

Recognising people using body shape analysis

Man-Leong Chan, *University of Southampton*

Abstract—We present a combination of 2 methods for analysing a person's body shape and ultimately attempt to identify subjects from a set of photos by recognition. Silhouette and limb position analysis provides us metrics of the person's standing posture, which is used in our recognition algorithm.

1 INTRODUCTION

Given a set of images each with a subject standing in front of a green screen with one facing forward and another facing sideways, we can recognise a person's body shape and attempt to identify them individually. This involves extracting the subject from an image, generating subject's metrics and finally identification.

2 IMAGE PROCESSING

2.1 Green screen and background removal

Our domain is a set of photos each with a subject in front of a green screen standing on a treadmill. Subject extraction requires the removal of the green screen as well as any background captured outside of the green screen area. A couple of samples of green screen color reveals the average screen color to be $rgb(56, 175, 93)$. Image area beyond the green screen area is cropped to remove excess background area.

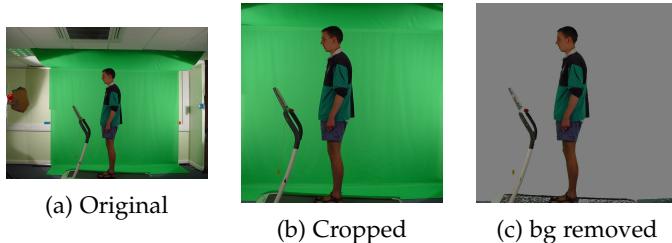


Fig. 1: Background removal

2.2 Skin tone extraction

Subject's limbs were extracted according to their skin tone. This extraction was done by analysing the HSV value of every pixel in the image. Where $0 < Hue < 25$, $0.50 < Saturation < 0.98$ and $Value > 0.22$. Our hue control selected skin tone from a range of color angles, it allowed the algorithm to be inclusive with a range of skin colour of different race and background[2]. Saturation controlled the range of colour sharpness which was crucial for our algorithm to differentiate between skin colour and shadows. The subject's skin surface under a reasonably well lit environment would have a higher saturation value than clothing in beige or brown.

Obviously tone extraction obviously relied on the subject exposing some part of their body. Primarily it required the



(a) Before extraction



(b) After extraction

Fig. 2: Skin tone extraction by HSV

subject to expose their face and hand for the best outcome. Therefore we are assuming that our subject does not wear a Niqab or gloves.

2.3 Limb Segmentation

Skin tone extraction gave us many fragments of skin color in different sizes. The larger fragments gave us more interesting ideas on what those limbs could be, whereas the smaller fragments scattered across the image could be noise which so happened to have the same HSV values as our subject's skin.[1]

In order to carry out further analysis, we had to outline our desired fragments which were potentially part of subject's limbs. A floodfill analysis was used to determine the outline of those potential targets. It determined fragments which were connected in this 2 dimensional pixel array and output the outline of each fragments. Fragments which are smaller than any desired features were ignored to reduce noise in the data.

Once the fragments were identified, they were subjected to feature analysis. Up to this point, the fragments are nothing but scattered data, feature analysis aims to identify which body part these fragments belong and this would give us useful data towards analysing subject's standing posture. Here we made a couple of assumptions.

- Firstly the subject is not raising their hands above their head at the moment when the picture was taken.

- Secondly the subject is not vertically rotated.

Then we ran a verticle analysis where we first identify the person's head followed by their arms and leg. We assumed not everyone were wearing short and exposing the skin from their legs, so identifying legs were not compulsory. Similarly, some subjects were wearing long-sleeves and others were wearing t-shirts, the variation between the detected limb size were accounted for.



Fig. 3: Limb identification

2.3.1 Statistics

Accuracy of limb identification

Total number of images	112
Number of limbs identified	391
False negatives	8
False positives	25
Error rate	8.4%

2.4 Silhouette extraction

Edge detection technique was used to extract the subject's silhouette. All three common edge detection technique - Sobel, Roberts, Susan was tested ultimately Roberts' provided the cleanest results. Sobel gave us some vibrant lines but it contained too much redundant data of the treadmill.

The result of Roberts' were able to outline our subject and some of the treadmill. We used the parameters of our limb segmentation to estimate our subject's location and ignored edges which were far away from the subject.

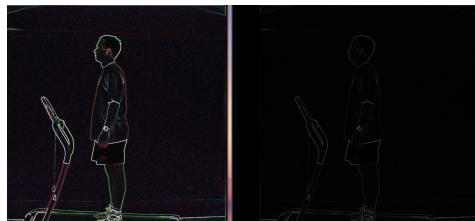


Fig. 4: Edge detection: Sobel (left) vs Roberts (right)

2.5 Results

The end result of both limb analysis and silhouette extraction has allowed us to gain a lot of metrics for further analysis into the subject's body shape as well as standing posture.



