

Color Based Object Recognition

T. Gevers and A.W.M. Smeulders

Faculty of WINS, University of Amsterdam, The Netherlands

email: gevers@wins.uva.nl

Images and recognition scheme can be experienced on-line at

<http://www.wins.uva.nl/research/isis/zomax/>

abstract

Assuming white illumination and dichromatic reflectance, we propose new color models $c_1c_2c_3$ and $l_1l_2l_3$ invariant to the viewing direction, object geometry and shading. Further, it is shown that $l_1l_2l_3$ is also invariant to highlights. Further, a change in spectral power distribution of the illumination is considered to propose a new photometric color invariant $m_1m_2m_3$ for matte objects.

To evaluate photometric color invariant object recognition in practice, experiments have been carried out on a database consisting of 500 images taken from 3-D multicolored man-made objects.

On the basis of the reported theory and experimental results, it is shown that high object recognition accuracy is achieved by $l_1l_2l_3$ and hue H followed by $c_1c_2c_3$ and normalized colors rgb under the constraint of white illumination. Finally, it is shown that solely $m_1m_2m_3$ is invariant to a change in illumination color.

1 Introduction

Color provides powerful information for object recognition. A simple and effective recognition scheme is to represent and match images on the basis of RGB histograms as proposed by Swain and Ballard [6]. This color-based recognition method has been extended by Funt and Finlayson [2] to become illumination independent by indexing on an illumination-invariant set of color descriptors. Furthermore, Healey and Slater [4] use illumination invariant moment descriptors for object recognition. The method fails, however, when objects are occluded as the moments are defined as an integral property on the (region) object as one.

Our aim is to analyze and evaluate various color models to be used for the purpose of object recognition by color-metric histogram matching according to the following criteria: 1. Robustness to a change in viewpoint; 2. Robustness to a change in object orientation; 3. Robustness to a change in the intensity and the direction of the illumination; 4. Robustness to a change in the color of the illumination; 5. High discriminative power; 6. Robustness to noise; 7. Robustness to object occlusion and cluttering.

The general application is considered of recognition of 3-D multicolored objects from 2-D color images.

This paper is organized as follows. In Section 2, the dichromatic reflectance under "white" reflection is introduced and new photometric invariant color features are proposed. The performance of object recognition by histogram matching differentiated for the various color models is evaluated and compared on an image database of 500 reference images in Section 3.

2 Photometric Color Invariance

In this paper, we concentrate on the following standard, *essentially different*, color features derived from RGB : intensity $I(R, G, B) = R + B + G$, RGB , normalized colors $r(R, G, B) = \frac{R}{R+G+B}$, $g(R, G, B) = \frac{G}{R+G+B}$, $b(R, G, B) = \frac{B}{R+G+B}$, hue $H(R, G, B) = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right)$ and saturation $S(R, G, B) = 1 - \frac{\min(R, G, B)}{R+G+B}$.

2.1 The Reflection Model

Consider an image of an infinitesimal surface patch. Using the red, green and blue sensors with spectral sensitivities given by $f_R(\lambda)$, $f_G(\lambda)$ and $f_B(\lambda)$ respectively, to obtain an image of the surface patch illuminated by a SPD of the incident light denoted by $e(\lambda)$, the measured sensor values will be given by Shafer [5]:

$$C = m_b(\mathbf{n}, \mathbf{s}) \int_{\lambda} f_C(\lambda) e(\lambda) c_b(\lambda) d\lambda + m_s(\mathbf{n}, \mathbf{s}, \mathbf{v}) \int_{\lambda} f_C(\lambda) e(\lambda) c_s(\lambda) d\lambda \quad (1)$$

for $C = \{R, G, B\}$ giving the C th sensor response. Further, $c_b(\lambda)$ and $c_s(\lambda)$ are the albedo and Fresnel reflectance respectively. λ denotes the wavelength, \mathbf{n} is the surface patch normal, \mathbf{s} is the direction of the illumination source, and \mathbf{v} is the direction of the viewer. Geometric terms m_b and m_s denote the geometric dependencies on the body and surface reflection respectively.

