APPENDIX

The representation of color C

Color input for computer vision is normally in RGB (red, green, blue) format, with typically one byte of data for each of the three input channels. However, just as the eye perceives color images subjectively as color, so it is convenient to match this using a suitable computer representation, such as HSI (hue, saturation, intensity). In HSI, the color is represented entirely by the hue channel; the brightness is represented by the intensity channel; and the saturation channel represents the *degree of coloration* present in the original RGB signal—i.e., the extent to which the color is diluted by white light.

Look out for:

- formulae defining the H, S, I parameters
- the periodic nature of the hue parameter and the complications it leads to
- the problems of achieving color constancy
- applications for which color-based segmentation is particularly beneficial.

It often happens that images have varying intensity because of random changes in illumination or because of shadows across the field of view. In such cases it can be useful to ignore the intensity and to rely on the hue parameter. More generally, it is desirable to utilize the whole HSI color space so that maximum information is available for color-based object segmentation.

C.1 INTRODUCTION

In the early days of computer vision, digitizers usually inputted gray-scale images and thus processing was restricted to the analysis of gray-scale images. However, over the past two decades, color input has become almost universal and processing capabilities have evolved to match this requirement. Color input is normally in RGB format, with typically one byte of data for each of the three input channels. However, just as the eye perceives color images subjectively as color, so it is convenient to match this via a suitable computer representation. In fact, a great variety of color representations can be used for the purpose, depending on the requirements and whether, for example, the results are to be presented to a computer image analyzer, a display or a printer. Here we consider only computer analysis, and choose the HSI representation, which is well adapted to the task. In particular, when converting from RGB to HSI, the actual color is represented entirely by the hue channel; the brightness is represented by the intensity channel;

this leaves the saturation channel to represent the *degree of coloration* present in the original RGB signal: i.e., the saturation channel shows the extent to which the color is diluted by white light.

It often happens that images have varying intensity because of random changes in illumination or because of shadows across the field of view. In such cases it can be useful to ignore the intensity information and to provide an interpretation based on color alone: in that case it can be more reliable to base an interpretation on the data issuing from the hue channel alone. This is the strategy that is normally adopted in applications such as inspection, surveillance, and face recognition when undesired intensity variations are to be expected.

However, more sophisticated solutions are possible and in principle necessary. This is because variations in the color of the ambient light can themselves lead to different apparent colors in the light reflected from objects. In spite of this, the human eye exhibits the property of *color constancy*, reflected light *not* appearing to change as a result of variations in the color of the ambient light. Though arduous, it is possible to achieve this by computer processing, but we shall not explore this possibility further here.

C.2 DETAILS OF THE HSI COLOR REPRESENTATION

As indicated above, it is common practice to concentrate on the hue parameter when the intensity parameter varies significantly because of random changes in illumination or because of shadows that may move across the field of view. Essentially, the hue parameter is independent of the intensity parameter as it varies in a plane in color space orthogonal to the intensity vector, so it holds the most undistorted and therefore the most meaningful information.

However, more generally, it is desirable to employ the whole HSI color space, so that the maximum information is available for image interpretation and in particular for color-based object segmentation. For example, we might wish to locate road signs by their colors, tomatoes by their red skins, or human faces using skin color recognition.

To carry out such tasks, we must first convert the incoming data from RGB to HSI format. First, we need to define the three HSI parameters, which we do using the formulae presented below. The most straightforward definition is that of intensity I: this is the mean light intensity, which is given by

$$I = \frac{1}{3}(R + G + B) \tag{C.1}$$

Hue H is a measure of the underlying color, and saturation S is a measure of the degree to which the color is *not* diluted by white light (S is zero for white light and is unity for *least* dilution of the colors). S is given by the following formula:

$$S = 1 - \frac{\min(R, G, B)}{I} = 1 - \frac{3 \min(R, G, B)}{R + G + B} \tag{C.2}$$

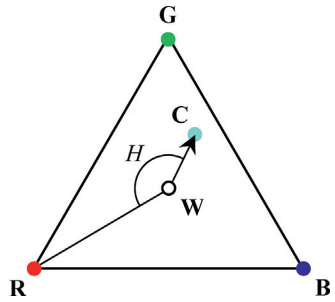

FIGURE C.1

View of the color triangle **RGB**. This contains all possible colors in all possible dilutions with white light. It is suspended between its three corners, which are points on the three color axes, indicated here by the colored dots marked as vectors **R**, **G**, **B**. White light is indicated by the vector **W** at the center of the color triangle. An arbitrary color **C** is located by hue angle H relative to the vector **R**–**W**. The diagram is 2-D and is spanned by the hue angle H and by the saturation level S (see text). Intensity I does not appear in the diagram as the latter is drawn for a constant intensity: in fact, intensity variation takes place along an axis orthogonal to the color triangle, passing through the white spot **W** at the center of the color triangle. Note that the whole side of the triangle opposite to the red dot has $R = 0$, and similarly for the sides opposite to the green and blue dots.

