

Imputing Human Descriptions in Semantic Biometrics

Daniel A. Reid and Mark S. Nixon
Information: Signals, Images, Systems
School of Electronics and Computer Science
University of Southampton
[dar1g09 | msn]@ecs.soton.ac.uk

ABSTRACT

Human identification at a distance has received significant interest due to the ever increasing surveillance infrastructure. Biometrics such as face and gait offer a suitable physical attribute to uniquely identify people from a distance. When linking this with human perception, these biometrics suffer from the semantic gap which is the difference between how people and how biometrics represent and describe humans. Semantic biometrics bridges this gap, allowing conversions between gait biometrics and semantic descriptions. One possible application of semantic biometrics is to automatically search surveillance footage for a person who best matches a given semantic description - possibly obtained from an eyewitness report. We now exploit patterns and structure within the physical descriptions to be able to predict occluded or erroneous data, thereby widening application potential. We show how imputation techniques can be used to increase accuracy and robustness of automatic semantic annotation of gait signatures.

Categories and Subject Descriptors

I.2.10 [Artificial Intelligence]: Vision and Scene Understanding; I.4.8 [Image Processing and Computer Vision]: Scene Analysis; H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval; I.4.9 [Image Processing and Computer Vision]: Applications

General Terms

Performance, Reliability

1. INTRODUCTION

With increase in terrorist and criminal activities, accurate human identification is critical. Quickly identifying humans at a distance from surveillance video would allow detection of known criminals. Popular biometrics capable of identifying humans from a distance, for example face[15] and gait[6] recognition, both suffer from the semantic gap. The

semantic gap is the difference between two representations, in this case the difference between how machines and how people recognize humans. It is common in eyewitness reports that a physical description of a criminal is available, this description cannot be translated to a machine understandable biometric. Semantic biometrics bridges this gap, allowing conversions between human descriptions and biometrics. One possible application for such a technique is to automatically search surveillance footage for people who best match a given human description.

Human descriptions inherently contain commonly co-occurring semantic labels, which can be thought of as structure. This structure occurs due to genetic, morphological and social factors and can be exploited to improve semantic biometrics. Extending analysis to footage of unconstrained environments introduces occlusion which can conceal a person's physical attributes from the camera. These occluded features can affect the automatic semantic annotation of the biometric data, leading to incorrect semantic labels. By utilizing the semantic structure we can compensate for missing visual features and correct erroneous semantic labels. Human descriptions are often considered unreliable and suffer under stressful situations. By exploiting the structure any likely discrepancies can be detected and corrected. The semantic structure can also improve semantic annotation under ideal conditions by increasing the accuracy of the latent semantic analysis technique used to annotate the biometric signatures.

The paper is organized as follows. Section 2 discusses the techniques used in the current state of the art. Imputation and the use of structure is explained in section 3 along with results showing the success of such an approach. Section 4 concludes the paper.

2. SEMANTIC BIOMETRICS

Semantic biometrics [8][9] utilizes latent semantic analysis to learn the relationships between gait signatures and semantic traits. By learning the relationships between the two representations, conversions can be made allowing automatic semantic labeling of gait signatures. Semantic biometrics is composed of three aspects, first deciding the most optimal semantic traits to allow identification, second to select the measured data to utilize, and finally the technique used to learn the relationships between the semantic traits and measured data.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MIFOR'10, October 29, 2010, Firenze, Italy.

Copyright 2010 ACM 978-1-4503-0157-2/10/10 ...\$10.00.

Table 1: Semantic traits and corresponding terms

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 Colour	[White, Tanned, Oriental, Black]
Facial Hair Colour	[None, Black, Brown, Blond, Red, Grey]
Facial Hair Length	[None, Stubble, Moustache, Goatee, Full Beard]
Hair Colour	[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]

2.1 Semantic Descriptions

To allow identification from eyewitness reports the physical properties described must be accurate, salient and reliable. Often human descriptions are considered to be unreliable by the eyewitness community [7], although research into semantic biometrics and soft biometrics have shown the benefit of such descriptions. Soft biometrics are those which can distinguish between populations; semantic biometrics extends the set to allow discrimination within populations. Studies in soft biometrics have shown by considering features like race, height and weight the correct classification rate of biometric systems can be improved [1][3][14].

