Multivariate Gaussian Modelling of Object Colors for Image Analysis

Mehmet Celenk and Maarten Uijt de Haag
School of Electrical Engineering and Computer Science
Stocker Center, Ohio University, Athens, Ohio 45701 USA

Abstract

This paper describes an automatic object detection method for color scenes by modeling the object and the background colors with additive Gaussian noise. Image color histogram is treated as an estimate of the 3-D probability density function p(**x**) of image color vector $\mathbf{x}=[R\ G\ B]^t$, where t denotes the matrix transposition. For an image of a background and an object with the 3-D color distributions $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$, $p(\mathbf{x})$ is the sum or mixture of these two unimodal densities given by $p(\mathbf{x})=P_1p_1(\mathbf{x})+P_2p_2(\mathbf{x})$, where P_1 and P_2 are the a priori probabilities of the two colors. Here, $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ are modeled with additive 3-D Gaussian noise with two unknown parameters (i.e., mean vectors, μ_1 and μ_2, and covariance matrices, Σ_1 and Σ_2). The symmetry of 3x3 covariance matrices allows complete specification of the 3-D multivariate density $p(\mathbf{x})$ by 19 unknown parameters. The unknown parameters (μ's and Σ's) are estimated by minimizing the mean square error $e_{ms}(\mathbf{x})$ between the 3-D model $p(\mathbf{x})$ and the experimental 3-D histogram $h(\mathbf{x})$ obtained from the image data. A threshold vector $\mathbf{T}=[T_R\ T_G\ T_B]^T$ is defined so that all pixels with a color level below $|\mathbf{T}|$ are considered background points and all pixels with a level above $|\mathbf{T}|$ are considered object points. $\mathbf{T}$ is computed by minimizing the overall probability $e(\mathbf{T})$ of erroneously classifying image pixels along the R, G, and B color axes. The possibility of two solutions for T_R, T_G, and T_B indicates that two threshold values in each color dimension may be required to obtain the optimal boundary detection result. The proposed algorithm is tested using color images of various natural scenes, aerial photographs, indoor and outdoor pictures, and images downloaded from the internet. Most of these images contain complex shaped objects and texture surfaces. Computer experiments show that even for noisy pictures, the algorithm accurately separates the object from the background without human intervention.

Introduction and Background Information

Thresholding and clustering are two of the most commonly used approaches to image segmentation. Although the former method works best for gray-level images, its extention to color pictures has proven difficult. The major difficulty in this extension is the multi-dimensionality of the color data. Many techniques have been developed for image thresholding, each with its own merits and limitations.[1,2,3] Because of its mathematical tractability, Gaussian modeling of color distribution of images is highly desirable. In this regard, we describe an automatic object detection scheme applicable for a wide variety of imaging applications.

Description of the Algorithm

In the proposed method, we consider an image color histogram as an estimate of the 3-D color probability density function $p(\mathbf{x})$, where $\mathbf{x}=[R\ G\ B]^t$ is the 3x1 image color vector and t denotes the matrix transposition. It is assumed that a given image contains a background and an object region with the 3-D color distributions $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$. Based on this assumption we define $p(\mathbf{x})$ as the sum or mixture of these two unimodal densities. From the basic probability theory $p(\mathbf{x})$ can be expressed in a form

$$p(\mathbf{x})=P_1p_1(\mathbf{x})+P_2p_2(\mathbf{x}) \tag{1}$$

where P_1 and P_2 are the a priori probabilities of the two colors. Here, $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ are modeled with additive 3-D Gaussian noise with two unknown parameters (i.e., mean vectors, μ_1 and μ_2, and covariance matrices, Σ_1 and Σ_2). The underlying multivariate unimodal normal densities are given by

$$p_1(\mathbf{x})=(2\pi)^{-3/2}|\Sigma_1|^{-1/2}\exp[(\mathbf{x}-\mu_1)^T\Sigma_1^{-1}(\mathbf{x}-\mu_1)] \tag{2}$$

$$p_2(\mathbf{x})=(2\pi)^{-3/2}|\Sigma_2|^{-1/2}\exp[(\mathbf{x}-\mu_2)^T\Sigma_2^{-1}(\mathbf{x}-\mu_2)] \tag{3}$$

This model is of particular interest because of its mathematical tractability. The symmetry of 3x3 covariance matrices allows complete specification of the 3-D multivariate density $p(\mathbf{x})$ by 19 unknown parameters (i.e., 9 for the μ_1 and Σ_1 of the background, 9 for the μ_2 and Σ_2 of the object, and 1 for the a priori probabilities since

$$P_1 + P_2 = 1 \tag{4}$$

By minimizing the mean square error $e_{ms}(\mathbf{x})$ between the 3-D model $p(\mathbf{x})$ and the experimental 3-D histogram $h(\mathbf{x})$ (obtained from the image data) we estimate the unknown parameters (μ's and Σ's).

