Color Angular Indexing

Graham D. Finlayson[1], Subho S. Chatterjee[2] and Brian V. Funt[2]

[1] Department of Computer Science, University of York, York YO1 5DD, UK
[2] School of Computing Science, Simon Fraser University, Vancouver, Canada

Abstract. A fast color-based algorithm for recognizing colorful objects and colored textures is presented. Objects and textures are represented by **just six numbers**. Let r, g and b denote the 3 color bands of the image of an object (stretched out as vectors) then the color angular index comprises the 3 inter-band angles (one per pair of image vectors). The color edge angular index is calculated from the image's color edge map (the Laplacian of the color bands) in a similar way. These angles capture important low-order statistical information about the color and edge distributions and invariant to the spectral power distribution of the scene illuminant. The 6 illumination-invariant angles provide the basis for angular indexing into a database of objects or textures and has been tested on both Swain's database of color objects which were all taken under the same illuminant and Healey and Wang's database of color textures which were taken under several different illuminants. Color angular indexing yields excellent recognition rates for both data sets.

1 Introduction

Various authors[13, 12, 10] (beginning with Swain and Ballard[13]) have found that the color distributions of multi-colored objects provide a good basis for object recognition. A color distribution, in the discrete case, is simply a three-dimensional histogram of the pixel values using one dimension for each color channel. For recognition, the color histogram of a test image is matched in some way (the matching strategy distinguishes the methods) against model histograms stored in the database. The closest model defines the identity of the image. Swain has shown that excellent recognition is possible so long as objects are always presented in the same pose and the illumination color is held fixed. Varying either the pose or illumination color can cause the color distribution to shift so as to prevent recognition.

Our goal is to develop a color-distribution-based descriptor that is concise, expressive and illuminant invariant. Swain and Ballard's histogram descriptors are expressive—they provide good recognition rates so long as the illuminant color is held fixed. However, they are not concise since each histogram is represented by the counts in 4096 bins. A descriptor requiring only a few parameters should speed up indexing performance and be useful in other more computationally intensive recognition algorithms (e.g. Murase and Nayar's [9] manifold method). Of course, we would prefer to decrease the match time without decreasing the match success.

Healey and Slater[4] have proposed a representation based on a small set of moments of color histograms[14]. They show that when illumination change is well described by a linear model (i.e., image colors change by a linear transform when the illumination changes) certain high-order distribution moments are illuminant invariant. Unfortunately, in the presence of noise (even in small amounts) this high-order information is not very stable and as such their representation may not be very expressive.

Requiring a full 3-by-3 linear model of illuminant change limits the kind of illuminant-invariant information that can be extracted from a color distribution. However, the linear model is in fact over general since only a small subset of the possible linear transforms correspond to physically plausible illuminant changes[2]. In particular if, as is usually the case, the sensitivities of our camera are relatively narrow-band then the images of the same scene viewed under two different illuminants are related (without error) by 3 simple scale factors. Each color pixel value (r_i, g_i, b_i) in the first color image becomes $(\alpha r_i, \beta g_i, \gamma b_i)$ in the second (where r_i, g_i and b_i denote the ith pixel in the red, green and blue image bands respectively and α, β and γ are scalars). This means that the images differing only in terms of the scene illuminant are always related by a simple 3-parameter diagonal matrix.

Under a diagonal model of illuminant change, the 3 angles between the different bands of a color image provide a simple illuminant independent invariant. To see this think of each image band as a vector in a high-dimensional space. When the illumination changes each vector becomes longer or shorter but its orientation remains unchanged. As well as being invariant to illumination change we show that the color angles encode important low-order statistical information.

Of course the camera sensors are not necessarily narrow-band and as such color angles might not be stable across a change in the illuminant. Nonetheless, good invariance can be attained if the angles are calculated with respect to a special *sharpened* color image. The sharpened color image, which exists for all sensor sets[2, 1], is created by taking linear combinations of the original color bands.

Angular indexing using just the 3 color angles suffice if the object database is small but performance breaks down for larger databases. To bolster recognition we develop a second angle invariant called the color-edge angle. Each band of the input color image is filtered by a Laplacian of Gaussian mask to generate a color edge map. The inter-band angles calculated with respect to this edge map are once again illuminant invariant and encode important information about the color edge distribution which when combined with the original 3 color angles leads to excellent recognition rates. Swain and Ballard demonstrated the richness of the color histogram representation for objects; and indexing on the 4096 histogram bins, they achieved almost flawless recognition for a database of 66 objects. Using color angular indexing, we attain very similar recognition rates indexing on **only 6 angles**.