The most complete study into salient and reliable physical descriptions for eyewitness testimonies was some time ago [4]. The most reliable descriptors were discovered in a two step process. The experiment started by asking participants to describe people within still images and videos. This generated 1,238 descriptors from the video footage alone. From these descriptors, 23 of the most common and distinct were selected to produce either 5 point bipolar scales or dichotomous items. Two groups of participants then labelled videos of people based on these descriptors to discover their reliability. To improve distinction between descriptors, redundancy was discovered and removed. This resulted in a reduced set of 13 of the most reliable descriptors.

Previous work into semantic biometrics [8][9] used 23 categorical traits (table 1) based largely on MacLeod’s work. Each is universal, distinct, easily discernible at a distance and largely permanent. The majority of traits focus on body proportions and global aspects like skin colour. Descriptions of faces are generally inaccurate and hence have been excluded from the 23 traits [12]. Categorical labels are used to avoid value judgments [5] which have been shown to be often inaccurate. This research will also use these 23 semantic traits due to their proven recognition capability.

2.2 Biometric Signatures

To determine the population’s subjective understanding of a particular semantic label it must be paired with visual features. These visual features can be thought of as a visual signature of the subject. This does not necessarily have to be unique to the subject but must encompass the person’s physical attributes and be a standard representation. This allows the latent semantic analysis technique (see section 2.3) to identify which visual features result in the person being labeled with a certain semantic description. For example which visual features are important in being perceived as ‘tall’ or ‘skinny’. By learning the structure between the visual and semantic features the technique can be used to automatically label people based on their physical characteristics.

It is vital that the visual signature encompasses the physical features which contribute to a semantic label being assigned to the subject. If these features are not present the technique will be unable to link the semantic term to a physical trait and hence will not identify what makes someone ‘tall’. Previously [8] used a gait signature as the visual feature for subjects, this was suitable for learning the majority of the 23 semantic traits although currently only works from a fronto-parallel viewpoint.

As stated earlier we aim to extend semantic biometrics to real world situations. This would undoubtedly involve surveillance cameras recording people in unconstrained environments. This brings several problems when generating a suitable visual signature, including lighting, resolution and frame rate limitations. The most detrimental factor in semantic biometrics is the unconstrained view point of the subject. People are not constrained to walk in one direction with respect to the camera, this means that the perceived physical appearance of the subject will be affected due to self occlusion. It is worth noting that the semantic descrip-

tions are inherently view invariant, although the biometric signatures from which they are obtained are not.

The gait signature used in [8] has been proven to contain the desired information, but only from certain view points. Obviously signatures obtained from different viewpoints will contain different views of the subject’s physical appearance, which may be beneficial or detrimental to certain semantic labels. In this paper structure will be used to counter this problem. By knowing how successfully a semantic label’s corresponding physical features are captured from a certain viewpoint, the possible detrimental effects can be countered using imputation approaches which are detailed in section 3. For this reason the current average gait signature will be used, allowing a suitable testbed in which to evaluate the developed imputation techniques.

2.3 Latent Semantic Analysis

Latent semantic analysis (LSA) [2] creates a vector space model which discovers latent semantic structure based on the occurrences of features within documents. In the case of semantic biometrics the documents are the subjects and the features are the pixels within the biometric signature and the semantic terms. By determining the co-occurrence between the features we can detect which gait features occur commonly with which semantic labels and hence deduce the relationships between the two representations.

The space is constructed with the use of a set of subjects with semantically annotated gait signatures. [8] obtained semantic labels for each subject within the Soton gait database [10]. Each subject was labeled by multiple annotators using an online gait annotator system. These semantic labels will be used within this research so that comparisons can be made to the original approach. Each subject produces a feature vector containing the 4096 visual features obtained from the gait signature and 113 semantic terms denoting each semantic label. Each training subject’s feature vector is combined to create the term-document matrix \mathbf{O} .

