Segmentation [5]

Qi Zhao

September 2, 2018

1. Introduction

The goal of segmentation is to assign a discrete label $\{w_n\}_{n=1}^{N}$ which takes one of K values $w_n \in \{1,2,...,K\}$ to each of the N pixels in the image so that regions that belong to the same object are assigned the same label. The segmentation model depends on observed data vectors $\{x_n\}_{n=1}^{N}$ at each of the N pixels that would typically include the RGB pixel [2] values, the (x,y) position of the pixel and other information characterizing local texture.

It will frame this problem as unsupervised learning [1]. In other words, and it does not have the luxury of having training images where the state of the world is known. It must both learn the parameters θ and estimate the world states $\{w_i\}_{i=1}^{I}$ from the image data $\{x_n\}_{n=1}^{N}$.

To fit this model, the parameters $\theta = \{\lambda_k, mu_k, \Sigma_k\}_{k=1}^{K}$ use the EM algorithm [4]. To assign a class to each pixel, it then finds the value of the world state that have the highest posterior probability given the observed data in Equation 1.

$$\hat{w_i} = argmax_{w_i}[Pr(w_i|x_i)] \tag{1}$$

Figure 1 shows results from this model and a similar mixture model based on t-distributions [3] from Sfikas et al. (2007). The mixture models manage to partition the image quite well into different regions. Unsurprisingly, the t-distribution results are rather less noisy than those based on the normal distribution.

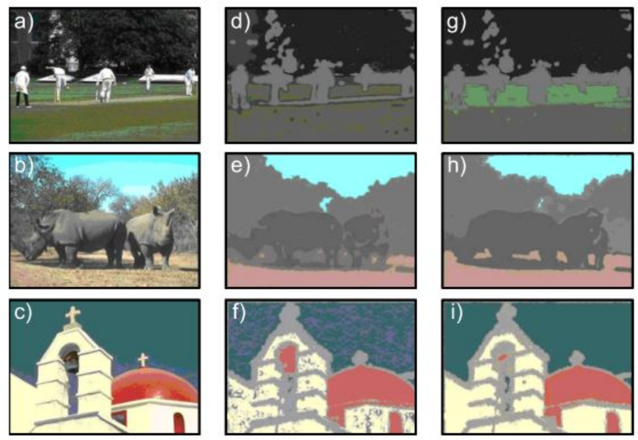

Figure 1. Segmentation. a-c) Original images. d-f) Segmentation results based on a mixture of five normal distributions. The pixels associated with the k^{th} component are colored with the mean RGB values of the pixels that are assigned to this value g-i) Segmentation results based on a mixture of K t-distributions. The segmentation here is less noisy than for the MoG model. Results from Sfikas et al. (2007) IEEE 2007.

References

[1] T. Hofmann. Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 42(1-2):177–196, 2001. 1

[2] Q. A. Kester. Image encryption based on the RGB pixel transposition and shuffling. *International Journal of Computer Network & Information Security*, 5(7):43–50, 2013. 1

[3] D. Li, K. Ota, M. Dong, and Y. Li. Traces and t-distribution based multi-copy routing in delay tolerant networks. In *International Conference on Mobile Ad-Hoc and Sensor Networks*, pages 222–227, 2017. 1

[4] X. Li, L. Girin, and R. Horaud. An EM algorithm for audio source separation based on the convolutive transfer function. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 56–60, 2017. 1

[5] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 39(4):640–651, 2017. 1