Biometrics Coursework

Dávid Máté

***Abstract*—This study presents a biometric system designed for subject recognition and verification, focusing solely on body and head shapes, without the use of face recognition. Through experimentation, the system's performance is evaluated, revealing promising capabilities in accurately identifying and verifying subjects. This showcases the potential of body shape biometrics as a feasible alternative, particularly beneficial when subjects are not facing the camera. However, limitations exist, notably sensitivity to changes in clothing and posture. Overall, this work contributes to advancing biometric authentication techniques and highlights practical implications and challenges associated with using body shape cues in biometric systems.**

I. Introduction

BIOMETRIC systems are all around us and continue to increase in importance. These systems utilise unique physical or behavioural characteristics of individuals for automatic identification or authentication purposes. There are many use cases, such as border security, law enforcement and surveillance. Common modalities include face, fingerprint and iris-based recognition.

This report presents a biometric system that can automatically recognise and verify individuals based on their body shape alone. Section II provides a detailed description of the algorithm, explaining the subject segmentation, features and classification model used by the system. Section III presents the results for recognition and verification tasks on a part of the Large Southampton Gait Database. In Section IV we analyse these results, looking at the strengths and weaknesses of the system. We also suggest some possible improvements and future directions. To conclude the report, we provide a summary of our main findings and conclusions.

II. Method

A. Silhouette extraction

Fig. 1: Extraction using simple thresholding

To extract features from images of individuals, the silhouette first needs to be segmented.

Fig. 2: Extraction using background subtraction

The image is cropped to [320:1580, 750:1800] from 2240x1680, to only include the green background which makes segmentation a lot easier.

The first method considered is simple thresholding: a mask is applied to a range of green values, which removes the background. The image is then thresholded to create a binary image. However, this doesn't remove the treadmill, as seen in Fig 1. To resolve this issue, background subtraction is used instead. The camera angle changes slightly for a few of the images in the dataset, which is why aligning the subject image to the background image is needed for accurate subtraction.

The subject image is aligned to match the perspective of the background image using keypoint detection and homography estimation techniques. Firstly, it converts both images to grayscale and detects keypoints and descriptors using the ORB algorithm. Then, it matches descriptors between the images using the *BFMatcher* from OpenCV with Hamming distance and cross-checking. After sorting and extracting the best matches, it estimates a homography matrix using RANSAC to find the transformation between the images. Finally, it warps the image to align with the perspective of the background using the computed homography matrix, resulting in an aligned image.

After background subtraction, the resulting image is converted to grayscale and thresholded with threshold=40, resulting in a binary image of the subject's silhouette, as seen in Fig 2. This method performs better than the previous, however it still fails to extract the subject's shoes, due to most shoes being close in colour to the treadmill behind. As most subjects change shoes between the training and test datasets, this wouldn't be a descriptive feature. The silhouette is therefore simply cropped at the bottom.

The active contours method was also considered, however defining the initial contour for each subject was not possible.

B. First method

First of all, Hu moments are used to describe the shape of silhouettes. Fourier series of 10th order are used to approximate the contour of the subject, with the coefficients being used as features. The system utilises the implementation in the PyEFD package. Another feature is the height of the silhouette, calculated by finding the bounding box using OpenCV's *boundingRect(contour)*. Height is an important feature of this system, as it doesn't change with the viewpoint or based on the clothes the subject is wearing, like most other descriptors.

Finally, we use MediaPipe's *Pose* module for pose estimation. Pose landmarks are detected, as illustrated in Fig 3, which are then used to calculate the length of the lower leg, upper leg and torso. As these are visible in both front and side view images, they are ideal viewpoint-invariant features. The considerable variation in body proportions renders them valuable features for classification purposes. Other features, such as Histogram of Gradients (HOG) and SIFT were tried but provided no noticeable performance benefits.

All of these features are concatenated to create the feature vector representing a subject. To reduce the dimensionality, Principal Component Analysis (PCA) is applied, resulting in a final feature vector with length=40.

The biometric system applies a Support Vector Machine (SVM) to classify the images. The following hyperparameters are used, found using grid search: 'C': 1e-06, 'gamma': 'auto', 'kernel': 'rbf'.

Fig. 3: Landmark detection for estimating body proportions

C. Improved method

The method previously presented has a major flaw. If an individual changes clothes, like going from loose trousers to shorts, their silhouette changes drastically. Changing their pose by raising their arms to hold the paper has a similar effect. Different viewpoints of the same subject also lead to different silhouettes. This is why using object descriptors such as Hu moments and Fourier series for the whole body leads to poor classification performance.

An improvement of the previous method is presented next, which uses the shape of the head exclusively, making it robust to changes in clothing and posture. The goal of the system was to recognise individuals without using facial features. While this method doesn't extract features from the face but the head, the previous system is included as a baseline for fairness.

The head is extracted using the following simple process: starting from the top of the image, the first white pixel is found and a bounding box of size 250x200 is placed around it, as illustrated in Fig 4. This approach has several limitations, which are described along with potential solutions in Section IV.

