Skin Segmentation using a skin color model

Moses Modupi
1614669

October 1, 2020

School of Computer Science & Applied Mathematics
University of the Witwatersrand

Introduction

The objective of the skin segmentation is to separate regions of skin and non-skin in colour images. This is done by classifying each pixel in the image as either being skin or not skin. This problem can be applied to a wide range of problems in computer vision, robotics and human-computer interaction. Examples of this are face detection [Cai and Goshtasby 1999] and filtering crude images on the internet [Fleck *et al.* 1996]. Skin segmentation can be done using a skin colour model. The model should be able to classify a given pixel as either skin or non-skin. We expect the model to have low false acceptance rate and false rejection rate. This means the model has to be able to correctly classify a wide range of skin colours and be able to discard much of the environment.

Background Work

Several challenges plague the problem of skin segmentation. Using a skin colour model requires a labelled dataset of skin images. The dataset has to contain a varied sample of skin colours. The quality of the dataset will have an impact on the performance of the skin colour model. Particularly when it comes to the false rejection of skin pixels. Mahmoodi *et al.* [2015] presents benchmarks in performance regarding different datasets.

The next challenge is variations in lighting that can make skin pixels marked otherwise. The choice of colour space can help mitigate the effects of varying lighting. [Saini and Chand 2013] uses RGB colour space. However, the RGB colour space represents the colour and brightness of the image [Yang and Waibel 1996]. Yang and Waibel [1996] uses the Chromatic colour space. The HSV and YCbCr colour spaces are also options where we can isolate the brightness component of the pixels [Mahmoodi 2017].

The skin model can be based on the Bayesian decision rule for minimum cost Phung *et al.* [2003]. The Bayesian model classifies a pixel as skin if:

$$\frac{P(\mathbf{c}|skin)}{P(\mathbf{c}|nonskin)} = \frac{\lambda_{fd}}{\lambda_{fr}} \frac{P(nonskin)}{P(skin)} \tag{1}$$

Where λ_{fd}, λ_{fr} are the costs of false acceptance and false rejection respectively. This model then requires training on datasets of skin images and non-skin images. In contrast, the model could be in the form of a gaussian distribution that presents the likelihood of a pixel being skin [Yang and Waibel 1996].

Proposed Strategy

In this paper, we use the colour model presented in the literature presented by yang Yang and Waibel [1996]. This approach fits a Gaussian distribution to a dataset of skin colours. We will be fitting our distribution using the SFA dataset [Casati *et al.* 2013]. The SFA dataset contains a set of 1118 cropped skin images of size 35x35 pixels. To fit the dataset to a Gaussian, we first convert the dataset to the YCbCr colour space, discard the Y component, then find the mean colour and covariance matrix. The following plot shows the Gaussian I were able to fit on the SFA dataset.

