Silhouette Portrait Based Human Identification System

Yaodong Cui (29746396)

Abstract—**In many real-life situations, we need to identify human with images that are not suitable for traditional biometric methods. For instance, in many crime scenes, the quality of surveillance footage is very poor or the criminals coved their faces. Therefore, face or gait recognition may not be feasible. In this report, we propose a new method for human identification using only the silhouette of their upper body shape. We will extract subject's portrait(upper body shape) using K-Means clustering for color-based image segmentation. The extracted portrait will then be transformed into binary image form for contour extraction. We will then use Fourier descriptors to generate a label vector for each portrait's contour. We will repeat this process for the test images. Finally, we will implement K-nearest neighbor based on these Fourier descriptors. We will explore an extended version of this method, where other features of the portrait will be added to the feature vector for KNN classification.**

Index Terms—**body shape, Color-Based Segmentation, K-Means Clustering, KNN, Fourier descriptors, image classification,**

I. A INTRODUCTION OF SILHOUETTE PORTRAIT BASED HUMAN IDENTIFICATION SYSTEM

A. The disadvantages of current biometric system

It is common sense that the more information one got, the better judgment one can make based on that information. This is mostly true for the biometric system too. Face recognition, which utilizes many features of the human face can reach an astonishing high recognition rate. However, in real-world we often need to work in situations where information is heavily limited. In these situations, rich data based biometric systems could fail.

More and more people today try to hide their identity from surveillance cameras through various ways of disguises. This is even more prominent in the crime scene. Many criminals not only cover their faces but also change their clothes after committing the crime. And the fact that many surveillance camera can only produce poor quality footage worsen the situations. The lack of useful information and intentional introduced harmful information can mislead and render current rich data based biometric system useless. Therefore, it is essential to develop a different biometric system that can work with minimal information and is robust against disguises.

B. Reasons for using silhouette portrait method

Human is known for recognizing famous people through their silhouette. Many famous icons are specially designed in this way. For instance, the Silhouette portraits of the Queen and Sherlock Holmes showed in Fig. 1.

Figure 1: Silhouette portraits of the Queen and Sherlock Holmes

The benefits of using silhouette portrait can be concluded into three major aspects.

Firstly, much useful information, such as height, is actually preserved or can be preserved in silhouette portrait method. For instance, surveillance cameras have fixed angle and positions. Therefore, people of different height have drastically different silhouette portrait, as shown in Fig. 2.

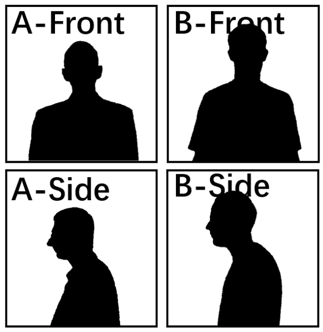

Figure 2: The front and side silhouette portraits of two male
The height information are not lost in silhouette portraits. As can be seen from subject A and B, different height leads to different contours.

Secondly, many potentially harmful and misleading information is discarded. In real-life footages, subjects are always moving. And their movements can change their whole body shape dramatically. Or in some cases, criminals try to change their cloth entirely. Most of these invalid or harmful pieces of information have little influence in the silhouette portraits,[2] as can be seen from Fig. 3.

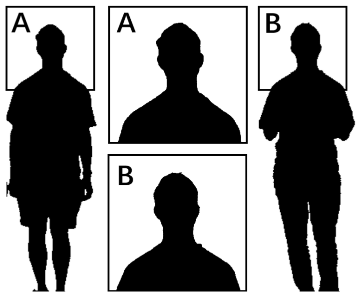

Figure 3: The silhouettes of the whole body and the portrait of the same male subjects from training and testing set
The silhouettes of the whole body are quit different This figure exhibit the harmful information that may hinter identification if using the silhouette of the whole body.

Thirdly, this method is robust against multiple disguises. Try to hide one's identity by wearing a big hat or a mask can actually help our identification method. This is because the disguise itself make their silhouette portrait more unique. As can be seen in Fig. 1, a hat will create more features in silhouette portrait, making it easier for recognition.