As a second comparison, we evaluated recognition using Healey and Wang's color texture data set. It consists of 10 colored textures viewed under 4 illumin-

ants, but we added images of the same textures viewed under 5 more orientations giving a grand total of 240 images. Indexing with the 6 combined angles delivers almost perfect recognition for the texture dataset.

2 Background

2.1 Color Image Formation

The light reflected from a surface depends on the spectral properties of the surface reflectance and of the illumination incident on the surface. In the case of Lambertian surfaces, this light is simply the product of the spectral power distribution of the light source with the percent spectral reflectance of the surface. For the theoretical development, we henceforth assume that most surfaces are Lambertian. Illumination, surface reflection and sensor function, combine together in forming a sensor response:

$$\rho_k^x = \int_w S^{x'}(\lambda) E^{x'}(\lambda) R_k(\lambda) d\lambda \tag{1}$$

where λ is the wavelength, R_k is the response function of the kth sensor class (r, g or b), $E^{x'}$ is the incident illumination and $S^{x'}$ is the surface reflectance function at location x' on the surface which is projected onto location x on the sensor array. We assume the illumination does not vary spectrally over the scene, so the index x' from $E(\lambda)$ can be dropped.

In the sections which follow we denote an image by I. The content of the jth color pixel in I is denoted (r_j^I, g_j^I, b_j^I). we will assume that there are M pixels in an image.

2.2 Color Histograms

Let $H(I_1)$ and $H(I_2)$ be color histograms of the images I_1 and I_2. $H_{i,j,k}(I_1)$ is an integer recording the number of colors in I_1 which fall in the ijkth bin. The mapping of color to bin is usually not 1-to-1 but rather color space is split into discrete regions. For example, Swain and Ballard[13] split each color channel into 16 intervals giving $16 \times 16 \times 16 = 4096$ bins in each histogram.

To compare histograms $H(I_1)$ and $H(I_2)$ a similarity measure is computed. Swain calculates similarity of a pair of histograms as their common intersection:

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} \min(H_{i,j,k}(I_1), H_{i,j,k}(I_2)) \tag{2}$$

where N is the number of bins in each color dimension. While other methods of comparing histograms have been suggested, e.g. [10], they are all similar in the sense that they consist of many bin-wise operations. There are several problems with similarity calculated in this way. First, each distribution is represented by a feature vector of N^3 dimensions (the number of bins in a histogram), and this

is quite large. The larger the feature space the slower the match. Secondly, it is unlikely that all the information in a distribution will be useful in calculating a match. Lastly the color histogram depends on the color of the light. Moving from red to blue illumination causes the color distributions to shift and results in poor match success[3].

2.3 Statistical Moments

If color distributions are described well by a small number of statistical features then comparing these features should suffice for determining distribution similarity. Suppose we must characterize the color distribution in an image I by one single color. A good candidate (and the obvious one) is the mean image color,

$$\underline{\mu}(I) = \frac{1}{M}\sum_{j=1}^{M}(r_j^I, g_j^I, b_j^I)^T \tag{3}$$

where M is the number of pixels in the image. The variance or spread of colors about the mean also captures a lot of information about the color distribution. The variance in the red channel $\sigma_r^2(I)$ is defined as

$$\sigma_r^2(I) = \frac{1}{M}\sum_{i=1}^{M}(r_i^I - \mu_r(I))^2 \tag{4}$$

The covariance $\sigma_r\sigma_g(I)$ between the red and green color channels is defined as

$$\sigma_r\sigma_g(I) = \frac{1}{M}\sum_{i=1}^{M}(r_i^I - \mu_r(I))(g_i^I - \mu_g(I)) \tag{5}$$

Similarly, $\sigma_g^2(I)$, $\sigma_b^2(I)$, $\sigma_r\sigma_b(I)$ and $\sigma_g\sigma_b(I)$ can be defined. These variances and covariances are usually grouped together into a covariance matrix $\Sigma(I)$:

$$\Sigma(I) = \begin{bmatrix} \sigma_r^2(I) & \sigma_r\sigma_g(I) & \sigma_r\sigma_b(I) \\ \sigma_r\sigma_g(I) & \sigma_g^2(I) & \sigma_g\sigma_b(I) \\ \sigma_r\sigma_b(I) & \sigma_g\sigma_b(I) & \sigma_b^2(I) \end{bmatrix} \tag{6a}$$

Suppose we represent an image I as an $M \times 3$ matrix, where the ith row contains the ith rgb triplet. If the mean image color is the zero-vector, $\underline{\mu}(I) = (0,0,0)^T$, then Equation (6a) can be rewritten in matrix notation:

$$\Sigma(I) = \frac{I^T I}{M} \tag{6b}$$

where ^{T} denotes matrix transpose. The covariance relationship given in Equation (6b) will prove very useful in subsequent discussion.

