Body Shape Biometric System: An Active Contour Modelling and Segmentation Approach (2000)

Jordan Spooner
MSc Computer Science
30068649

Abstract— **Biometrics play an important role in society by identifying individuals through contact and non-contact methods. Body shape is a non-contact biometric used in this paper that is obtained and used in KNN and SVM classifiers. The proposed biometric assesses the capability of body shape as a unique identifier and shows that clothes and pose variations present issues to this biometric. The KNN model outperforms the SVM model and shows how side-profiles of a subject have the potential for recognition with the use of more complicated methods such as cloth invariance. (90 words)**

I. Introduction

Each person possesses unique and identifiable attributes, commonly referred to as biometrics. Some widely recognized biometrics are body shape [10, 11], face, fingerprint, and walking gait [3, 4, 8]. The fingerprint requires contact with a sensor to obtain its signature and, while the face can be used to identify a subject with significant results, this biometric requires a high enough resolution [12] and the subjects face to be facing the camera to ensure that the identifiable features can be extracted [10, 11]. The use of biometrics has many applications and can be used to aid humans in surveillance and potentially replace them in the future [12].

Many images obtained from security cameras, used for surveillance, often contain subjects with obscured faces and weighted items that can affect the ability of face recognition and walking gait biometric systems; this may be due to hidden facial features [11] and altered walking gait patterns. Factors that may affect ones walking gait are camera shyness and human psychology [3]. While the database used on the proposed method in this paper does not contain image sequences of walking gaits, the subjects may also be affected by this factor as they may sub-consciously change their stance while a camera is pointed at them. The proposed method aims to recognize and classify individuals based on their body shape; a non-contact biometric that does not require high resolution images.

The rest of this paper is organized into five further sections; section 2 introduces the field of contactless biometrics such as gait and body shapes; section 3 describes the proposed method; section 4 talks about the results and discussion and, finally, section 5 is the conclusion.

II. Literature Review

Body shapes are unique [11] and useful for recognition; this is consistent with [11] where a human identification system achieved a correct recognition rate of 77.5%. However, this doesn't consider cloth variations, which introduces ambiguity [1]. Data collection techniques to introduce cloth invariance include laser-scanning [10] and multi-view images with varying poses where the clothes constraints vary on the human body [2, 5].

Alternatively, [2] mitigates the cloth issue by estimating the body shape from anthropometric markers regressed onto the subject and then optimizing the anthropometric distances and markers with respect to anthropometric constraints to form close representations to a statistical body model. [13] assumed that with more manual measured data they generate better anthropometric estimated parameters; however, this requires time/ resources to annotate subjects. Additionally, anthropometric distances and markers don't consider non-skeletal information such as muscle/ fat distribution [6], disregarding external shapes of the subject that is unique and unpredictable. This is important for identifying subjects of different sex as females have a different fat distribution compared to males [10].

When considering silhouettes, cloth invariant methods need to be applied to remove the noise that they add. [2] exploits temporal information from multi-view images to estimate the body shape of the subject under their clothing using an SMPL body model, while [9] uses a Convolutional Neural Network (CNN) in conjunction with a body contour key point dataset applied to one image to obtain under-cloth contour key points. [8] also uses deep learning networks, e.g. a CNN for human recognition, though, using gait information; their method is also cloth invariant, and viewpoint invariant.

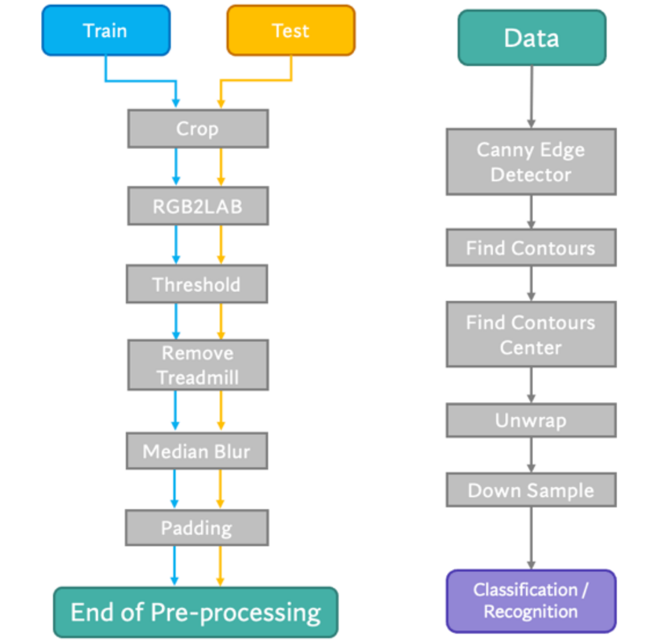
As the training dataset contains 44 front facing subjects and 44 corresponding side facing subjects, use of a deep network wouldn't be appropriate. [10] compares SVM, KNN and ANN models, and shows the SVM to be the best classification algorithm for gender recognition using body shapes. [12] also shows the capability of the SVM classifier where it outperforms KNN for human identification, when using color and shape-based features. Due to the high dimensionality of images, PCA has been applied [4], despite it removing large amounts of data, it is valuable in estimating body shapes [5].