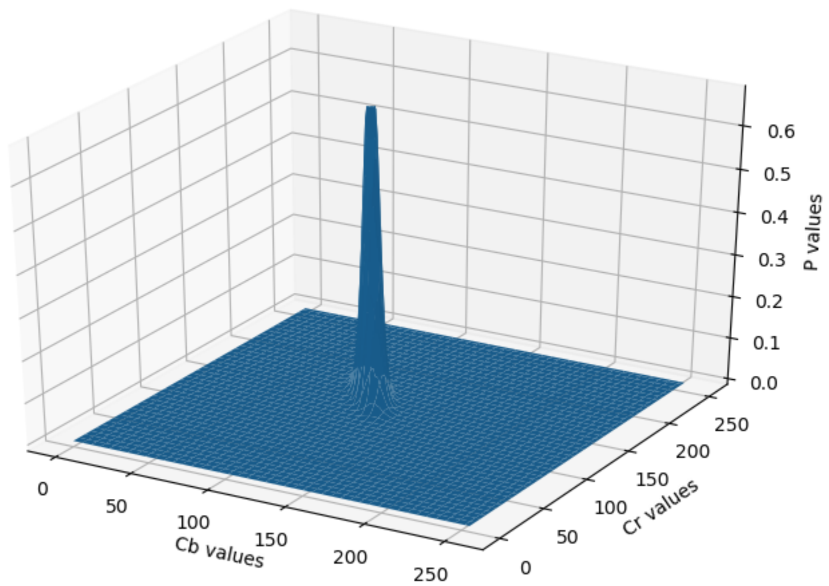

Figure 1: A Gaussian fit of the SFA skin dataset

From Figure 1, the high peak of the Gaussian suggests that the colour space of the human skin is concentrated on a small range of the colour space.

As we have mentioned. The colour model can fall prey to the effects of images with uneven lighting conditions. The introduction of the YCbCr colour space helps with this since the light intensity of colours is encoded in the Y component (otherwise referred to as the luma component) which can be discarded. This does not completely solve the problem. The next thing I tried is to reduce the variations in color intensities in the image using some of the region based segmentation methods. I applied the K-means clustering and Simple linear iterative clustering (SLIC) superpixel algorithm to reduce the variation of the image. These methods should have an averaging effect to the images. This prior segmentation was done on the RGB image. Figure 2 shows the original image as well as results of applying K-means and SLIC. As we can see, K-means clustering results in a blander image with less variation in color. SLIC segments the image in spatial segments and assigns colors of each pixel in a segment to the average pixel color. SLIC as the appeal of being more computationally efficient than K-means. Figure 3 shows an image where each pixel is the likelihood (between 0 and 1) that it is a skin pixel. The likelihood of color $\mathbf{x}$ is given by the Gaussian distribution in equation 2.

$$likelihood(\mathbf{x}) = e^{\textbf{(x-u)}^T \textbf{C(x-u)}} \tag{2}$$

Where **u, C** are the mean and Covarience matrix. The denominator constant is discarded since the determinant of the Covarience matrix is close to 0 and would make the likelihoods greater than 1. Figure 3d, 3e and 3f show the likelihood images of the original image, K-means segmented image and SLIC image. We see that the pixels on the cheeks are not recognised by the skin colour model when the origina image is used. K-means segmentation was able to assign the cheek pixels to some other colour the model can recognise. K-means was also able to give us higher likelihoods in the skin regions by assigning them to mean colors the model saw to be likely skin. SLIC only averages pixels in each segment. This means that the segments in cheeks will take on the same pixel colour which explains why it is ineffective. To turn the image into a binary image, we use the Otsu thresholding algorithm. The results are shown in figure 4. The last step is to remove blobs from the image. This is done by finding connected components from the image and removing components that have an area that is less than 5% of the largest component. We can see the effects in Figure 5.

Figure 2: a) Original Image, b) original segmented by K-means with 30 means. c) original segmented by slic with 964 segments