The occurrence matrix will contain underlying structure between the gait features and semantic terms. This structure is hidden under a majority of irrelevant occurrences between features. By removing the irrelevant relationships (noise) the 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 occurrence matrix. Singular Value Decomposition (SVD) is utilized to factorize the matrix allowing a rank reduced version to be determined. First factorizing the matrix \mathbf{O} into three matrices such that:

$$\mathbf{O} = \mathbf{TSD}^T \tag{1}$$

Where \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 integral information within \mathbf{O} and remove the noise. The reduced rank k determines how many dimensions the data is condensed to and ultimately how much information is lost. The diagonal matrix \mathbf{S} consists of r diagonal values, these 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.

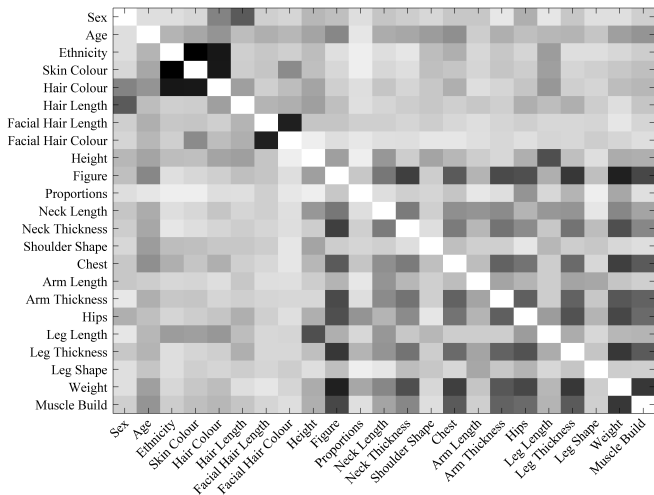


Figure 1: Correlation (Pearson’s r) between semantic features

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 the features and documents. The distance between these entities will correspond to the 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 semantic annotation is achieved by inserting a test subject’s gait signature into the semantic space. The closest, and hence most relevant, semantic terms are then assigned to the subject resulting in a semantically annotated gait signature.

3. IMPUTATION

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 common on females), genetics (black hair common within people of Asian descent) or the morphology of the human body (taller people more likely to have longer legs). This structure offers a basis to improve automatic semantic labeling and to make the system more robust to missing semantic labels or occluded visual features.

To verify that structure is present within the semantic features a correlation matrix was produced. This shows correlation between semantic terms based upon their occurrences within the Soton gait database. It is worth noting that the categorical data being used within the semantic labels contains no ordering. For this reason each possible ordering between the labels was enumerated and the maximum correlation was deemed to be the most representative of the relationship between the two traits. Figure 1 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 colour, hair colour 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 colour and hair colour. 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 not a relationship between the features only that it is not prevalent within the dataset currently being used.

One interesting and necessary use for structure is to allow inference to predict missing semantic labels. Imputation techniques are a statistical approach used to predict missing variables. Using such techniques missing semantic features can be predicted utilizing the structure within the semantic labels. Within this section two applications of imputation will be discussed. The first attempts to predict missing semantic labels to provide a solution for view invariance and the issues encountered when obtaining human descriptions from witnesses. The following section introduces an advanced semantic space insertion method used to improve the automatic semantic annotation of gait signatures.

3.1 Missing Semantic Labels

Occlusion is a major concern when applying semantic biometrics to unconstrained environments. Visual features can be concealed by the scenery, the person’s body (self occlusion) or covariates such as bags, hats and clothing. Structure can be utilized to predict the semantic labels normally obtained from the occluded visual data, to provide a more complete description of the person. Similarly human descriptions obtained from witnesses are often unreliable and incomplete, by utilizing structure these erroneous or missing labels can be assessed and the most probable labels based on previous experience can be used to refine the description.

Four techniques have been explored for providing such a solution. The first is to make use of the semantic space built during training. By inserting the subject’s feature vector into the space we can determine which terms, belonging to the missing trait, are closest. As explained before, the distances between a subject’s feature vector and semantic terms determines their similarity. Therefore the closest term belonging to the missing trait is the most relevant and hence most suitable, and is assigned to the subject. The problem with such a technique is that the subject’s feature vector must be complete to allow insertion. To complete the feature vector any missing features must be replaced with zeros. These zeros affect the positioning of the subject within the semantic space. The positioning error causes unrelated semantic terms to be closer to the subject leading to incorrectly ascribed values.