II. IMPLEMENTATION

In this section, we will discuss the implementation of our method in detail. This method has five main steps, namely preliminary process, segmentation, contour extraction, feature extraction (Fourier description) and classification.

A. *Preliminary Process of The Image*

We will first crop the original image to a certain size. All images from training and testing dataset can be cropped using a predefined strategy. We use a fixed crop strategy for two reasons. Firstly, the camera settings remain largely unchanged in this database, which provides a fixed angle and field. Secondly, all subjects in the data set are confined relatively small areas in the picture.

By cropping the image we can get rid of much irrelevant information in the original picture. This process will help us segment the portrait.

B. *Color-based Image segmentation*

The training and testing images in this dataset all have a green background. This property means we can implement color-based segmentation to extract the foreground.

We will first convert the cropped image from RGB color space to Lab color space. The Lab color space describes colors in three dimensions, namely luminosity axis, red-green axis and the blue-yellow axis. In Lab color space we can measure the Euclidean distance between colors.

The distance between colors allowed us to use K-means clustering to classify different object according to their colors.[1] In this process, we only need to cluster two classes of colors, namely green and others. This will allowed us to separate the background from the foreground(subject's portrait), as shown in Fig. 4, panel 3.

C. *Contour extraction*

After segmentation, we will set the foreground (subject) to black, as shown in Fig. 4, panel 4. This process increases the contrast of the contour's edge and suppresses other edges. We will then convert it to grayscale images for contour extractions. The extracted contour is shown in Fig. 4, panel 5.

Figure 4: The process flow of extracting the contour of subject's portrait

D. *Fourier descriptor*

We will use Fourier descriptor (FD) to extract contour's feature. The Fourier descriptor can characterize a contour using a sequence of frequency components (Fourier coefficient). The FD will first define a representation of the curve and then use Fourier theory to expand it. The FD is invariant to scaling and rotation.[3][4] If we use x(t) and y(t) to denote the coordinates of the subject's contour, we can form a complex number using the cumulative angular function(equation. 1).

$$c(t) = x(t) + jy(t) \tag{1}$$

$$c(m) = \sum_{t=-M/2}^{M/2} c[t]e^{j2\pi mk/N} \tag{2}$$

Then we can use equation.2 to calculate the Fourier expansion. In order to make the FD invariant to rotation, we need to divide every Fourier coefficient by the DC-coefficient(the first coefficient). However, because the settings of our camera system, we do not have the problem of rotation. Therefore, we omit the rotation invariant part and include the DC-coefficient in the feature vector. This will increase the correction rate.

E. *K-nearest neighbors classification*

After obtaining the feature vector (Fourier description), we will use the K-nearest neighbors (KNN) for classification. The KNN classify a subject using a majority vote of its K nearest neighbors.[6] In this design, the KNN will return indexes of the K closest positives and the distance between the test and these positives.

The use of KNN also means even the test subject is not in the database, the KNN classifier will return a classification result. Therefore, we will use a maximum distance to filter the result from the KNN classifier. We will reject results from the KNN if it exceeds the maximum distance.

III. Results

In this section, we will implement the FD based classification system and its improved method, where we combined simple scalar features (area feature) with the FD feature vectors. We will evaluate and compare their performance via classification rates and equal error rates.

A. *Evaluation of FD based method and combined method*

As can be seen from the table I and Fig. 5, the combined method (blue line) yields a better overall performance compared to FD only method. The combined method (Fourier description & area feature) has a top 1 correct classification rate of 59.1%, and it reaches 100% at top 16 guesses. The FD only method has a lower top 1 correct rate at 50%, and it reaches 100% at top 17 guesses.