Considering the neutral interface reflection (NIR) model (assuming that $c_s(\lambda)$ has a constant value independent of the wavelength) and "white" illumination, then $e(\lambda) = e$ and $c_s(\lambda) = c_s$. Then, we propose that the measured sensor values are given by:

$$C_w = em_b(\mathbf{n}, \mathbf{s})k_C + em_s(\mathbf{n}, \mathbf{s}, \mathbf{v})c_s \int_{\lambda} f_C(\lambda) d\lambda \quad (2)$$

for $C_w \in \{R_w, G_w, B_w\}$ giving the red, green and blue sensor response under the assumption of a white light source. $k_C = \int_{\lambda} f_C(\lambda) c_b(\lambda) d\lambda$ is a compact formulation depending on the sensors and the surface albedo.

If the integrated white condition holds (as we assume throughout the paper) $\int_{\lambda} f_R(\lambda) d\lambda = \int_{\lambda} f_G(\lambda) d\lambda = \int_{\lambda} f_B(\lambda) d\lambda = f$, we have:

$$C_w = C_b + C_s = em_b(\mathbf{n}, \mathbf{s})k_C + em_s(\mathbf{n}, \mathbf{s}, \mathbf{v})c_s f \quad (3)$$

2.2 Reflection with White Illumination

Photometric Invariant Color Features for Matte, Dull Surfaces According to the body reflection term of eq. (3), $C_b = em_b(\mathbf{n}, \mathbf{s})k_C$, a uniformly painted surface (i.e. with fixed k_C) may give rise to a broad variance of RGB values due to the varying circumstances induced by the image-forming process. The same argument holds for intensity I .

In contrast, normalized color rgb is insensitive to surface orientation, illumination direction and intensity as can be seen from:

$$r(R_b, G_b, B_b) = \frac{em_b(\mathbf{n}, \mathbf{s})k_R}{em_b(\mathbf{n}, \mathbf{s})(k_R + k_G + k_B)} = \frac{k_R}{k_R + k_G + k_B} \quad (4)$$

only dependent on the sensors and the surface albedo. Equal arguments hold for g and b .

Saturation S is an invariant for the set of matte, dull surfaces illuminated by white SPD mathematically specified by:

$$S(R_b, G_b, B_b) = 1 - \frac{\min(em_b(\mathbf{n}, \mathbf{s})k_R, em_b(\mathbf{n}, \mathbf{s})k_G, em_b(\mathbf{n}, \mathbf{s})k_B)}{em_b(\mathbf{n}, \mathbf{s})(k_R + k_G + k_B)} = 1 - \frac{\min(k_R, k_G, k_B)}{(k_R + k_G + k_B)} \quad (5)$$

Similarly, hue H is an invariant for matte, dull surfaces:

$$H(R_b, G_b, B_b) = \arctan\left(\frac{\sqrt{3}em_b(\mathbf{n}, \mathbf{s})(k_G - k_B)}{em_b(\mathbf{n}, \mathbf{s})(k_R - k_G) + (k_R - k_B)}\right) = \arctan\left(\frac{\sqrt{3}(k_G - k_B)}{(k_R - k_G) + (k_R - k_B)}\right) \quad (6)$$

In fact, any expression defining colors on the same linear color cluster formed by the body reflection vector in RGB -space are photometric color invariants for dichromatic reflectance for matte surfaces under white light. To that end, we propose the following photometric color invariant model:

$$c_1 = \arctan\left(\frac{R}{\max\{G, B\}}\right), c_2 = \arctan\left(\frac{G}{\max\{R, B\}}\right), c_3 = \arctan\left(\frac{B}{\max\{R, G\}}\right) \quad (7)$$

denoting the angles of the body reflection vector and consequently being invariants for matte, dull objects:

$$c_1(R_b, G_b, B_b) = \arctan\left(\frac{em_b(\mathbf{n}, \mathbf{s})k_R}{\max\{em_b(\mathbf{n}, \mathbf{s})k_G, em_b(\mathbf{n}, \mathbf{s})k_B\}}\right) = \arctan\left(\frac{k_R}{\max\{k_G, k_B\}}\right) \quad (8)$$

only dependent on the sensors and the surface albedo. Equal arguments hold for c_2 and c_3 .