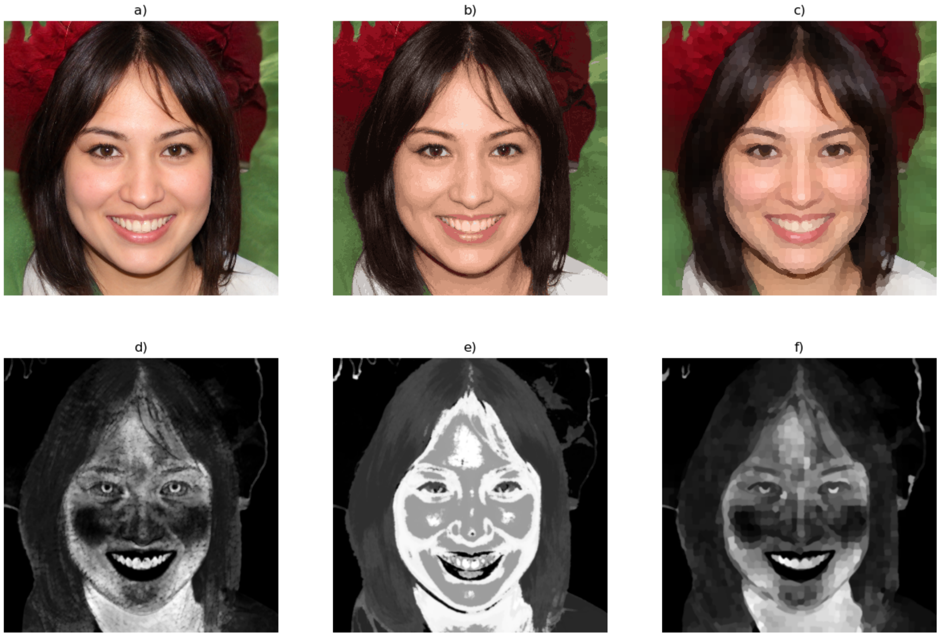

Figure 3: a) Original Image, b) original segmented by K-means with 30 means. c) original segmented by slic with 964 segments, d) Likelihood image of a, e) Likelihood image of b, f) Likelihood image of c