If all the parameters are known, a quadratic optimal decision surface can be formed in the color space so that the object area can be separated from the background region. The underlying parameters (mean vector and covariance matrix) may be estimated in the maximum likelihood (ML) sense. However, this requires a priori knowledge about the scene in question, which may not be available prior to the operation of an automatic object detection system.

Instead, we propose a method that relies on slicing this 3-D model distribution along the respective color coordinate axes resulting in three one-dimensional (1-D) mixture normal distributions, each corresponding to a particular color component (i.e., red, green, and blue), as given by

$$p(R) = P_1 p_1(R) + P_2 p_2(R) \tag{5a}$$

$$p(G) = P_1 p_1(G) + P_2 p_2(G) \tag{5b}$$

$$p(B) = P_1 p_1(B) + P_2 p_2(B) \tag{5c}$$

where $p_1(R)$, $p_1(G)$, $p_1(B)$ are the 1-D Gaussian distributions of the background and $p_2(R)$, $p_2(G)$, $p_2(B)$ are those of the object, respectively. The underlying color density functions for the background (p_1) and the object (p_2) are given by

$$p_1(R)=(2\pi)^{-1/2}(\sigma_{R1})^{-1}\exp\{-(R-\mu_{R1})^2/2\sigma_{R1}^{\ 2}\} \tag{6a}$$

$$p_1(G)=(2\pi)^{-1/2}(\sigma_{G1})^{-1}\exp\{-(G-\mu_{G1})^2/2\sigma_{G1}^{\ 2}\} \tag{6b}$$

$$p_1(B)=(2\pi)^{-1/2}(\sigma_{B1})^{-1}\exp\{-(B-\mu_{B1})^2/2\sigma_{B1}^{\ 2}\} \tag{6c}$$

$$p_2(R)=(2\pi)^{-1/2}(\sigma_{R2})^{-1}\exp\{-(R-\mu_{R2})^2/2\sigma_{R2}^{\ 2}\} \tag{7a}$$

$$p_2(G)=(2\pi)^{-1/2}(\sigma_{G2})^{-1}\exp\{-(G-\mu_{G2})^2/2\sigma_{G2}^{\ 2}\} \tag{7b}$$

$$p_2(B)=(2\pi)^{-1/2}(\sigma_{B2})^{-1}\exp\{-(B-\mu_{B2})^2/2\sigma_{B2}^{\ 2}\} \tag{7c}$$

where $(\mu_{R1}, \sigma_{R1}^{\ 2})$, $(\mu_{G1}, \sigma_{G1}^{\ 2})$, $(\mu_{B1}, \sigma_{B1}^{\ 2})$ and $(\mu_{R2}, \sigma_{R2}^{\ 2})$, $(\mu_{G2}, \sigma_{G2}^{\ 2})$, $(\mu_{B2}, \sigma_{B2}^{\ 2})$ are the unknown means and variances of the R, G, and B model distributions, respectively. Notice that this system has a total of 13 unknowns; i.e., means ($\mu_{R1},\mu_{G1},\mu_{B1}$ and $\mu_{R2},\mu_{G2},\mu_{B2}$) and variances ($\sigma_{R1}^{\ 2}, \sigma_{G1}^{\ 2}, \sigma_{B1}^{\ 2}$ and $\sigma_{R2}^{\ 2}, \sigma_{G2}^{\ 2}, \sigma_{B2}^{\ 2}$) of the background and the object and the a priori probabilities, which is 6 parameters less than that of the 3-D mixture distribution model.