SVD imputation [11] was designed to counter the limitations observed in the previous technique. By inserting only the present data into the semantic space the missing data does not affect the positioning. Once the position has been obtained the missing data can be deduced from the subject’s position. Using this technique any missing data can be rebuilt utilizing the structure implicitly learnt by the LSA technique from previous co-occurrence data.

A technique was developed to utilize the known correlation between different semantic traits (figure 1). If the missing semantic trait is highly correlated with another trait

Table 2: Observations of Hair Colour and Skin Colour

	Black	Blond	Brown	Grey	Red	Dyed
Black	1	0	0	0	0	0
Oriental	23	0	0	0	0	0
Tanned	6	0	1	0	0	0
White	1	17	54	2	1	2

Table 3: Percentage of observations of Hair Colour and Skin Colour

	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

then it is beneficial to exploit this relationship to predict the missing term. The technique uses a similar method as the k nearest neighbour (k NN) classification technique and was inspired by work within [13]. 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 neighbour’s trait is weighted by the correlation between that trait and the subject’s missing trait. This favours neighbours with the same labels for traits with a strong relationship with the missing trait. The similarity between two subjects is determined as shown in equation 2 where X is the neighbour, Y is the subject with missing trait i . N is the total number of semantic traits and the matrix \mathbf{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 neighbours.

$$Similarity(X) = \frac{1}{N} \sum_{j=1}^N |\mathbf{C}_{i,j}| (1 - |X_j - Y_j|) \quad (2)$$

Correlation cannot determine relationships between terms and traits, though it is adequate for determining linear relationships between traits. Table 2 shows the observations of skin colour and hair colour 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 colour contain the least variance over their occurrences with hair colour. 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 when looking for the k nearest neighbours. If table 2 is converted to percentages showing the distribution of a term over the trait hair colour (see table 3) the variance can be easily identified. 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 neighbour’s similarity.

The similarity between two subjects is determined as shown in equation 4 where X is the neighbour, Y is the subject with missing trait i , which is composed of T_i terms. The matrices \mathbf{O} and \mathbf{P} contain the observations (table 2) and percentages of observations (table 3) respectively between terms, such that $\mathbf{O}_{i,j}$ details the observations of term j with term l . N is the total amount of semantic terms within the semantic feature vector.

$$H(j) = 1 - \sum_{l=1}^{T_i} -\mathbf{P}_{j,l} \log \mathbf{P}_{j,l} \quad (3)$$

$$\text{Similarity}(X) = \frac{1}{N} \sum_{j=1}^N H(j)(1 - |X_j - Y_j|) \quad (4)$$

This entropy approach was extended further to provide more robust operation. The most obvious addition to such a technique is to provide the metric with some measure of reliability. For example in table 2 and 3 we can see that black skin occurs with black hair 100% of the time. This is very misleading as only one person with black skin was observed and there is no confidence in the assumption that all people with black skin have black hair. To take this into account the reliability is gauged by the amount of observations of the term. Equation 5 shows the additional weighting function, where S represents the number of subjects within the training data.

$$R(j) = \frac{1}{S} \sum_{l=1}^{T_i} \mathbf{O}_{j,l} \quad (5)$$

Giving us a novel two part weighting function as seen in equation 6.

$$\text{Similarity}(X) = \frac{1}{N} \sum_{j=1}^N H(j)R(j)(1 - |X_j - Y_j|) \quad (6)$$

Figure 2 shows results from an experiment testing each technique’s success. The experiment was to predict a single missing semantic trait using the subject’s remaining semantic traits. This was performed by artificially removing every trait in turn from every subject within the test set and analyzing the accuracy of the rebuilt data. It can be seen that the semantic space is not suitable for rebuilding missing semantic terms. Both techniques utilizing the space (SVD imputation and semantic space distance) provided inaccurate results featuring high variance. Pearson’s R correlation metric together with a k NN approach featured less variation but was unable to detect term-trait relationships, resulting in a lower average accuracy compared to the SVD imputation technique. Finally the entropy approach utilizing the novel two part weighting process achieved the most successful results, a higher average accuracy and a smaller variation was observed. Figure 3 shows the accuracy of rebuilding each semantic trait using the entropy based approach. It can be observed that the least successful traits are those which include lengths and heights. This may be due to the normalization process when gathering the gait signature, which removes height information. The most successful traits are skin colour and ethnicity, this is likely due to their strong correlation with other traits allowing accurate predictions of missing data.

