Color Angle Invariants for Object Recognition

G. D. Finlayson, S. S. Chatterjee and B. V. Funt
School of Computing Science
Simon Fraser University, Vancouver, B. C. Canada

Abstract

We present a fast, color-based algorithm for recognizing objects viewed under an unknown illuminant. Objects are indexed by just three numbers: the angles of the object's color distribution. If R, G and B denote the 3 color bands of the image of an object (stretched out as vectors) then the angular index comprises the 3 inter-band angles (one per pair of vectors). In the general case the distribution of colors, and in turn the angular index, will depend on the color of the illuminant. If, however, the original color bands are transformed by a *sharpening* transform[2,3] before computing the distribution angles, then we show that the angular index is illuminant independent. Indexing using angles calculated post sharpening delivers excellent recognition for a variety of illuminations.

Introduction

Swain and Ballard[10] developed an algorithm, called color-indexing, which sets out to recognize objects using only color information; objects are represented by color histograms (i.e. the distribution of colors present in an object) and recognition proceeds by histogram matching. Good recognition is possible since color histograms are very robust features of objects since they are invariant to translation and rotation about the optical axis and change only slowly as a function of rotation about other axes. Moreover color histograms are relatively stable to object occlusion and changes of object scale. However, the color histogram method breaks down completely if illumination is allowed to vary since the color histogram depends on the color of the light. Funt and Finlayson[4] have shown that by histogramming color ratios taken between nearby image locations instead of the raw colors, the histogram technique can be made invariant to changes in the illumination. This Color Constant Color Indexing (CCCI) algorithm performs much better than straight color indexing under an illuminant change and almost equally well for a fixed illuminant. However, it is computationally more expensive and it is limited by the fact that ratios at low intensity levels can become dominated by noise. A problem common to both color-indexing and CCCI is that histograms are compared a bin at a time which can be an expensive operation (typically a histogram has thousands of bins and comparison requires thousands of operations).

More recently, Healey and Slater[5] have used moment invariants computed from color histograms for recognition in the presence of illumination changes. Their algorithm is based on the assumption that when the illumination changes the colors in an image shift by a linear transform. When this assumption holds (in general it does[8]) a linear transform must relate the color histograms of an object viewed under two different illuminations. Taubin and Cooper[11] have developed efficient algorithms for the computation of invariants of centralized moments; applied to color histograms, these are invariant under an affine transformation. It follows then that these invariants provide an illumination independent index useful for color object recognition. Moreover, centralized moments effectively circumvent the ratio accuracy limitation of CCCI. Healey and Slater demonstrate that a small set of 6 moment invariants supports good recognition when the object database is small. Unfortunately centralized moments capture only a small amount of the total information available in the histogram; specifically all the low-frequency information is lost. We predict therefore that moment indexing will return poor performance for larger object sets. Results corroborating this prediction are presented later.

The motivation for using moments is that colors shift by a linear transform under a change in illumination. While a linear model accurately describes illumination change it is in fact too general a model since most linear transforms never occur in practice. In this paper, we model illumination change by a von Kries type model: colors under different illuminants are related by a diagonal matrix. What this means is that under an illumination change every pixel within an image band is scaled by a single *von Kries* coefficient. Based on this simpler model of illumination change we propose a new index for object recognition: the angles of the color distributions in an image.

An image band containing N pixels can be thought of as an N-vector. Thus, it is immediate that under the von Kries model the length of each N-vector (or image band) changes with illumination but its orientation remains fixed. From this it follows that the angles between the image-bands are independent of the illuminant.These illumination-invariant angles comprise the angular index which we will use as our object descriptor. Because the descriptor consists of just 3 numbers it promises very efficient indexing. Moreover most of the low-frequency color information is preserved; from which it follows that object recognition on reasonably large data sets should be possible.

We criticized the linear model by saying that it was too powerful. Equally the von Kries model is widely believed to be too simple; this is especially true for computational models operating with human cone sensors[12]. However recent work by Finlayson et al. has shown that while a von Kries model may not be appropriate for the original color image it will always be appropriate for a *sharpened* color image which is created by taking a linear combination of the original color bands.