Photometric Invariant Color Features for Both Matte and Shiny Surfaces Note that under the given conditions (the NIR model), the color of highlights is not related to the color of the surface on which they appear, but only on the color of the light source. Thus for the white light source, the surface reflection color cluster is on the diagonal grey axis of the basic RGB -color space corresponding to intensity I . For a given point on a shiny surface, the contribution

of the body reflection component and surface reflection component are added together. Hence, the observed colors of the surface must be inside the triangular color cluster in the RGB -space formed by the two reflection components.

Because H is a function of the angle between the reference line and color point, all possible colors of the same shiny uniformly colored surface have to be of the same hue mathematically specified as:

$$H(R_w, G_w, B_w) = \arctan\left(\frac{\sqrt{3}(k_G - k_B)}{(k_R - k_G) + (k_R - k_B)}\right) \quad (9)$$

Only dependent on the sensors and the surface albedo. Obviously other color features depend on the contribution of the surface reflection component and hence are sensitive to highlights.

In fact, any expression defining colors on the same linear triangular color cluster, formed by the two reflection components in RGB -space, are photometric color invariants for the dichromatic reflectance under white light.

To that end, a new photometric color invariant model $l_1 l_2 l_3$ is proposed uniquely determining the direction of the linear triangular color cluster: $l_1 = \frac{(R-G)^2}{(R-G)^2+(R-B)^2+(G-B)^2}$, $l_2 = \frac{(R-B)^2}{(R-G)^2+(R-B)^2+(G-B)^2}$, $l_3 = \frac{(G-B)^2}{(R-G)^2+(R-B)^2+(G-B)^2}$ the set of normalized color differences which is, similar to H , an invariant for the set of matte and shiny surfaces.

2.3 Reflection with Colored Illumination

The Reflection Model We consider the body reflection term of the dichromatic reflection model:

$$C_c = m_b(\mathbf{n}, \mathbf{s}) \int_{\lambda} f_C(\lambda) e(\lambda) c_b(\lambda) d\lambda \quad (10)$$

for $C = \{R, G, B\}$ where $C_c = \{R_c, G_c, B_c\}$ gives the red, green and blue sensor response of a matte infinitesimal surface patch under unknown spectral power distribution of the illumination.

Suppose that the sensor sensitivities of the color camera are narrowband with spectral response be approximated by delta functions $f_C(\lambda) = \delta(\lambda - \lambda_C)$, then we have:

$$C_c = m_b(\mathbf{n}, \mathbf{s}) e(\lambda_C) c_b(\lambda_C) \quad (11)$$

By simply filling in C_c in the color feature equations, we can see that all color feature values change with a change in illumination color.

Color Constant Feature for Matte, Dull Surfaces Funt and Finlayson [2] proposes a simple and effective illumination-independent color feature for the purpose of object recognition. The method runs short, however, when images are contaminated by shading and highlights. To that end, we propose a color constant feature not only independent of the illumination color but also discounting shading cues:

$$m(C_1^{\mathbf{x}_1}, C_1^{\mathbf{x}_2}, C_2^{\mathbf{x}_1}, C_2^{\mathbf{x}_2}) = \frac{C_1^{\mathbf{x}_1} C_2^{\mathbf{x}_2}}{C_1^{\mathbf{x}_2} C_2^{\mathbf{x}_1}}, C_1 \neq C_2 \quad (12)$$

expressing the color ratio between two neighboring image locations, for $C_1, C_2 \in \{R, G, B\}$ where \mathbf{x}_1 and \mathbf{x}_2 denote the image locations of the two neighboring pixels.