Many available systems are viewpoint invariant as they use laser-scans [5], [10] or use multi-view images [6]. [8] requires prior knowledge about the viewing angles. However, the dataset provided only gives images of the subjects' front and side on, therefore prior information is known about the viewpoint, and viewpoint invariance isn't considered.

III. Method

A. Pre-processing

The goal of the pre-processing stage is to produce a silhouette of the subject to be unwrapped by the feature extraction steps. The subject isn't the only object in the image; therefore, a series of steps are required to ensure they are the only object that is extracted, fig 1 a).



a) b)

Figure 1. Abstractly shows the steps taken for the a) pre-processing and the b) feature extraction stages.

The set of images were cropped by manually inspecting where the subject could stand with respect to the treadmill (centered) in each image. Ensuring that the background only contained the green backdrop for accurate foreground segmentation. RGB to LAB conversion then occurred before binary thresholding as this provided optimal thresholding results; even when subjects wore similar color clothes to the background. The user and the subjects' feet are often the same color as the treadmill, moreover, people often share foot size; therefore, this part of the image was cropped. The treadmill handle remains and is located using template matching; the pixels based in its most probable location are converted to white pixels. Consequently, the body parts of some subjects positioned with the treadmill handle are removed. Background subtraction is an alternative, however insufficient background information led to inconsistent extraction results of the subjects. The penultimate pre-processing step, median blur, aimed to reduce the affect that clothes have on the edges of the silhouette, and removed any background noise. Lastly, the bottom of each image was padded with white pixels to prevent the active contour model defining the edge of the image as a contour.

Figure 2. Four cropped images of a front-facing subject in the training dataset; a) RGB image, b) LAB image, c) thresholded image after RGB2LAB conversion, d) thresholded image with treadmill removed using template matching.

B. Feature Extraction

The objective of feature extraction is to unwrap the contour of the silhouette for classification and recognition. Fig 1 b) outlines the steps for feature extraction.

$$\bar{x} = \frac{M_{10}}{M_{00}} \tag{1}$$

$$\bar{y} = \frac{M_{01}}{M_{00}} \tag{2}$$

$$M_{ij} = \sum_x \sum_y x^i y^j I(x,y) \tag{3}$$

Combination of the Canny edge detector with active contour modelling ensures a complete contour of the subject's silhouette; some edges are missed without the edge detector [14]. Artefacts sometimes remain after pre-processing, fig 3 a). Therefore, only the largest contour is considered, and its geometric center is calculated using equations (1) and (2). This center is used as the point of reference for unwrapping the silhouette, using the Euclidean distance between the center and each contour point. Each subject may have a different number of contour points; therefore, each unwrapped representation is down sampled to 2000 contour points before normalization. Fig 3 shows the contours, centers, and corresponding unwrapped silhouettes.