Angle Invariants

The light reflected from a surface depends on the spectral properties of the surface reflectance and illumination incident on the surface. We will restrict our discussion here to Lambertian surfaces. In an imaging system, light reflected from a surface falls onto a planar array of sensors in the camera. Each location x on the array has k classes of sensors. The value (ρ^x_k) at each sensor output is given by the integral of its response function multiplied by the light and reflectance:

$$\rho^x_k = \int_\omega S^{x'}(\lambda) E^{x'}(\lambda) R_k(\lambda) d\lambda \tag{1}$$

where λ is the wavelength, R_k is the response function of the k^{th} sensor class, $E^{x'}(\lambda)$ is the incident illumination and $S^{x'}(\lambda)$ is the surface reflectance function at location x' on the surface which is projected onto location x on the sensor. We further assume here that the illumination does not vary spectrally over the given surface, and so drop the index x' from $E(\lambda)$. Under a von Kries model of illumination change the sensor responses under two different illuminants are assumed to be related by a diagonal matrix.

$$\underline{\rho}^x \approx D\rho^x \tag{2}$$

here $\underline{\rho}^x$ represents the k sensor outputs at location x on the sensor array under a different illumination $E(\lambda)$. The diagonal matrix D contains the von Kries coefficients taking sensor responses between illuminants. Note that the same diagonal matrix D maps the entire image.

In practice the von Kries model will in general only approximately account for a changing illuminant and the relationship in Eq. (2) can be quite imprecise[12]. The exception to this is if the spectral sensitivities of the vision system are very narrow-band. In this circumstance Eq. (2) is exact. In fact the von Kries model is quite accurate so long as there exists linear combinations of sensitivities which are narrow-band. This observation forms the basis of Finlayson et al's[3] spectral sharpening method. They show that if the sensor response functions are first transformed to a more narrow-band (or sharper) sensor basis then the accuracy of Eq. (2) is improved. Indeed a von Kries type model is quite adequate for all sensor sets[2]. Von Kries plus sharpening can be written as

$$T\underline{\rho}^x \approx TD\rho^x \tag{3}$$

where T denotes the sharpening transformation of the original sensor response functions. Let W be a $3\times N$ matrix representing the set of sharpened sensor responses for a 3 sensor imaging environment with N array elements. We can then write

$$\underline{W} \approx DW \tag{4}$$

where $\underline{W}$ is the set of responses under a different illumination. We see from Eq. (4) that a change in illumination corresponds to a change in length for each of the rows of W. Upon normalizing the rows of W and $\underline{W}$ to unit length, the diagonal matrix D in Eq. (4) reduces to the identity matrix. It follows from the above that the angles between the three N-dimensional row vectors are invariant to changes in the illumination. Using subscripts and superscripts to index row and columns respectively then these angles are computed as

$$\phi_{(i,j)} = \cos^{-1}\left(\sum_{x=1}^{N} W_i^x W_j^x\right) \tag{5}$$

where $W_k^x = \frac{W_k^x}{|W_k|}$ (each row is normalized to unit length). The $\phi_{(i,j)}$ thus computed form our color distribution descriptors. These can be computed with $3N$ multiplications and $3N$ additions from the sharpened responses. This compares favorably with Healey and Slater's moment index which requires $21N$ multiplications to construct.

Experimental Results

To test distribution angle indexing we begin by generating a set of synthetic Mondrian images using human cones[13] and sharpened-cone responses[3]. A set of 21 Mondrians under five different illuminants[6,13] (D48, D55, D75, D100 and CIE standard A) were generated. Each Mondrian comprised between 4 and 10 reflectances randomly drawn from the 24 Macbeth color checker chips[9], 12 ceramic reflectances[1] and 4 Krinov[7] natural reflectances. Images rendered under illuminant D55 were used to create the model database of angle indices, and the remaining 84 images were used to evaluate recognition performance. Matching is performed as follows: first, the angle index is calculated for each test image using Eq. (5) and second, each index is compared to those in the model database. If the minimum Euclidean distance between the image angle index and the model database occurs for the correct answer (i.e. the same Mondrian under D55) then the ranking is 1; if the correct answer is the second smallest match then the ranking is 2 and so on. The match rankings calculated for entire test set are shown in Table 1.

With the sharpened responses all the images are perfectly matched to the database; whereas, with the cone responses—where the diagonal assumption breaks down—only 74 images are perfectly matched, with 8 ranked under second place and the remaining 2 ranked as third matches in the database.

As a second test we constructed a database of 55 angular indices one per object for Swain and Ballard's object database (each index is calculated from a color image of each object). A test set of angular indices are generated for 24 of these objects viewed in different positions and with small degrees of occlusion and deformation; however, the illumination color was held constant. These test indices are then compared to the model indices and the rank of the correct match calculated. Table 2 summarizes the match rankings for angular indexing. Color Constant Color Indexing[4] and Healey and Slater's[5] algorithm have also been implemented and run on the same data; ranking results from running these algorithms are also shown in Table 2. Our algorithm does reasonably well in matching most of the objects within the top 3 places—only 1 object returns a greater than third place match. This match success is quite remarkable given that we are indexing a large database (55 objects!) with just 3 numbers. Moreover comparison is comparable with CCCI (which matches based on the thousands of histogram bin counts), although CCCI manages to rank all but 2 in the first place. Healey's algorithm does much worse with 7 images ranked greater than third place.