By minimizing the mean square (MS) error ($e_{ms}(R)$,$e_{ms}(G)$,$e_{ms}(R)$) between the mixture 1-D densities, p(R), p(G), p(B), and the 1-D histograms, h(R), h(G), h(B), obtained from the image data we estimate the unknown parameters (μ's and σ^2's) for the 1-D mixture model, which leads to three optimal thresholds (T_R,T_G,T_B), one per color component, for automatic image thresholding. The MS equations are given by

$$e_{ms}(R) = \Sigma^{n}_{\ j=1}\ [p(R_j)-h(R_j)]^2/n \tag{8a}$$

$$e_{ms}(R) = \Sigma^{n}_{\ j=1}\ [p(G_j)-h(G_j)]^2/n \tag{8b}$$

$$e_{ms}(B) = \Sigma^{n}_{\ j=1}\ [p(B_j)-h(B_j)]^2/n \tag{8c}$$

where an n-point histogram is assumed. In general, analytically determining the above mentioned parameters that minimize the MS error is not a simple matter. Even for the Gaussian case, the straight-forward computation of equating the partial derivatives to 0 leads to a set of simultaneous transcendental equations that usually can be solved only by numerical procedures. Since the gradient is easily computed, a conjugate gradient or Newton's method for simultaneous nonlinear equations may be used to minimize e_{ms}'s. With either of these iterative methods, starting values must be specified. Assuming the a priori probabilities to be equal, P_1=P_2, are sufficient. Starting values for the means and variances are determined by detecting modes in the red, green, and blue histograms or simply by dividing these color histograms into two parts about their mean values, and computing means and variances of the two parts to be used as starting values.

Here we assume, without loosing generality, that the dark colored regions correspond to the background and the bright colored regions correspond to objects. In this case $\mu_{R1}<\mu_{R2}$, $\mu_{G1}<\mu_{G2}$, $\mu_{B1}<\mu_{B2}$ and a set of thresholds (T_R,T_G,T_B) are defined so that all pixels with a color level below (T_R,T_G,T_B) are considered background points and all pixels with a level above (T_R,T_G,T_B) are considered object points. To determine the desired threshold values, we define the probability of (erroneously) classifying an object point as a background point by the integral equations

$$E_{R1}(T_R) = {}_{-\infty}\!\int^{T}\!R\ p_2(R)dR \tag{9a}$$

$$E_{G1}(T_G) = {}_{-\infty}\!\int^{T}\!G\ p_2(G)dG \tag{9b}$$

$$E_{B1}(T_B) = {}_{-\infty}\!\int^{T}\!B\ p_2(B)dB \tag{9c}$$

Similarly, the probability of classifying a background point as an object point may be written as

$$E_{R2}(T_R) = {}_{-\infty}\!\int^{T}\!R\ p_1(R)dR \tag{10a}$$

$$E_{G2}(T_G) = {}_{-\infty}\!\int^{T}\!G\ p_1(G)dG \tag{10b}$$

$$E_{B2}(T_B) = {}_{-\infty}\!\int^{T}\!B\ p_1(B)dB \tag{10c}$$

Therefore the overall probability of error is defined by

$$E(T_R) = P_2 E_{R1}(T_R) + P_1 E_{R2}(T_R) \tag{11a}$$

$$E(T_G) = P_2 E_{G1}(T_G) + P_1 E_{G2}(T_G) \tag{11b}$$

$$E(T_B) = P_2 E_{B1}(T_B) + P_1 E_{B2}(T_B) \tag{11c}$$

To find the threshold values for which this overall error is minimal requires

$$dE(T_R)/dT_R = 0 \tag{12a}$$

$$dE(T_G)/dT_G = 0 \quad (12b)$$

$$dE(T_B)/dT_B = 0 \quad (12c)$$

This yields

$$P_1 p_1(R) = P_2 p_2(R) \quad (13a)$$

$$P_1 p_1(G) = P_2 p_2(G) \quad (13b)$$

$$P_1 p_1(B) = P_2 p_2(B) \quad (13c)$$

Applying these results to the Gaussian densities, taking logarithms, and simplifying, gives the quadratic equations

$$A_R T_R^2 + B_R T_R + C_R = 0 \quad (14a)$$

$$A_G T_G^2 + B_G T_G + C_G = 0 \quad (14b)$$

$$A_B T_B^2 + B_B T_B + C_B = 0 \quad (14c)$$

where

$$A_R = \sigma_{R1}^2 - \sigma_{R2}^2$$

$$B_R = 2(\mu_{R1}\sigma_{R2}^2 - \mu_{R2}\sigma_{R1}^2)$$

$$C_R = \mu_{R2}^2\sigma_{R1}^2 - \mu_{R1}^2\sigma_{R2}^2 + 2\sigma_{R1}^2\sigma_{R2}^2 \ln(\sigma_{R2}P_1/\sigma_{R1}P_2)$$