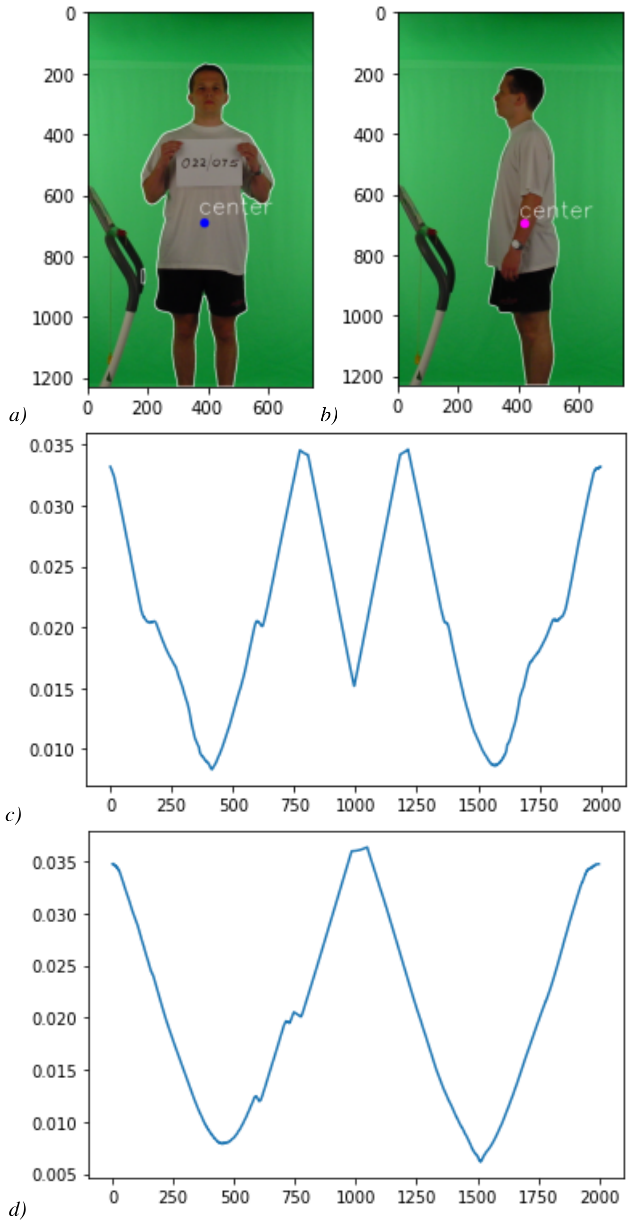


Figure 3. The first two images show a contour surrounding the subject with the geometric center of the corresponding contour for the a) front facing profile and the b) side profile. The last two images are the unwrapped contours for the c) front facing profile and the d) side profile.

PCA is appropriate to avoid overfitting as there are 2000 features and only 88 samples in the training set. However, literature does not use PCA for dimensionality reduction, but uses it for body shape alignment [10], therefore the proposed method is compared with and without PCA.

C. Recognition and Verification

The final unwrapped silhouettes are inputs for a multi-class SVM and KNN classifier. [10, 12] showed, respectively, that SVM outperformed KNN for gender classification and for full-body person recognition. Deep learning models e.g. CNN's were considered, given the initial input being an image, however, deep models require more data to prevent over fitting. [8] creates a generative model to generate artificial data justifying CNN's, however, the dataset would remain insufficient.