Having three color components of two locations, color ratios obtained from a *RGB*-color image are:

$$m_1 = \frac{R^{\mathbf{x}_1} G^{\mathbf{x}_2}}{R^{\mathbf{x}_2} G^{\mathbf{x}_1}}, m_2 = \frac{R^{\mathbf{x}_1} B^{\mathbf{x}_2}}{R^{\mathbf{x}_2} B^{\mathbf{x}_1}}, m_3 = \frac{G^{\mathbf{x}_1} B^{\mathbf{x}_2}}{G^{\mathbf{x}_2} B^{\mathbf{x}_1}} \quad (13)$$

For the ease of exposition, we concentrate on m_1 based on the *RG*-color bands in the following discussion. Without loss of generality, all results derived for m_1 will also hold for m_2 and m_3 .

If we assume that the color of the illumination is locally constant (at least over the two neighboring locations from which ratio is computed), the color ratio is independent of the illumination color, and also a change in viewpoint, the surface geometry, and illumination intensity as follows from:

$$m_1 = \frac{(m_b^{\mathbf{y}_1}(\mathbf{n}, \mathbf{s})e^{\mathbf{y}_1}(\lambda_R)c_b^{\mathbf{y}_1}(\lambda_R))(m_b^{\mathbf{y}_2}(\mathbf{n}, \mathbf{s})e^{\mathbf{y}_2}(\lambda_G)c_b^{\mathbf{y}_2}(\lambda_G))}{(m_b^{\mathbf{y}_2}(\mathbf{n}, \mathbf{s})e^{\mathbf{y}_2}(\lambda_R)c_b^{\mathbf{y}_2}(\lambda_R))(m_b^{\mathbf{y}_1}(\mathbf{n}, \mathbf{s})e^{\mathbf{y}_1}(\lambda_G)c_b^{\mathbf{y}_1}(\lambda_G))} = \frac{c_b^{\mathbf{y}_1}(\lambda_R)c_b^{\mathbf{y}_2}(\lambda_G)}{c_b^{\mathbf{y}_2}(\lambda_R)c_b^{\mathbf{y}_1}(\lambda_G)} \quad (14)$$

only dependent on the surface albedo, where \mathbf{y}_1 and \mathbf{y}_2 are two neighboring locations on the object's surface not necessarily of the same orientation.

Taking logarithms of both sides of equation 12 results for m_1 in:

$$\ln m_1(R^{\mathbf{x}_1}, R^{\mathbf{x}_2}, G^{\mathbf{x}_1}, G^{\mathbf{x}_2}) = \ln R^{\mathbf{x}_1} + \ln G^{\mathbf{x}_2} - \ln R^{\mathbf{x}_2} - \ln G^{\mathbf{x}_1} \quad (15)$$

The color ratios can be seen as differences at two neighboring locations \mathbf{x}_1 and \mathbf{x}_2 in the image domain:

$$d_{m_1}(\mathbf{x}_1, \mathbf{x}_2) = \ln R^{\mathbf{x}_1} + \ln G^{\mathbf{x}_2} - \ln R^{\mathbf{x}_2} - \ln G^{\mathbf{x}_1} \quad (16)$$

When these differences are taken between neighboring pixels in a particular direction, they correspond to finite-difference differentiation.

The results obtained so far for m_1 hold also for m_2 and m_3 , yielding a 3-tuple $(\mathcal{G}_{m_1}(\mathbf{x}), \mathcal{G}_{m_2}(\mathbf{x}), \mathcal{G}_{m_3}(\mathbf{x}))$ denoting the gradient magnitude for every neighborhood centered at \mathbf{x} in the image.

For pixels on a uniformly painted region, in theory, all three components will be zero whereas at least one the three components will be non-zero for pixels on locations where two regions of distinct color meet.

3 Color Based Object Recognition: Experiments

The database consists of $N_1 = 500$ reference images of multicolored domestic objects, tools, toys, etc.. Objects were recorded in isolation (one per image) with the aid of the SONY XC-003P CCD color camera (3 chips) and the Matrox Magic Color frame grabber. Objects were recorded against a white cardboard background. Two light sources of average day-light color are used to illuminate the objects in the scene. A second, independent set (the test set) of recordings was made of randomly chosen objects already in the database. These objects, $N_2 = 70$ in number, were recorded again (one per image) with a new, arbitrary position and orientation with respect to the camera (some recorded upside down, some rotated, some at different distances (different scale)).