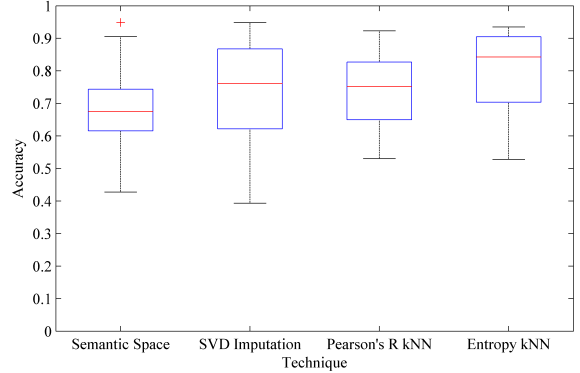


Figure 2: Results from rebuilding semantic data based on remaining semantic traits

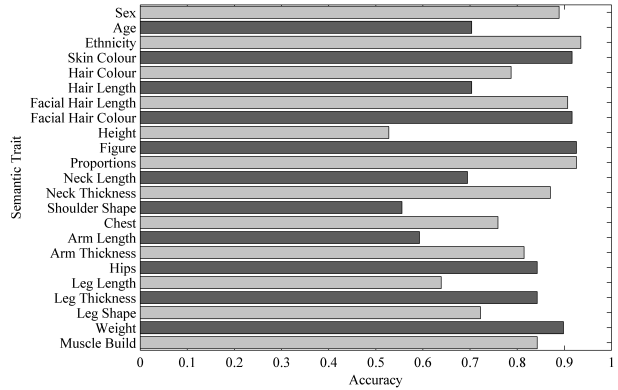


Figure 3: Rebuilding semantic data using entropy based weighted k NN

3.2 Semantic Space Insertion

Inserting a subject’s gait signature into the semantic space facilitates automatic semantic annotation. To allow insertion the subject’s feature vector must be complete. The missing semantic features are initially assigned to zero to permit insertion. Once the position is obtained, the closest semantic terms within the space are used to generate the semantic feature vector for the subject. The semantic features assigned to zero ultimately affect the subject’s positioning within the semantic space, leading to incorrect annotations. The impact of this misplacement can be minimized by predicting the missing semantic feature vector based on the visual feature before insertion. This is not as accurate as annotating via the semantic space but it does allow the subject to be positioned more accurately by generating a reasonable semantic feature vector. Once positioned the imputed semantic feature vector is replaced using the semantic space annotation method discussed earlier. A semantic space based imputation approach is required to understand the complex relationship between the gait signature and semantic feature vector. For this reason the SVD imputa-

tion technique discussed within section 3.1 is used. Figure 4 shows the comparison between imputing the missing semantic data and using zeros. It can be seen that utilizing the structure learnt by the LSA algorithm allows the semantic data to be adequately initialized, leading to a more robust insertion method over all possible dimensions of the semantic space.

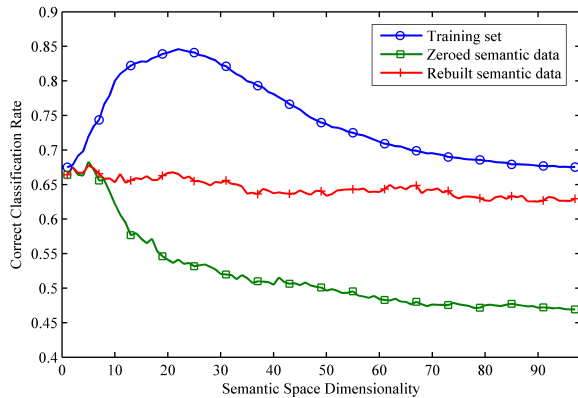


Figure 4: Classification accuracy over all possible dimensionalities of the semantic space

4. CONCLUSIONS

Semantic biometrics is a new and exciting field of research closing the semantic gap between human descriptions and biometrics. Human descriptions can now be utilized to search for people automatically, bypassing manual semantic labeling and the problems associated with the process. Current semantic biometric techniques use LSA to relate 23 categorical semantic traits with an average gait signature, facilitating automatic semantic annotation of gait signatures.

