Skin Detection [1]

Qi Zhao

July 28, 2018

1. Introduction

The goal of skin-detection algorithms is to infer a label w $\in$ {0,1} denoting the presence or absence of skin at a pixel, based on the RGB [2] measurements x = [x^R, x^G, x^B] at that pixel. This is a useful precursor to segmenting a face or hand, or it may be used as the basis of a crude method for detecting prurient content in web images. Taking a generative approach, it can be described the likelihoods as in Equation 1.

$$Pr(x|w = k) = Norm_x[\mu_k, \sigma_k] \tag{1}$$

and the prior probability [4] over states as in Equation 2.

$$Pr(w) = Bern_w[\lambda] \tag{2}$$

In the learning algorithm, it estimates the parameters μ_0, μ_1, Σ_0, Σ_1 from training data pairs $\{w_i, x_i\}_{i=1}^{I}$ where the pixels have been labeled by hand. In particular we learn μ_0 and Σ_0 from the subset of the training data where w_i = 0 and μ_1 and Σ_1 from the subset where w_i = 1. The prior parameter is learned from the world states $\{w_i\}_{i=1}^{I}$ alone.

To classify a new data point x as skin or non-skin it can apply Bayes rule [3] in Equation 3 and denote this pixel as skin if Pr(w = 1|x) $>$ 0.5. Figure 1 shows the result of applying this model at each pixel independently in the image. Note that the classification is not perfect: there is genuinely an overlap between the skin and non-skin distributions and this inevitably results in misclassified pixels. The results could be improved by exploiting the fact that skin areas tend to be contiguous regions without small holes.

$$Pr(w = 1|x) = \frac{Pr(x|w = 1)Pr(w = 1)}{\sum_{k=0}^{1} Pr(x|w = k)Pr(w = k)} \tag{3}$$

The RGB data are naturally discrete with x^R, x^G, x^B $\in$ {0,1,...,255}, which is basic of the skin detection model on this assumption. For example, modeling the three color channels independently, the likelihoods become in Equation 4.

$$Pr(x|w = k) = Cat_{x^R}[\lambda_k^R]Cat_{x^G}[\lambda_k^G]Cat_{x^B}[\lambda_k^B] \tag{4}$$

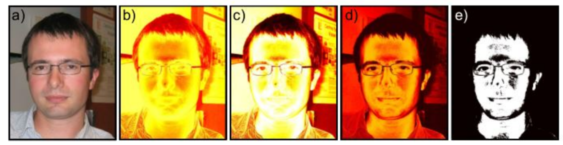

Figure 1. Skin detection. For each pixel we aim to infer a label w $\in$ {0,1} denoting the absence or presence of skin based on the RGB triple x. Here we modeled the class conditional density functions Pr(x|w) as normal distributions. a) Original image. b) Log likelihood (log of data assessed under class-conditional density function) for non-skin. c) Log likelihood for skin. d) Posterior probability of belonging to skin class. e) Thresholded posterior probability Pr(w|x) $>$ 0.5 gives estimate of w.

2. Conclusions

Assume that the elements of the data vector are independent as naïve Bayes. Of course, it is not necessarily valid in the real world. To model the joint distribution [5] of the R,G, and B components, it might combine them to form one variable with 256^3 entries and model this with a single categorical distribution. And it is more practical to quantize each channel to fewer levels (say 8) before combining them together.

References

[1] A. Albiol, L. Torres, and E. J. Delp. Optimum color spaces for skin detection. In *IEEE International Conference on Image Processing*, pages 122–124, 2001. 1

[2] L. Bo, X. Ren, and D. Fox. *Unsupervised feature learning for RGB-D based object recognition*. Springer, 2013. 1

[3] D. M. Grether. Bayes' rule as a descriptive model: the representativeness heuristic. *The Quarterly Journal of Economics*, 95(3):537–557, 1980. 1

[4] H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London*, 186(1007):453–461, 1946. 1

[5] M. S. Longuet-Higgins. On the joint distribution of the periods and amplitudes of sea waves. *Journal of Geophysical Research*, 80(18):2688–2694, 1975. 1