IV. Results and Discussion

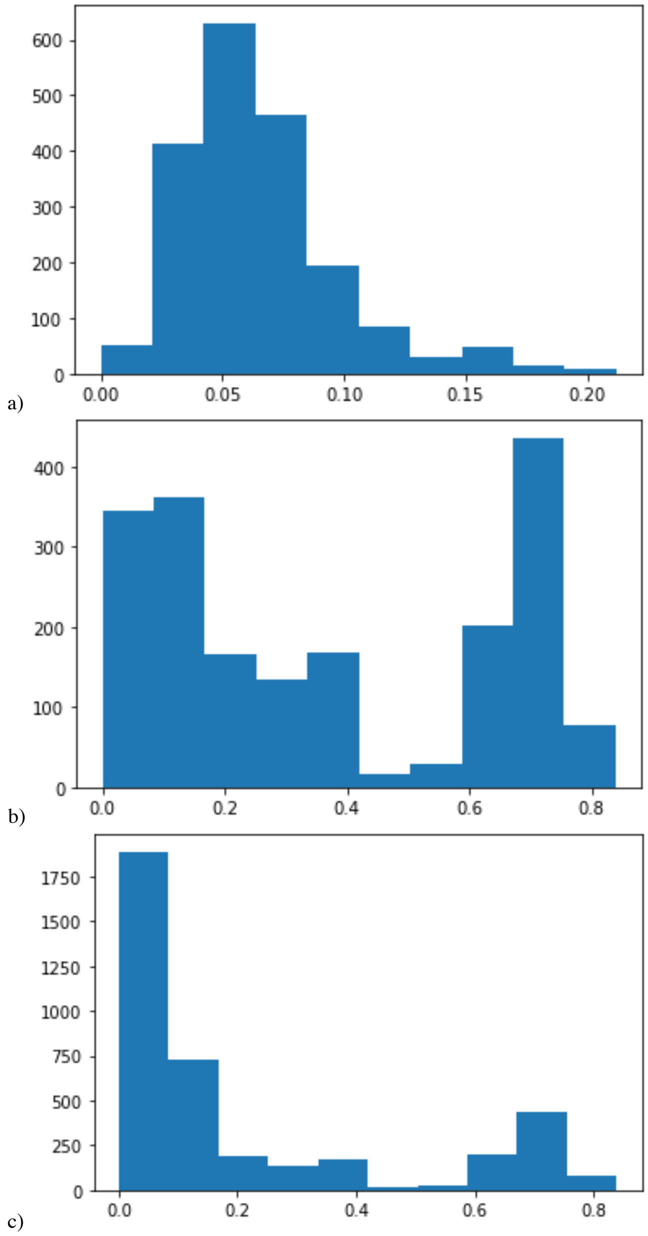


Figure 4. shows 3 histograms of (Euclidean) distances for each item of data i against item of data j where i and j are elements of a dataset of a) side profiles, b) front profiles, c) all profiles.

An analysis of the distances between each of the side profiles shows a positively skewed distribution of distances, fig 4 a). The set of distances between each subject for their front profiles appears to have less structure and has a sparser subject space; the range of distances between the different subjects is greater, fig 4 b). Combining these sets of distances shows that the side profile set of distances dominates the overall histogram, fig 4 c); this is expected as the distance histogram of the side profiles is denser.

For the classification task, three additional datasets were generated to fully assess the performance of the KNN and SVM models. KNN applied to all these datasets clearly outperforms the SVM multi-class classifier, table 1. It is also evident that by applying SVM and KNN to the side-profile dataset, the accuracy increases significantly; the accuracy decreases significantly when applied to the front-profile datasets. The front-facing subjects holding a sign in the training dataset suggests that this drastically affected the classification capabilities of the SVM and KNN models as for the side-profiles with consistent poses, the classification rate is extremely higher.

TABLE I. CORRECT CLASSIFICATION RATES FOR RECOGNITION

	Dataset			
Model	**All Profiles**	**Side Profiles**	**Front Profiles**	**All Profiles PCA50**
SVM	0.0909	0.3636	0.0000	0.0909
KNN	0.2727	0.4545	0.0909	0.2727

The performance of the SVM multi-class classifier proved incapable for recognition. This is because there are a vast number of dimensions and the only 88 training images with 44 different users and 2 different profiles each. Additionally, the intra-class variation is large as the side profile and front profile share the same label while also appearing different, see fig 3 c) and d) where the unwrapping of the front and side silhouettes shows a visual difference between the two representations.

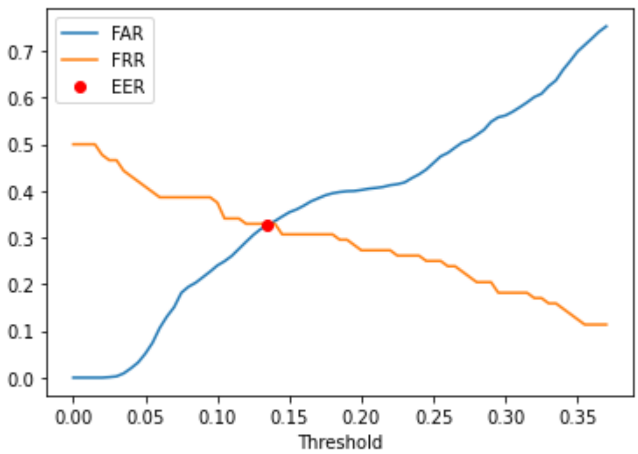


Figure 5. graph showing the False Error Rate (FAR), False Rejection Rate (FRR) and the point of intersection (EER) where FRR = FAR at x = 0.135, y = 0.327.

The KNN outperformed the SVM in the recognition task, thus, KNN was considered for verification. Due to this large inter class variation when considering the side and front profile of the same person coming from the same class, KNN for verification considers each image as its own class. The range of threshold values were 0 to 0.4 using a 0.005 increase.

The verification task shows that as the threshold value increases, the False Acceptance Rate (FAR) increases and the False Acceptance Rate (FRR) decreases, fig 5. The Equal Error Rate (EER) is calculated at a threshold of 0.0135 and FAR = FRR = 0.327. This means that the biometric system verifies a subject incorrectly at a 0.654 rate; this is not a secure biometric system. The reasons behind this poor verification rate are due to the cloth variation, and because each front-facing individual in the training dataset is holding a sign. Therefore, a new individual with different clothes and who is not holding a sign when facing the camera will be far away from their training image in the feature space. Although the threshold can be increased to allow for these individuals to be accepted, it significantly increases the rate of false acceptance. Inspection of the histogram of distances, fig 4 c), further suggests that this biometric system is not appropriate as it shows a high proportion of distances between different subject's profile signatures to be lower than the threshold, suggesting a large proportion of false acceptances.