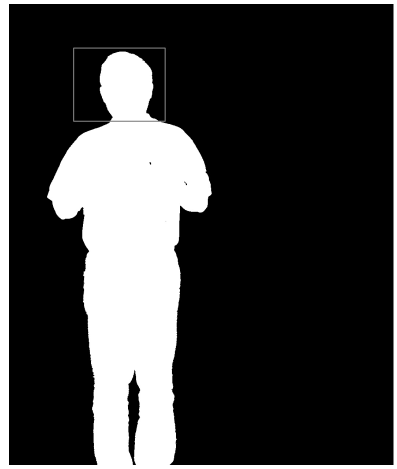

Fig. 4: Face detection

Hu moments and Fourier Series coefficients are then calculated for the binary image of the head, along with its height, width, area and perimeter. While these features will still change with the camera angle for a given subject, recognition accuracy is improved if the angle is the same in the training and test sets.

III. Results

This section summarizes the results of the improved biometric system for subject recognition and verification tasks on a part of the Large Southampton Gait Database. Fig 5 shows the distributions of intra-class and inter-class euclidean distances of the feature vectors. There is a large overlap between the histograms, which is explained by the structure of the dataset: each class (individual) only appears twice, from different viewpoints. The change in viewpoint leads to vastly different body shapes, which leads to different feature vectors and large intra-class variances.

A. Subject Recognition

For subject recognition, the classifier is given the feature vector of an individual and predicts the most likely identity from the set of individuals it was trained on.

The test dataset contains a subset of the subjects found in the training dataset, most with different clothing. The improved system achieves precision=0.73, recall=0.68, f1=0.68 and a rank-1 Correct Classification Rate (CCR) of 68%. On the other hand, the baseline method only achieves 45%, showing the importance of the modified features.

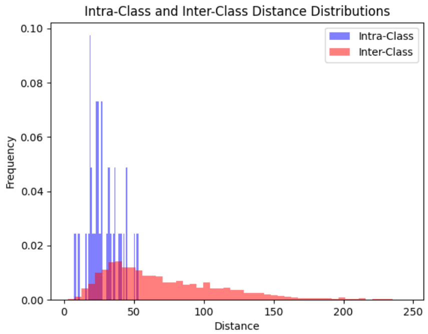

Fig. 5

B. Subject Verification

Subject verification is a binary classification task: the classifier takes 2 feature vectors (which are concatenated). These correspond to the claimed identity and the real identity of the individual. The dataset is constructed by taking 2 pairs of samples for each class: a genuine pair (where the classes are the same) and an impostor pair (where the classes are different). This is done instead of taking every possible combination of individuals to keep the dataset balanced. Otherwise, the dataset has many times more impostor pairs.

Figure 6 shows the ROC curve of the biometric system for verification, with the red dot representing the Equal Error Rate point and the dashed line being random guessing. The classifier achieves a Correct Classification Rate of 91% at the EER. The system demonstrates significantly improved performance on this task as it is inherently easier: it is binary classification, while recognition is a multi-class task.

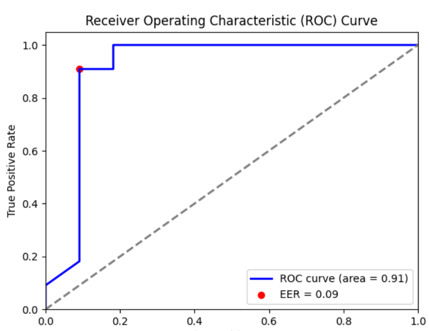

Fig. 6: Subject Verification ROC curve

IV. Discussion

The biometric system achieves a decent rank-1 recognition accuracy of 68% and a good verification accuracy of 91%, especially considering the dataset size. It does have several limitations, however. As the test dataset is exceedingly limited in size, the presented performance values may not generalise to a larger dataset. The accuracy of the system could be improved by applying a better pose estimation algorithm which is more robust to changes in viewpoint. The order of the Fourier Series could also be increased to better approximate the contours, given more training samples.

The head detection algorithm used is too specific and will only work for this dataset and this range, due to the fixed bounding box dimensions. The algorithm would fail to detect the head already if the subject raised their arm above their head. However, this can be improved by replacing it with a neural network trained for object detection, such as YOLOv8. Finally, another possible future direction is adapting the system presented to outdoor images, where silhouette extraction is significantly more difficult.

V. Conclusion

The biometric system presented showcases reasonable performance in subject recognition and verification, leveraging solely body and head shape cues, without relying on conventional face recognition models. Furthermore, the system achieves this without large amounts of training data, unlike large face recognition models. The study highlights the usefulness of body shape biometrics, especially in cases where subjects are not facing the camera and conventional methods would underperform. Nonetheless, it is essential to acknowledge the inherent limitations of such approaches, notably susceptibility to variables like changes in attire and posture, which could impede accuracy under certain conditions. The system is also not robust to viewpoint changes but would perform well when the viewpoint is constant. The presented techniques may also be useful when combined with other methods.