Table 1. Performance of Human Cone responses vs sharpened responses on synthetic data

Cone Responses	Rankings			
	1	2	3	> 3
Sharpened	84	0	0	0
Human Cones	74	8	2	0

Table 2. Database of 55 real objects

Algorithm	Rankings			
	1	2	3	>3
Angular	16	5	2	1
CCCI	22	2	0	0
Healey	7	7	3	7

Table 3. Database of 13 real objects

Algorithm	Rankings			
	1	2	3	>3
Angular	20	3	3	0
CCCI	24	2	0	0
Healey	11	6	6	3

In Table 3, we compare the performance of the 3 algorithms for a second set of 13 objects where illumination color is varied. The database in this case comprises features derived for images taken under a whitish artificial light. The 26 test features are derived from the same objects, viewed in different positions and under either bluish or reddish illumination. The angle-invariant method matches 20 out of the 26 test images perfectly, though all the objects are identified within the top 3 rankings. CCCI does a little better correctly identifying 24 of the 26 images. Healey's algorithm performs significantly worse matching just 11 images correctly, although 3 images have a match ranking greater than 3.

Conclusion

A new method of color based object recognition has been presented based on the angle between the 3 bands of a color image where each band is viewed as a vector. The method's primary virtue lies in the speed with which the angle invariants can be computed and the simplicity with which they can be compared to angle invariants stored in the database of known objects. It is significantly faster than CCCI and comparable to Healey's algorithm in the amount of time taken to compute and match the angle invariants. As shown by experiments using both real and synthetic data, the three angle invariants provide very good matching results, nearly always identifying the correct object within the top three.

References

1. R. S. Berns and K. H. Petersen. Empirical modelling of systematic spectrophotometric errors. *Color Res. Appl.*, vol **4**, pp 243, 1988.
2. G. D. Finlayson, M. S. Drew and B. V. Funt. "Color constancy: generalized diagonal transforms suffice". *Journal of the Optical Society of America, A*, vol **11**, no. 11, pp. 3011-3019, Nov. 1994.
3. G. D. Finlayson, M. S. Drew and B. V. Funt. "Spectral Sharpening: sensor transformations for improved color constancy". *Journal of the Optical Society of America, A*, vol. **11**, no. 5, pp. 1553-1563, May 1994.
4. B. V. Funt and G. D. Finlayson. "Color constant color indexing". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol **17**, no. 5, pp. 522-528, May 1995.
5. G. Healey and D. Slater. "Global color constancy: recognition of objects by use of illumination invariant properties of color distributions". *Journal of the Optical Society of America, A*, vol. **11** no. 11, pp. 3003-3010, Nov. 1994.
6. D. B. Judd, D. L. MacAdam, and G. Wyszecki. "Spectral distribution of typical daylight as a function of correlated color temperature". *Journal of the Optical Society of America,* vol. **54**, pp. 1031-1040, Aug. 1964.
7. E. L. Krinov. "Spectral Reflectance properties of natural formations". *Technical Translation TT-439, National Research Council of Canada*, 1947.
8. L. T. Maloney. "Evaluation of linear models of surface spectral reflectance with small numbers of parameters". *Journal of the Optical Society of America, A,* vol. **3**, pp. 1673-1683, 1986.
9. C. S. McCamy, H. Marcus and J. G. Davidson. "A color-rendition chart". *J. App. Photog. Eng.*, pp. 95-99, 1976.

10. M. J. Swain and D. H. Ballard. "Color Indexing", *International Journal of Computer Vision*, vol. **7**, no. 11, pp. 11-32, 1991.
11. G. Taubin and D. Cooper. "Object Recognition based on moment (or algebraic) invariants". In J. Mundy and A. Zisserman, eds., *Geometric Invariance in Computer Vision*, pp. 375-397. MIT Press, Cambridge, Mass., 1992.
12. G. West and H. Brill. "Necessary and sufficient conditions for von Kries chromatic adaptation to give color constancy". *Journal of Math. Biol*, vol. **15**, pp. 249-258, 1982.
13. G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulas*. Wiley, New York, 2nd edition, 1982.