Table I: The Correct Classification Rates of Top N Guesses

Models	Top 1	Top 3	Top 5
FD Based only	50.0% (11/22)	72.7% (16/22)	81.8% (18/22)
FD & area feature	59.1% (13/22)	77.3% (17/22)	81.8% (18/22)

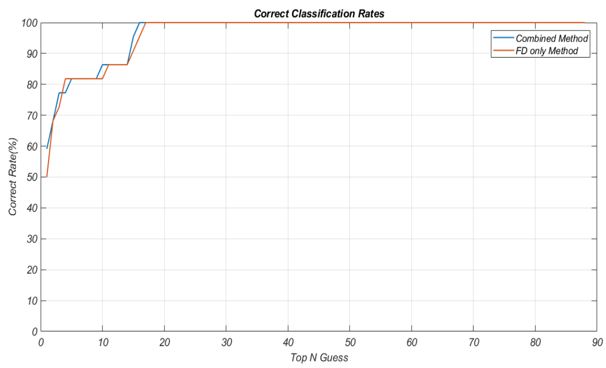

Figure 5: The Correct Classification Rate (top 1 to top 88 guesses) of the FD based method and combined method

The Fourier Description based method is marked in orange line. The combined method is marked in blue line.

In this design, we use Minkowski distance (p=2) to measure the feature distance between two points in a normed vector space, as shown in the equation. 3.

$$d(x,y) = (\sum_{i=0}^{n-1} |x_i - y_i|^p)^{1/p} \tag{3}$$

The Minkowski distance, when p set to 2, is known as the Euclidean distance. The between-class distance (blue bar) and within-class distance (red bar) of one test subject are shown in Fig. 6. As can be seen from the Fig. 6, the within-class distance distance is significantly smaller than any between-class distance.

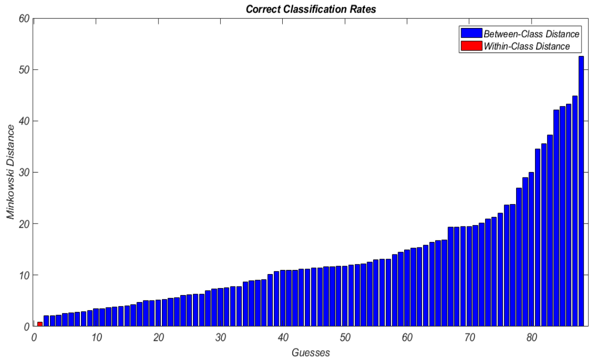

Figure 6: The histogram of between-class distance (blue bar) and within-class distance (red bar)

We use the Euclidean distances as the threshold for calculating the False Rejection Rate (FRR) and False Acceptance Rate (FAR). The intersection of FRR and FAR is the Equal Error Rate (EER), as shown in Fig. 7. The EER of the combined method is 0.38.

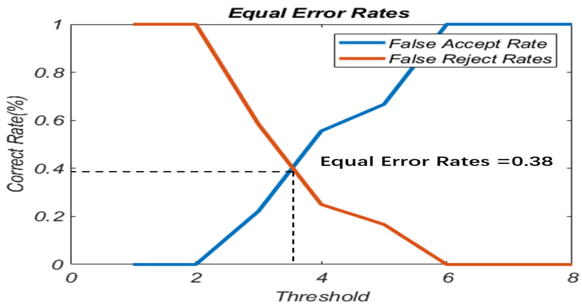

Figure 7: The plot of False Rejection Rate (FRR), False Acceptance Rate (FAR) and their convergence at Equal Error Rates (EER)

B. *Evaluation of simple scalar feature based method*

In this section, we will explore simple scalar features and a combined model of FD and simple scalar features. We will use KNN to classify the test images using each simple scalar feature and compare their performance using the correct classification rates.

Common simple scalar features includes area $A(s)$, perimeter length $P(s)$, compactness $C(s)$ and irregularity $I(s)$. In this design, we also include the neck length of each subject. The compactness can be calculated using equation 4. The irregularity can be calculated using equation 5.

$$C(s) = \frac{4\pi A(s)}{(P(s))^2} \tag{4}$$

$$I(s) = \frac{\pi max((x_j - \bar{x})^2 + (y_j - \bar{y})^2 +)}{A(s)} \tag{5}$$

A KNN classifier is used to classify the subjects using one scalar feature at a time. As can be seen from table II, the area and perimeter length are the most suitable scalar features. A line chart of correct classification rate is shown in Fig. 5. It is clear from Fig. 5, the area feature (yellow line) outperformed other features.