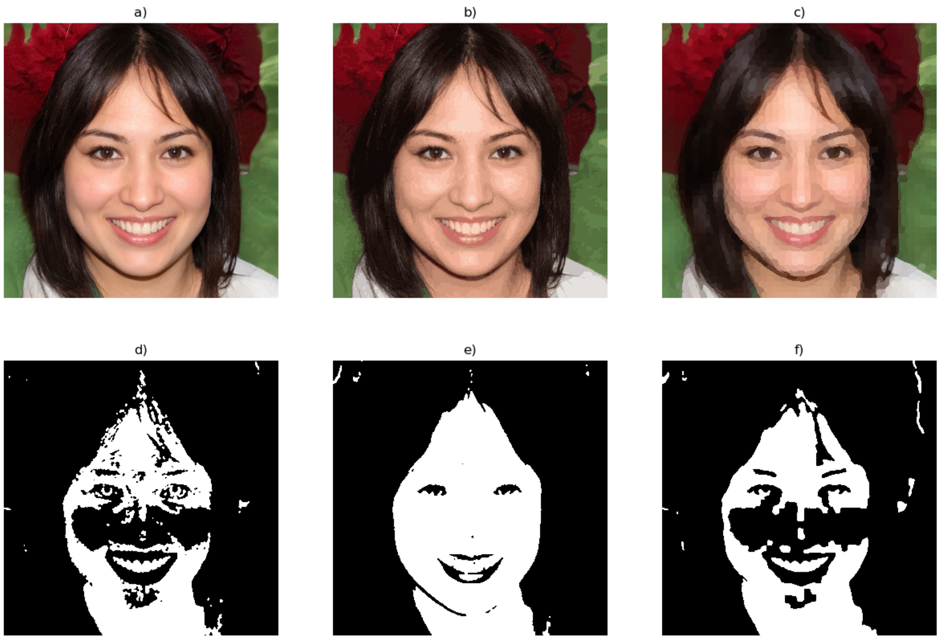

Figure 4: Binary images achieved by Otsu threshold

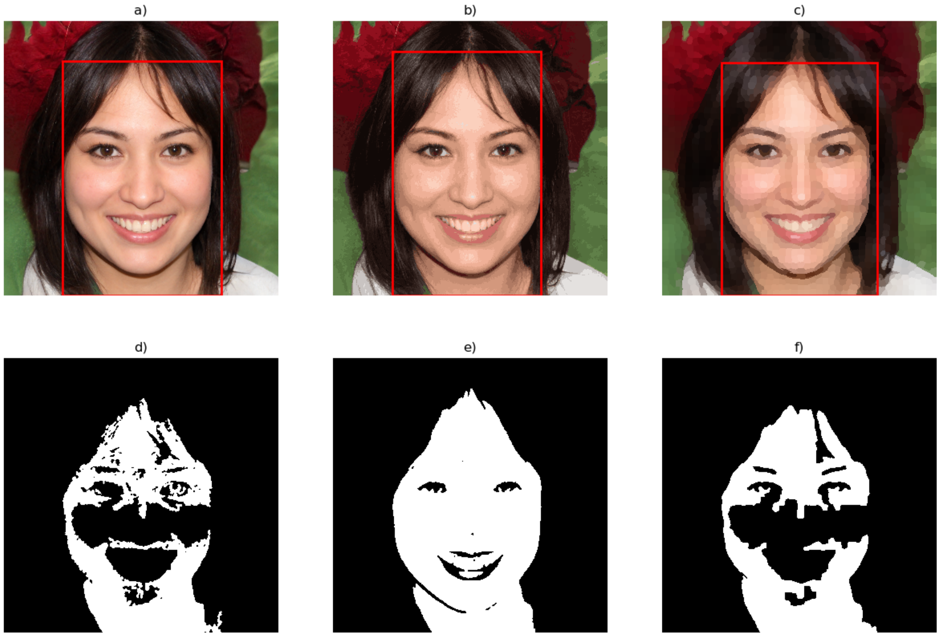

Figure 5: Final Segmentation

Results

The model was tested using a variety of images sourced from the internet. The images are grouped into three categories. The first category is for portrait images. The second is for images with 1 person but with skin exposure that extends further than just the face. The Last category is for group images. We only have 3 images from each category. The results from each can be seen in Figures 6, 7 and 8. In these figures:

- the first row shows the original images,
- the second row shows the segmented images with no prior segmentation routine applied
- the third row shows the segmented images with K-Means as the prior segmentation routine
- the fourth shows the segmented images with the SLIC applied as the prior segmentation routine.

From these figures, we can see that K-Means does fairly well compared to not applying prior segmentation as well as using SLIC. However, the averaging effect of SLIC does produce smoother results than applying our model to images with no prior segmentation. However, our model does not fair well in images will busy environments as we see in the results for group images. The model fails to distinguish some of the background elements as non-skin. I would seem as if future work in skin segmentation could benefit greatly from background/foreground separation. This would decrease the number of false acceptance we get from the background of the image.

Conclusion

In this project, I sought out to implement skin segmentation using a skin colour model. K-means clustering proved to be an effective way to deal with the problem of having varying lighting in images. This was achieved by applying segmentation to the input image using K-means. This method also proved to be computationally expensive and unsuitable for real-time systems. We attempted to use the more efficient SLIC algorithm instead. However, the resulting images did not bring much improvement on a prior segmentation free approach.