This paper describes several advances to the state of the art. By utilizing structure within the semantic labels, issues with view invariance and the subjective and unreliable nature of semantic labels can be minimized. Occlusion is a significant issue leading to physical features not being observed and the related semantic labels not being generated correctly. A possible solution is to assess which semantic labels will be affected by the missing visual features and predict them based on the correctly obtained semantic labels. A developed imputation technique using an entropy weighted k NN approach showed an 79% average accuracy at rebuilding a missing semantic trait. This technique can also be used to improve erroneous or incomplete human descriptions.

The automatic semantic annotation of gait signatures has also been improved. Previously test subject’s semantic feature vector would be filled with zeros to allow insertion into the semantic space. These zeros affected the final annotation of the subject, decreasing the robustness of the technique. The SVD imputation technique was successfully used to predict the most likely semantic feature vector, based on the subject’s gait signature. This approach achieved an average 13% accuracy increase over all possible dimensionalities of the semantic space.

Structure is inherent within physical descriptions and can be exploited to improve both the semantic labels received and the automatic annotation of often incomplete gait signatures. Future research will focus on improving the semantic space used to automatically annotate gait signatures and the ability to rebuild occluded visual features.

5. REFERENCES

- [1] H. Ailisto, E. Vildjiounaite, M. Lindholm, S. Makela, and J. Peltola. Soft biometrics - combining body weight and fat measurements with fingerprint biometrics. *Pattern Recognition Letters*, 27(5):325–334, 2006.
- [2] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman. Indexing by latent semantic analysis. *Journal of the American society for information science*, 41(6):391–407, 1990.
- [3] A. K. Jain, S. C. Dass, and K. Nandakumar. Soft biometric traits for personal recognition systems. In *International conference on Biometric Authentication*, pages 731–738, 2004.
- [4] M. D. MacLeod, J. N. Frowley, and J. W. Shepherd. Whole body information: Its relevance to eyewitnesses. In *Adult eyewitness testimony: Current trends and developments*, chapter 6. Cambridge University Press, 1994.
- [5] C. A. Meissner, S. L. Sporer, and J. W. Schooler. Person descriptions as eyewitness evidence. *Handbook of eyewitness psychology*, 2:3–34, 2007.
- [6] M. S. Nixon and J. N. Carter. Automatic Recognition by Gait. *Proceedings of the IEEE*, 94(11):2013–2024, 2006.
- [7] D. F. Ross, J. D. Read, and M. P. Togli. *Adult eyewitness testimony: Current trends and developments*. Cambridge University Press, 1994.
- [8] S. Samangoeei, B. Guo, and M. S. Nixon. The use of semantic human description as a soft biometric. In *2nd IEEE International Conference on Biometrics: Theory, Applications and Systems*, pages 1–7, September 2008.
- [9] S. Samangoeei and M. S. Nixon. Performing Content-Based Retrieval of Humans Using Gait Biometrics. In *SAMT*, pages 105–120, 2008.
- [10] J. Shutler, M. Grant, M. S. Nixon, and J. N. Carter. On a large sequence-based human gait database. *RASC*, pages 66–72, 2002.
- [11] O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein, and R. B. Altman. Missing value estimation methods for DNA microarrays. *Bioinformatics*, 17(6):520–525, 2001.
- [12] P. J. Van Koppen and S. K. Lochun. Portraying perpetrators; the validity of offender descriptions by witnesses. *Law and Human Behavior*, 21(6):661–685, 1997.
- [13] I. Wasito and B. Mirkin. Nearest neighbour approach in the least-squares data imputation algorithms. *Information Sciences*, 169(1-2):1–25, 2005.
- [14] R. Zewail, A. Elsafi, M. Saeb, and N. Hamdy. Soft and hard biometrics fusion for improved identity verification. In *MWSCAS*, pages 225–228, 2004.
- [15] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *Acm Computing Surveys (CSUR)*, 35(4):399–458, 2003.