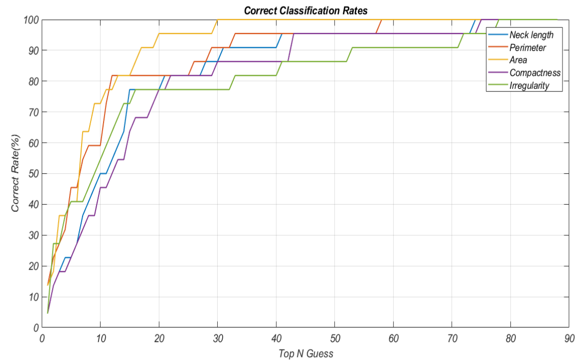

Figure 8: The Correct Classification Rate (top 1 to top 88 guesses) of different scalar features based method

Table II: The Correct Classification Rates of Top N Guesses

Models	Top 1	Top 3	Top 5
Area	13.6% (3/22)	31.4% (8/22)	41.0% (9/22)
Perimeter length	13.6% (3/22)	27.2% (6/22)	45.5% (10/22)
Irregularity	4.5% (1/22)	27.2% (6/22)	41.0% (9/22)
Neck Length	4.5% (1/22)	18.1% (4/22)	22.7% (5/22)
Compactness	4.5% (1/22)	18.1% (4/22)	22.7% (5/22)

When these simple scalar features are combined with the FD feature, the combination of area feature and FD feature can yield the best result (see Fig. 4).

IV. CONCLUSION

In summary, we propose a computer vision based human identification method using subject's silhouette portrait. This method employs a combined feature strategy (Fourier description and area feature) and uses KNN as the classifier. We explain the background and reason for using the subject's upper body silhouette in certain application areas. We then evaluate and compare different methods using performance metrics, such as correct classification rate, False Rejection Rate (FRR), False Acceptance Rate (FAR) and Equal Error Rate (EER).

A. *Advantages*

The silhouette portrait demonstrates robustness towards harmful information while maintaining a good recognition rate in our application areas. To be more specific, this method is insensitive to human movements, various of disguise, low-quality footage, and rotation. The potential application area of our method includes extreme situations, where it is impossible to implement the classic biometric method like facial recognition. In these extreme situations, our method demonstrates a good performance and robustness.

B. *Disadvantages*

Although our method can maintain a good classification rate in extreme situations. However, when conditions are ideal and more information, such as facial information, is available, a method that can utilize this information might outperform this method.

C. *Future Directions*

In theory, this method can be further improved by introducing more features into the feature vector. More features of silhouette portrait can be explored. However, simply adding more features into the feature vector can impair the overall performance. The combination of different features and their weights need to be tested comprehensively.

In real-world application area, the disadvantages of this method can be overcome by using an adaptive biometric system. An adaptive biometric system can choose to use one or more biometric information depending on conditions. Therefore, this system can utilize information efficiently. This ability can help it maintain a good performance in good and bad conditions.

REFERENCES

[1] MacQueen, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. 1. University of California Press. pp. 281–297. MR 0214227. Zbl 0214.46201. Retrieved 2009-04-07.

[2] Ismail Haritaoglu, David Harwood, and Larry S. Davis, "Ghost: A Human Body Part Labeling System Using Silhouettes", Fourteenth International Conference on Pattern Recognition, Brisbane, August, 1998.

[3] Hannu Kauppinen, Tapio Sepplinen and Matti Pietiktiinen "An Experimental Comparison of Autoregressive and Fourier Based Descriptors in 2D Shape Classification", Transactions on Pattern Analysis and Machine Intelligence, IEEE, February, 1995, V. 17, No.2, pp. 201-207.

[4] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," Signal Processing: Image Communication, vol. 17, pp. 825–848, 2002.

[5] S. Li and M.-C. Lee, "Effective Invariant Features for Shape Based," Journal of the American Society for Information Science and Technology,, vol. 56, pp. 729–740, 2005.

[6] Altman, N. S. (1992). "An introduction to kernel and nearest-neighbor nonparametric regression". The American Statistician. 46 (3): 175-185. doi: 10.1080/00031305. 1992.10475879