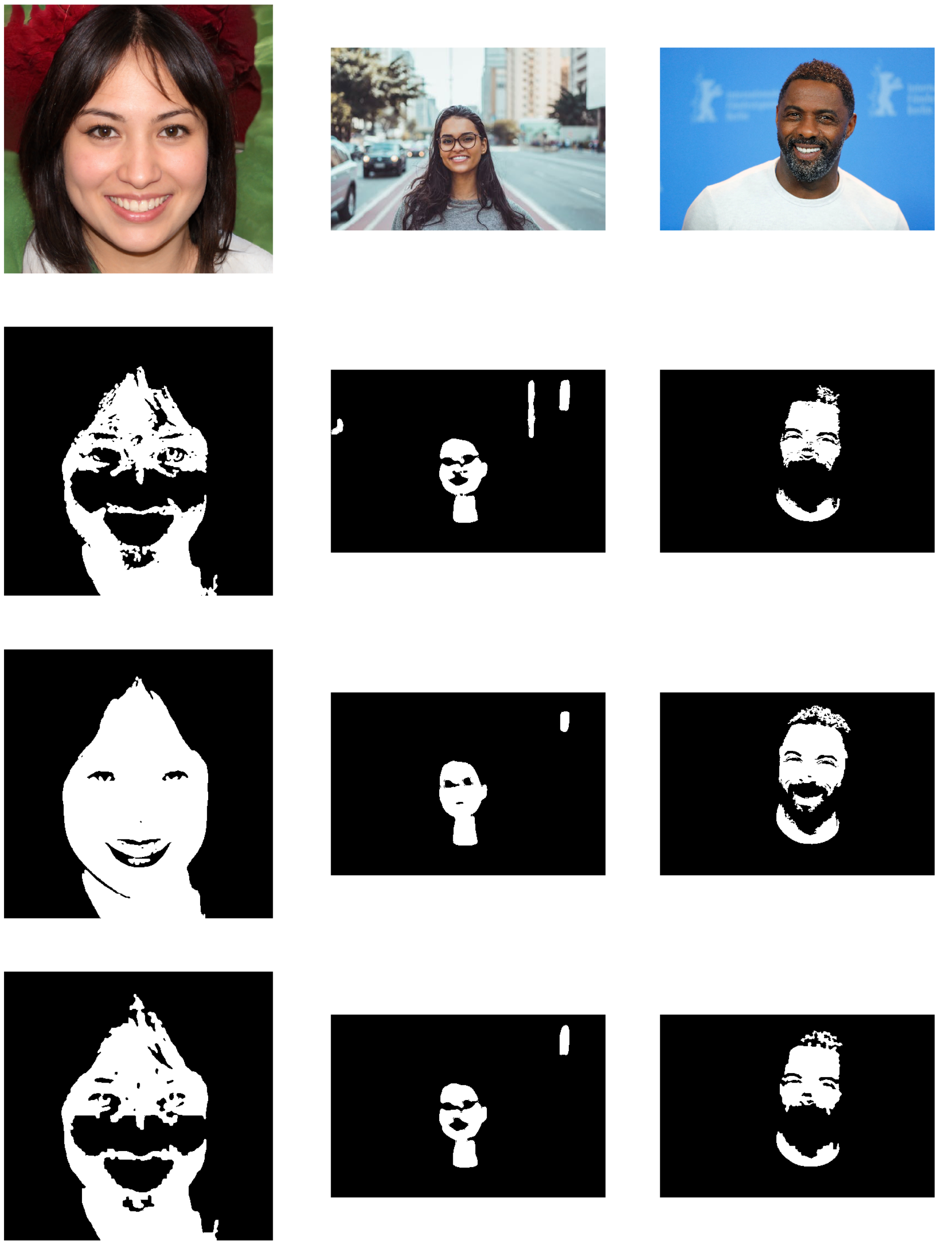

Figure 6: Segmentation on portrait images

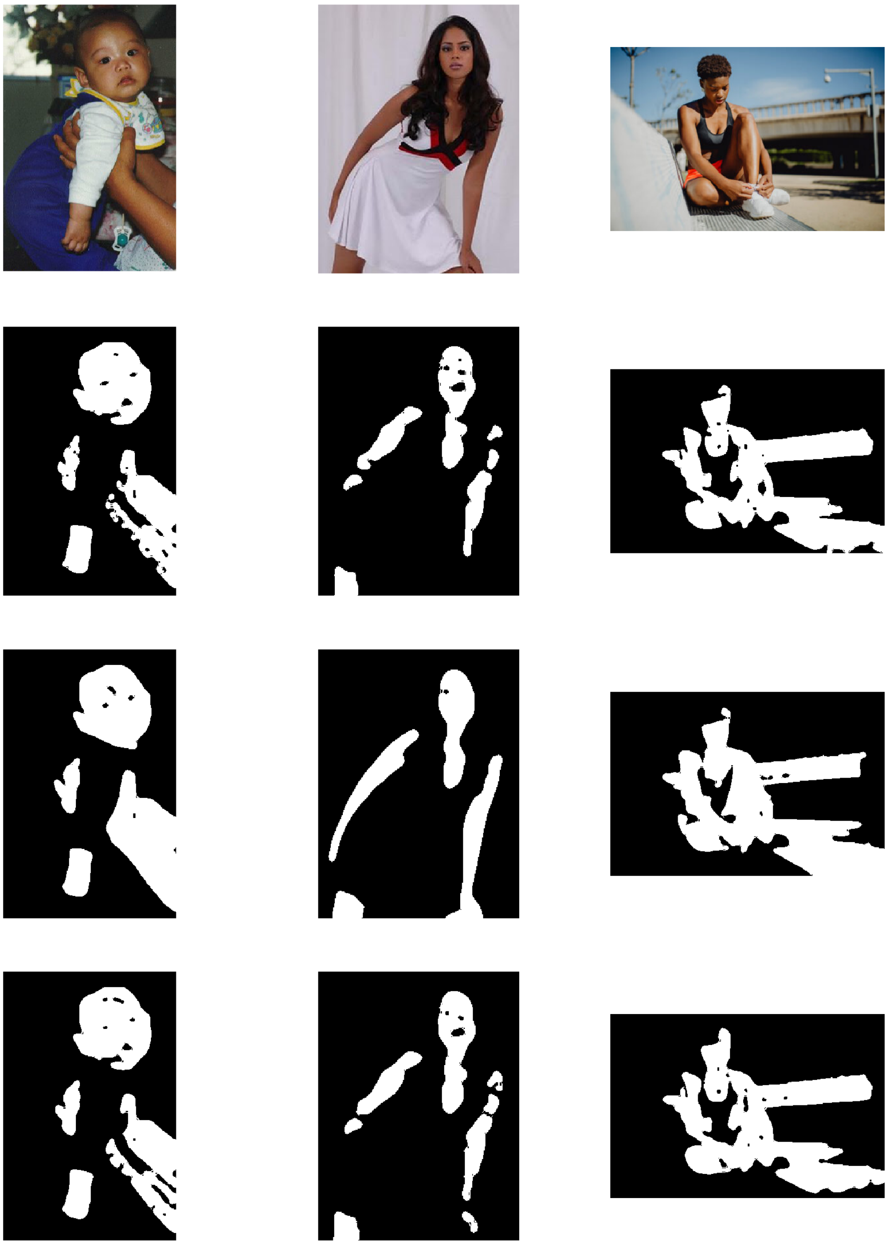

Figure 7: Segmentation on full body images

Figure 8: Segmentation on group images

References

[Cai and Goshtasby 1999] Junxia Cai and A Goshtasby. Detecting human faces in color images. *Image and Vision Computing*, 18(1):63–75, 1999.

[Casati *et al.* 2013] Joao Paulo Brognoni Casati, Diego Rafael Moraes, and Evandro Luis Linhari Rodrigues. Sfa: A human skin image database based on feret and ar facial images. In *IX workshop de Visao Computational, Rio de Janeiro*, 2013.

[Fleck *et al.* 1996] MM Fleck, DA Forsyth, and C Bregler. *Finding naked people, in 'ECCV'96: Proceedings of the 4th European Conference on Computer Vision-Volume II'*, 1996.

[Mahmoodi *et al.* 2015] M. R. Mahmoodi, S. M. Sayedi, F. Karimi, Z. Fahimi, V. Rezai, and Z. Mannani. Sdd: A skin detection dataset for training and assessment of human skin classifiers. In *2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, pages 71–77, 2015.

[Mahmoodi 2017] Mohammad Reza Mahmoodi. High performance novel skin segmentation algorithm for images with complex background. *arXiv preprint arXiv:1701.05588*, 2017.

[Phung *et al.* 2003] Son Lam Phung, Abdesselam Bouzerdoum, and Douglas Chai. Skin segmentation using color and edge information. In *Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings.*, volume 1, pages 525–528. IEEE, 2003.

[Saini and Chand 2013] Harpreet Kaur Saini and Onkar Chand. Skin segmentation using rgb color model and implementation of switching conditions. *Skin*, 3(1):1781–1787, 2013.

[Yang and Waibel 1996] Jie Yang and Alex Waibel. A real-time face tracker. In *Proceedings Third IEEE Workshop on Applications of Computer Vision. WACV'96*, pages 142–147. IEEE, 1996.