which makes it unity along the sides of the color triangle (i.e., where $R = 0$ or $G = 0$ or $B = 0$), and zero for white light ($R = G = B = I$). Note how the equation for S favors none of the R, G, B components. We emphasize that S does not express color itself but is a measure of the *proportion* of color relative to pure white. Fig. C.1 depicts the color triangle: its three corners are the points where the (R, G, B) axes (in RGB color space) cross a given constant I color plane.

Hue is defined as an angle H of rotation about the central white point **W** in the color triangle: it is the angle between the pure red direction (defined by the vector **R**−**W**) and the direction of the color **C** in question (defined by the vector **C**−**W**). The derivation of a formula for H is fairly complex and will not be attempted here. Suffice it to say that it may be determined by calculating cos H, which depends on the dot product $(\mathbf{C}-\mathbf{W}) \times (\mathbf{R}-\mathbf{W})$. The final result is

$$H = \cos^{-1}\left(\frac{\frac{1}{2}[(R-G)+(R-B)]}{\left[(R-G)^2+(R-B)(G-B)\right]^{1/2}}\right) \tag{C.3}$$

or 2π minus this value if $B > G$ (Gonzalez and Woods, 1992).

The above discussion shows that hue is a 1-D parameter. In fact, it is a *periodic* parameter as hue is an angle with a period of 2π. While this might appear to be only a marginal factor, it has the effect of complicating a number of image analysis computations. First, if the hue variable is to be smoothed using a 1-D moving average filter, care will have to be taken at the ends of the range 0 to 2π. Second, if the smoothing is to be done using a 1-D median filter, the operation cannot properly be defined because the hue values are double-valued at 0°. (This means that there is no unique ordering of the hue values by which to define a median value.)

Finally, the ranges of values of R, G, B, and I are given by the particular digitization used: we here assume that there is one byte of data for each of the RGB color parameters—in which case we have $0 \leq R, G, B, I \leq 255$.

C.3 A TYPICAL EXAMPLE OF THE USE OF COLOR

There is one area where color has a large part to play: that is in the automatic picking, inspection, and sorting of fruit. In particular, color is very important in the determination of apple quality. Not only is it a prime indicator of ripeness, but also it contributes greatly to physical attractiveness—not least with regard to readiness for eating. Clearly, for the computer to emulate human performance in assessing the appearance of fruit, it is useful to convert the RGB representation to the HSI domain before making any judgments about color.

In the work of Heinemann et al. (1995), discriminant analysis of color based on this approach gave complete agreement between human inspectors and the computer following training on 80 samples and testing on another 66 samples. However, a warning was given about maintaining lighting intensity levels identical to those used for training: in any such pattern recognition system, it is crucial that the training set be fully representative of the eventual test set.

When checking the color of apples, the hue is the important parameter. A rigorous check on the color can be achieved by constructing the hue distribution and comparing it with that for a suitable training set. The most straightforward way to carry out the comparison is to compute the mean and standard deviation of the two distributions to be compared and to perform discriminant analysis assuming Gaussian distribution functions. The standard theory for maximum likelihood thresholding (Section 4.5.3, Eqs. (4.19)−(4.22)) then leads to an optimum hue decision threshold.

Finally, note that full color discrimination would require an optimal decision surface to be ascertained in the overall 3-D color space. In general, such decision surfaces are hyperellipses and have to be determined using the Mahalanobis distance measure (see, e.g., Webb, 2002). However, in the special case of Gaussian distributions with equal covariance matrices, or more simply with equal isotropic variances, the decision surfaces become hyperplanes.

C.4 BIBLIOGRAPHICAL AND HISTORICAL NOTES

Space prevents a detailed study of the question of color: the reader is referred to more specialized texts for detailed information (e.g., Gonzalez and Woods, 1992; Sangwine and Horne, 1998). Examples of work on color inspection include food (Heinemann et al., 1995) and pharmaceutical products (Derganc et al., 2003). For early work on color constancy, see Forsyth (1990) and Finlayson et al. (2001).