Histograms are constructed on the basis of different color features representing the distribution of discrete color feature values in a n -dimensional color feature space, where $n = 3$ for RGB , rgb , $l_1l_2l_3$, $c_1c_2c_3$ and $m_1m_2m_3$, and $n = 1$ for I , S and H . During histogram construction, all pixels in a color image are discarded with a local saturation and intensity smaller than 5 percent of the total range. Consequently, the white cardboard background as well as the grey, white, dark or nearly colorless parts of objects as recorded in the color image will not be considered in the matching process. For comparison reasons in the literature, in this paper, the histogram similarity function is expressed by histogram intersection [6].

For a measure of match quality, let rank r^{Q_i} denote the position of the correct match for test image Q_i , $i = 1, \dots, N_2$, in the ordered list of N_1 match values. The rank r^{Q_i} ranges from $r = 1$ from a perfect match to $r = N_1$ for the worst possible match. Then, for one experiment, the average ranking percentile is defined by $\bar{r} = (\frac{1}{N_2} \sum_{i=1}^{N_2} \frac{N_1 - r^{Q_i}}{N_1 - 1}) 100\%$. The cumulative percentile of test images producing a rank smaller or equal to j is defined as $\mathcal{X}(j) = (\frac{1}{N_2} \sum_{k=1}^j \eta(r^{Q_i} == k)) 100\%$, where η reads as the number of test images having rank k .

For more information see [3]. The image database and the performance of the recognition scheme can be experienced within the ZOMAX system on-line at <http://www.wins.uva.nl/research/isis/zomax/>.

4 Results

In this subsection, we report on the recognition accuracy of the matching process for $N_2 = 70$ test images and $N_1 = 500$ reference images for the various color features. As stated, white lighting is used during the recording of the reference images in the image database and the independent test set. However, the objects were recorded with a new, arbitrary position and orientation with respect to camera. In Fig. 1 accumulated ranking percentile is shown for the various color features.

From the results of Fig. 1 we can observe that the discriminative power of $l_1l_2l_3$, H followed by $c_1c_2c_3$ and rgb is higher than the other color features achiev-

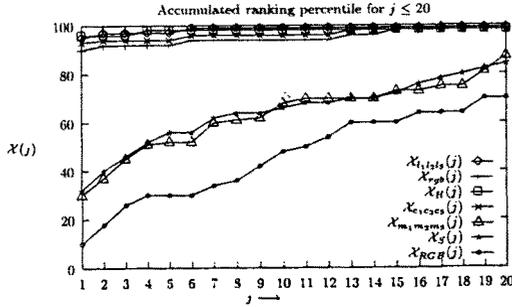


Fig. 1. The discriminative power of the histogram matching process differentiated for the various color features plotted against the ranking j . The cumulative percentile \mathcal{X} for H , $l_1 l_2 l_3$, $c_1 c_2 c_3$, rgb , S , $m_1 m_2 m_3$ and RGB is given by \mathcal{X}_H , $\mathcal{X}_{l_1 l_2 l_3}$, $\mathcal{X}_{c_1 c_2 c_3}$, \mathcal{X}_{rgb} , \mathcal{X}_S , $\mathcal{X}_{m_1 m_2 m_3}$ and \mathcal{X}_{RGB} respectively.

ing a probability of respectively 99, 98, 94 and 92 perfect matches out of 100. Saturation S and color ratio $m_1 m_2 m_3$ provides slightly worse recognition accuracy. As expected, the discrimination power of RGB has the worst performance due to its sensitivity to varying imaging conditions.

4.1 The Effect of a Change in the Illumination Intensity

The effect of a change in the illumination intensity is approximated by a multiplication of each RGB -color by a uniform scalar factor α . To measure the sensitivity of different color feature in practice, RGB -images of the test set are multiplied by a constant factor varying over $\alpha \in \{0.5, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.5\}$. The discrimination power of the histogram matching process differentiated for the various color features plotted against illumination intensity is shown in Fig. 2. As expected, RGB and I -color features depend on the illumination intensity.