Fig. 5: Silhouette with limb analysis

Positional data of each subject including their limbs were stored in a separate JSON file to aid subject identification in later parts.

3 SUBJECT IDENTIFICATION

3.1 Domain

3.1.1 Silhouette data

From our silhouette extraction, we can measure the width of a subject and occasionally the height of the subject. A simple metric that we can measure is the surface area of the person, obviously this is not a great idea as the subject's surface area depends on how close the subject is standing (scaling problem). A better idea was to calculate the ratio of height against their width. This can give us a brief estimate of the person's body ratio, in other words if the person is short and wide or tall and skinny. However we could not assume this measurement was always available. The noise generated from the bottom of the image did not allow us to collect reliable information of the subject's feet, therefore we cannot accurately determine the subject's height. One way to get around the problem is to uniformly measure each subject to the lowest point of their leg. This on the other hand struggles from inconsistency between subject wearing trousers and shorts, which affects if the lower limb data is available. Nevertheless, this data is used whenever available during our identification phase.

3.1.2 Front facing limb analysis

Instead of analysing the ratio between the subject's height and width, limb positional analysis provides much better results. We compared two main metrics when applicable.

- horizontal position of where the head is relative to their arms. Everyone has a slightly different standing posture, this determines if one of your arm is extending further than the other.

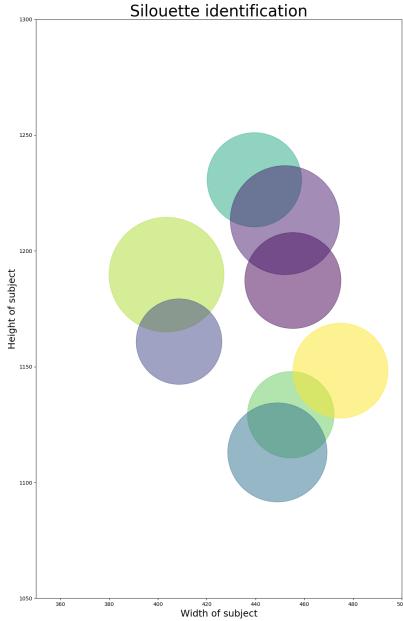


Fig. 6: Silhouette identification analysis

- the top position of the arm area. Again, depending on the subject's standing posture, one of their arm is naturally raised higher than the other. [3]

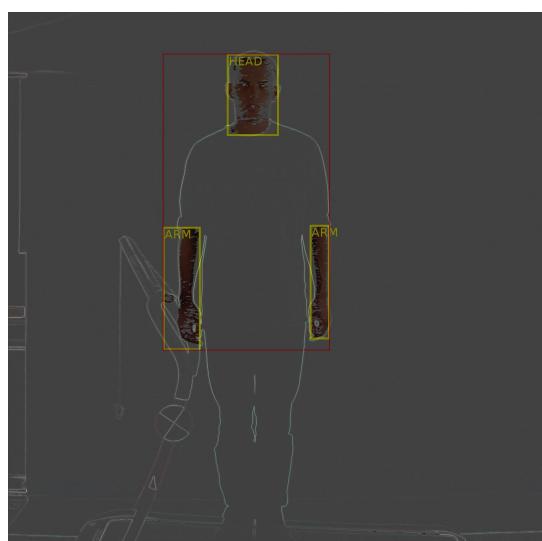


Fig. 7: Height difference of 5px between subject's arm

The diagram below demonstrate the head position being a better metrics than height-width ratio as there is no error bar involved. The image has been scaled by 25-times its original size for visual purposes.

After experimenting with vertical positional difference between arms, we discovered the maximum detected difference is only 12 pixel apart, therefore unsuitable for subject identification.

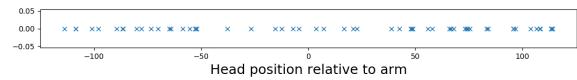


Fig. 8: Head position relative to arm (mid-points)

3.1.3 Side facing limb analysis

Side facing image analysis overall had much better results than front facing images, 51 out of 52 images identified all limbs correctly. This allowed us to run analysis with extra confidence.

4 CONCLUSION

The conclusion goes here.

ACKNOWLEDGMENTS

The authors would like to thank

REFERENCES

- [1] Marvin Project. 2018.
- [2] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat, and J. Jataklia. Human Skin Detection Using RGB, HSV and YCbCr Color Models. 2017.
- [3] Arpit Mittal, Andrew Zisserman, and Phil Torr. *Hand detection using multiple proposals*. 2011.