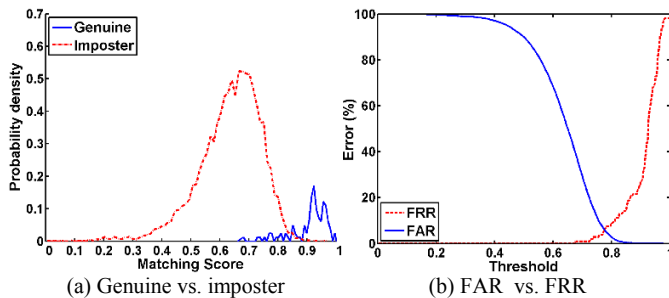


Figure 5: Intra-/inter-class distributions and error curves of *softCat-6*

5. Conclusions

In order to increase the capability of recognition and retrieval from surveillance imagery, we have studied the use of clothing for viewpoint invariant retrieval. By employing human labelling, soft clothing biometrics have been used to supplement body soft biometrics and they are capable of significantly increasing retrieval performance. Using comparative clothing traits and a small set of categorical clothing traits, which are highly correlated between multiple viewpoints, yielded more accurate human descriptions and can be used for viewpoint invariant subject

retrieval. The performance surpasses traditional and body soft traits when used in isolation. Although image-view variation can affect the observation of some clothing traits, the perception of other clothing traits remain invariant to viewpoint change and retain capability for subject retrieval. Here, the new clothing descriptions are shown to be relatively immune change from front to side view.

Proceeding from this research, a major future work can be to exploit our analysis and comprehension of clothing information, as soft biometrics, towards learning automated clothing label detection and description, which can pave the way for novel and effective biometric applications. We also need to consider the translation of these approaches to real-world imagery, and evaluation on a wider selection of publically available databases (e.g. ViPER, or APiS).

References

- [1] D. Reid, M. S. Nixon, and S. Stevenage, "Soft Biometrics; Human Identification using Comparative Descriptions," *IEEE TPAMI*, vol. 36, pp. 1216-1228, 2014.
- [2] J. Thornton, J. Baran-Gale, D. Butler, M. Chan, and H. Zwahlen, "Person attribute search for large-area video surveillance," in *HST*, 2011.
- [3] D. Adjeroh, C. Deng, M. Piccirilli, and A. Ross, "Predictability and correlation in human metrology," in *WIFS*, 2010.
- [4] S. Samangooei and M. S. Nixon, "Performing content-based retrieval of humans using gait biometrics," *Multimedia Tools and Applications*, vol. 49, pp. 195-212, 2010.
- [5] S. Denman, A. Bialkowski, C. Fookes, and S. Sridharan, "Determining operational measures from multi-camera surveillance systems using soft biometrics," in *AVSS*, 2011.
- [6] J. Zhu, S. Liao, Z. Lei, D. Yi, and S. Z. Li, "Pedestrian Attribute Classification in Surveillance: Database and Evaluation," in *ICCV*, 2013.
- [7] D. A. Vaquero, R. S. Feris, D. Tran, L. Brown, A. Hampapur, and M. Turk, "Attribute-based people search in surveillance environments," in *WACV*, 2009.
- [8] E. S. Jaha and M. S. Nixon, "Soft Biometrics for Subject Identification using Clothing Attributes," in *IJCB*, 2014.
- [9] E. S. Jaha and M. S. Nixon, "Analysing Soft Clothing Biometrics for Retrieval," in *BIOMET*, 2014.
- [10] A. Li, L. Liu, K. Wang, S. Liu, and S. Yan, "Clothing Attributes Assisted Person Re-identification," *IEEE CSVT*, 2014.
- [11] R. Layne, T. M. Hospedales, and S. Gong, "Person Re-identification by Attributes," in *BMVC*, 2012.
- [12] H. Chen, A. Gallagher, and B. Girod, "Describing clothing by semantic attributes," in *ECCV*, 2012.
- [13] K. Niinuma, U. Park, and A. K. Jain, "Soft biometric traits for continuous user authentication," *Trans. Inf. Forensic Secur.*, vol. 5, pp. 771-780, 2010.
- [14] A. C. Gallagher and T. Chen, "Clothing cosegmentation for recognizing people," in *CVPR*, 2008.
- [15] D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in *PETS*, 2007.
- [16] M. D. MacLeod, J. N. Frowley, and J. W. Shepherd, "Whole body information: its relevance to eyewitnesses," in *Adult eyewitness testimony*, ed: Cambridge Uni. Press, 1994.
- [17] J. Shutler, M. Grant, M. S. Nixon, and J. N. Carter, "On a large sequence-based human gait database," in *RASC*, 2002.
- [18] T. Joachims, "Optimizing search engines using clickthrough data," in *ACM SIGKDD*, 2002.
- [19] D. Parikh and K. Grauman, "Relative attributes," in *ICCV*, 2011.