4.2 The Effect of a Change in the Illumination Color

Based on the coefficient rule or von Kries model, the change in the illumination color is approximated by a 3×3 diagonal matrix among the sensor bands and is equal to the multiplication of each RGB -color band by an independent scalar factor [1]. Note that the diagonal model of illumination change holds in the case of narrowband sensors. To measure the sensitivity of the various color feature in practice with respect to a change in the color of the illumination, the R , G and B -color bands of each image of the test set are multiplied by a factor $\beta_1 = \beta$, $\beta_2 = 1$ and $\beta_3 = 2 - \beta$ respectively (i.e. $\beta_1 R$, $\beta_2 G$ and $\beta_3 B$) by varying β over $\{0.5, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.5\}$. The discrimination power of the histogram matching process differentiated for the various color features plotted against the illumination color is shown in Fig. 3. For $\beta < 1$ the color is reddish whereas bluish for $\beta > 1$.

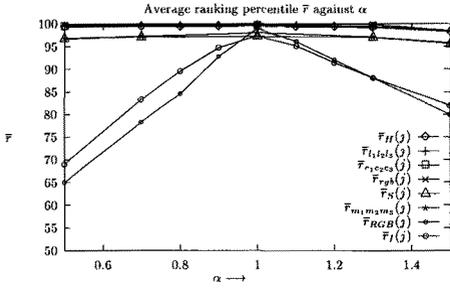


Fig. 2. The discriminative power plotted against the illumination intensity represented by variation as expressed by the factor α .

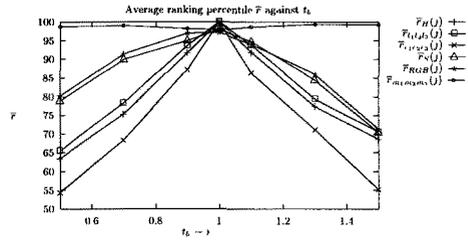


Fig. 3. The discriminative power plotted against the change β in the color composition of the illumination spectrum.

As expected, only the color ratio $m_1 m_2 m_3$ is insensitive to a change in illumination color. From Fig. 3 we can observe that color features H , $l_1 l_2 l_3$ and $c_1 c_2 c_3$, which achieved best recognition accuracy under white illumination, see Figures 1 and 2, are highly sensitive to a change in illumination color.

5 Conclusion

On the basis of the above reported theory and experiments, it is concluded that the proposed invariant $l_1 l_2 l_3$ followed by H are most appropriate to be used for photometric color invariant object recognition by color-metric histogram matching under the constraint of a white illumination source. When no constraints are imposed on the imaging conditions (i.e. the most general case), the newly proposed color ratio $m_1 m_2 m_3$ is most appropriate.

References

1. Finlayson, G. D., Drew M. S., and Funt B. V., *Spectral Sharpening: Sensor Transformations for improved Color Constancy* J. Opt. Soc. Am., 11(5), pp. 1553-1563, 1994.
2. Funt, B. V. and Finlayson, G. D., *Color Constant Color Indexing*, IEEE PAMI, 17(5), pp. 522-529, 1995.
3. Gevers, T., *Color Image Invariant Segmentation and Retrieval*, PhD Thesis, ISBN 90-74795-51-X, University of Amsterdam, The Netherlands, 1996.
4. Healey, G. and Slater D., *Global Color Constancy: Recognition of Objects by Use of Illumination Invariant Properties of Color Distributions*, J. Opt. Soc. Am. A, Vol. 11, No. 11, pp. 3003-3010, Nov 1995.
5. Shafer, S. A., *Using Color to Separate Reflection Components*, COLOR Res. Appl., 10(4), pp 210-218, 1985.
6. Swain, M. J. and Ballard, D. H., *Color Indexing*, Int. Journal of CV, Vol. 7, No. 1, pp. 11-32, 1991.