$$A_G = \sigma_{G1}^2 - \sigma_{G2}^2$$

$$B_G = 2(\mu_{G1}\sigma_{G2}^2 - \mu_{G2}\sigma_{G1}^2)$$

$$C_G = \mu_{G2}^2\sigma_{G1}^2 - \mu_{G1}^2\sigma_{G2}^2 + 2\sigma_{G1}^2\sigma_{G2}^2 \ln(\sigma_{G2}P_1/\sigma_{G1}P_2)$$

$$A_B = \sigma_{B1}^2 - \sigma_{B2}^2$$

$$B_B = 2(\mu_{B1}\sigma_{B2}^2 - \mu_{B2}\sigma_{B1}^2)$$

$$C_B = \mu_{B2}^2\sigma_{B1}^2 - \mu_{B1}^2\sigma_{B2}^2 + 2\sigma_{B1}^2\sigma_{B2}^2 \ln(\sigma_{B2}P_1/\sigma_{B1}P_2)$$

The possibility of two solutions for T_R, T_G, and T_B indicates that two threshold values in each color component may be required to obtain the optimal boundary detection result. Three special cases produce the following single set of closed form solutions:

Case 1:

If the variances are equal, $\sigma_R^2=\sigma_{R1}^2=\sigma_{R2}^2$, $\sigma_G^2=\sigma_{G1}^2=\sigma_{G2}^2$, $\sigma_B^2=\sigma_{B1}^2=\sigma_{B2}^2$, a single set of thresholds is sufficient:

$$T_R = (\mu_{R1}+\mu_{R2})/2 + [\sigma_R^2/(\mu_{R1}-\mu_{R2})]\ln(P_2/P_1) \quad (15a)$$

$$T_G = (\mu_{G1}+\mu_{G2})/2 + [\sigma_G^2/(\mu_{G1}-\mu_{G2})]\ln(P_2/P_1) \quad (15b)$$

$$T_B = (\mu_{B1}+\mu_{B2})/2 + [\sigma_B^2/(\mu_{B1}-\mu_{B2})]\ln(P_2/P_1) \quad (15c)$$

Case 2:

If the prior probabilities are equal, $P_1=P_2$, the optimal thresholds are the average of the means:

$$T_R = (\mu_{R1} + \mu_{R2})/2 \quad (16a)$$

$$T_G = (\mu_{G1} + \mu_{G2})/2 \quad (16b)$$

$$T_B = (\mu_{B1} + \mu_{B2})/2 \quad (16c)$$

Case 3:

If the color variances are all zero, $\sigma_R^2=\sigma_G^2=\sigma_B^2=0$, the same results as in the case 2 are obtained.

Assuming single thresholds, the object in question is detected by assigning a value of 1 to the image points (i,j) whose color values are above the threshold values; i.e., $R(i,j)\geq T_R$, $G(i,j)\geq T_G$, and $B(i,j)\geq T_B$. Similarly, the background area is extracted by labeling the image pixels (i,j) with 0, whose color values are below the thresholds; i.e., $R(i,j)<T_R$, $G(i,j)<T_G$, and $B(i,j)<T_B$. This is summarized by the following classification rules:

$$(i,j) \in \omega_1 \quad \text{if } R(i,j)\geq T_R, G(i,j)\geq T_G, B(i,j)\geq T_B \quad (17a)$$

$$(i,j) \in \omega_2 \quad \text{if } R(i,j)<T_R, G(i,j)<T_G, B(i,j)<T_B \quad (17b)$$

Here ω_1 and ω_2 represent the object and background pixel sets in the image plane, respectively.

Experimental Results and Conclusions

The proposed algorithm is tested using color images of various natural scenes, aerial photographs, indoor and outdoor pictures, and images downloaded from the internet. Most of these images contain complex shaped objects and texture surfaces. Computer experiments show that even for noisy pictures, the algorithm accurately separates the object from the background without human intervention. This verifies the basic premise that the 1-D Gaussian approximation to object and background color densities for the selected set of images is adequate. Furthermore, we also conclude as an empirical evidence that the minimum mean square approach can be used effectively to estimate the parameters (means and variances) of a color image from its respective color histograms.

References

1. R. C. Gonzalez and R. E. Woods, Digital image processing, Addison-Wesley, Reading: MA, 1992, pg.443-458.
2. R. M. Haralick and L. G. Shapiro, Computer and robot vision, Vol. I, Addison-Wesley, Reading: MA, 1992, pg.14-28.
3. M. Celenk and M. Uijt de Haag, "Optimal thresholding for color images," *Proc. SPIE Nonlinear Image Processing IX*, Vol. **3304**, pp.250-259 (1998).