V. CONCLUSION

In conclusion, the proposed method correctly segments the human subject. However, due to the location of some subjects, the process to remove the top of the treadmill also removes parts of subject's silhouette, affecting their unwrapped silhouette; clothes also affect this unwrapping. It is evident that median blur was not enough to provide cloth invariance as the classification and verification results were unsatisfactory. Although issues in the dataset, such as inconsistent front-facing poses and the limited number of images for each subject, the biometric system shows some capability of recognizing individuals by their side profile, with a recognition rate of 0.4545 with KNN. In contrast to [10, 12], KNN outperformed the SVM multi-class classifier in all the scenarios, and the results showed that PCA did not improve the performance of the system; however, SVM failed due to the limited number of images for each person. Future implementations using body shapes should aim for a database with consistent poses and a foreground with one object of interest, as opposed to a secondary object with similar colors to the subject's clothing/footwear. Furthermore, body pose, and statistical body shape optimization methods seen in [1], [2], should be implemented to create cloth invariant techniques that estimate the subjects body shape for more accurate recognition and verification.

REFERENCES

[1] A. Godil, P. Grother and S. Ressler, "Human identification from body shape", Fourth International Conference on 3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings.. Available: 10.1109/im.2003.1240273.

[2] C. Zhang, S. Pujades, M. Black and G. Pons-Moll, "Detailed, accurate, human shape estimation from clothed 3D scan sequences", *arXiv*, 2017. Available: 10.48550/ARXIV.1703.04454.

[3] J. Shutler, J. Grant, M. Nixon and J. Carter, "On a Large Sequence-Based Human Gait Database", *ourth International Conference on Recent Advances in Soft Computing, Nottingham, United Kingdom.*, pp. 66-72, 2022

[4] Y. Guan, C. Li and Y. Hu, "Robust Clothing-Invariant Gait Recognition", *2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2012. Available: 10.1109/iih-msp.2012.84

[5] W. Chang and Y. Wang, "Seeing through the appearance: Body shape estimation using multi-view clothing images", *2015 IEEE International Conference on Multimedia and Expo (ICME)*, 2015. Available: 10.1109/icme.2015.7177402

[6] Y. Shigeki, F. Okura, I. Mitsugami and Y. Yagi, "Estimating 3D human shape under clothing from a single RGB image", *IPSJ Transactions on Computer Vision and Applications*, vol. 10,

[7] . D. Song et al., "Clothes Size Prediction from Dressed-Human Silhouettes", *Next Generation Computer Animation Techniques*, pp. 86-98, 2017. Available: 10.1007/978-3-319-69487-0_7

[8] J. Luo and T. Tjahjadi, "View and Clothing Invariant Gait Recognition via 3D Human Semantic Folding", *IEEE Access*, vol. 8, pp. 100365-100383, 2020. Available: 10.1109/access.2020.2997814

[9]] S. Lu, F. Lu, X. Shou and S. Zhu, "DeepProfile: Accurate Under-the-Clothes Body Profile Estimation", *Applied Sciences*, vol. 12, no. 4, p. 2220, 2022. Available: 10.3390/app12042220

[10] J. Tang, X. Liu, H. Cheng and K. Robinette, "Gender Recognition Using 3-D Human Body Shapes", *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 6, pp. 898-908, 2011. Available: 10.1109/tsmcc.2011.2104950

[11] N. Rashid, M. Yahya and A. Shafie, "Human Identification at a Distance Using Body Shape Information", *IOP Conference Series: Materials Science and Engineering*, vol. 53, p. 012058, 2013. Available: 10.1088/1757-899x/53/1/012058

[12] C. Nakajima, M. Pontil, B. Heisele and T. Poggio, "Full-body person recognition system", *Pattern Recognition*, vol. 36, no. 9, pp. 1997-2006, 2003. Available: 10.1016/s0031-3203(03)00061-x

[13] Stancic, Ivo & Supuk, Tamara & Cecić, Mojmil. (2009). *Computer vision system for human anthropometric parameters estimation*. *WSEAS Transactions on Systems*. *8*. 430-439

[14] M. Baştan, S. Bukhari and T. Breuel, "Active Canny: edge detection and recovery with open active contour models", *IET Image Processing*, vol. 11, no. 12, pp. 1325-1332, 2017. Available: 10.1049/iet-ipr.2017.0336