The mean color $\underline{\mu}(I)$ is called the first-order moment of the distribution of image colors and the covariance matrix $\Sigma(I)$ is composed of the second order moments. Third and higher order moments can be calculated in an analogous manner. For example the third order moment of the r color channel is

$s_r(I) = \frac{1}{M}\sum_{i=1}^{M}(r_i^I - \mu_r(I))^3$. This moment captures the *skew* of the red response distribution and is measure of the degree of symmetry about the mean red response. The green and blue skews $s_g(I)$ and $s_b(I)$ are similarly defined. In general the nth order moments of a color distribution is defined as:

$$\frac{1}{M}\sum_{i=1}^{M}(r_i^I - \mu_r(I))^\alpha (g_i^I - \mu_g(I))^\beta (b_i^I - \mu_b(I))^\gamma \tag{7}$$

where $\alpha, \beta, \gamma > 0$ and $\alpha + \beta + \gamma = n$.

Roughly speaking low-order moments give a coarse description of the distribution. More and more distribution details are unveiled as one progresses through the higher moments[11]. Two observations stem from this. First, for color-based object recognition low-order moments capture the most useful information. For example, low-order moments are less effected by confounding processes such as image noise or highlights.

Stricker and Orengo[12], have presented experimental evidence that color distributions can be represented by low-order moments. They show that the features $\underline{\mu}(I)$, $\sigma_r^2(I)$, $\sigma_g^2(I)$, $\sigma_b^2(I)$, $s_r(I)$, $s_g(I)$ and $s_b(I)$ provide an effective index into a large image database. Unfortunately these low-order moments of color distributions are suitable for recognition only if objects are always viewed under the same colored light. A white piece of paper viewed under reddish and bluish lights have predominantly red and blue color distributions respectively. What we really need is descriptors of the low-order information that do not depend on the illuminant in this way.

2.4 Finite Dimensional Models

Both illuminant spectral power distribution functions and surface spectral reflectance functions are described well by finite-dimensional models of low dimension. A surface reflectance vector $S(\lambda)$ can be approximated as:

$$S(\lambda) \approx \sum_{i=1}^{d_S} S_i(\lambda)\sigma_i \tag{8}$$

where $S_i(\lambda)$ is a basis function and $\underline{\sigma}$ is a d_S-component column vector of weights. Similarly each illuminant can be written as:

$$E(\lambda) \approx \sum_{j=1}^{d_E} E_j(\lambda)\epsilon_j \tag{9}$$

where $E_j(\lambda)$ is a basis function and $\underline{\epsilon}$ is a d_E dimensional vector of weights.

Given a finite-dimensional approximations to surface reflectance and illumination, the color response eqn. (1) can be rewritten as a matrix equation. A *Lighting Matrix* $\Lambda(\underline{\epsilon})$[7] maps reflectances, defined by the $\underline{\sigma}$ vector, onto a corresponding response vector:

$$\underline{p} = \Lambda(\underline{\epsilon})\underline{\sigma} \tag{10}$$

where $\Lambda(\underline{\epsilon})_{ij} = \int_\omega R_i(\lambda)E(\lambda)S_j(\lambda)d\lambda$. The lighting matrix depends on the illuminant weighting vector $\underline{\epsilon}$ to specify $E(\lambda)$ via eqn. (9). If surface reflectances are 3-dimensional then every $\Lambda(\underline{\epsilon})$ is a 3×3 matrix. It follows that response triples obtained under one light can be mapped to those of another by a 3×3 matrix.

$$\underline{p}^1 = \Lambda(\underline{\epsilon}^1)\underline{\sigma}\,,\ \underline{p}^2 = \Lambda(\underline{\epsilon}^2)\underline{\sigma} \Rightarrow \underline{p}^2 = \Lambda(\underline{\epsilon}^2)[\Lambda(\underline{\epsilon}^1)]^{-1}\underline{p}^1 \tag{11}$$

Studies[6, 8] have shown that a 3-dimensional model is quite reasonable. Thus it follows that the color distributions of the same surfaces viewed under two illuminants are linearly related to a good approximation.

2.5 Illuminant Invariant Moments

Taubin and Cooper [14] have recently developed efficient algorithms for the computation of vectors of affine moment (or algebraic) invariants of functions. These vectors are invariant to affine transformations of the function which, as Healey and Slater observed, may make them a suitable illuminant-invariant representation for color distributions.

There are two steps in calculating Taubin and Cooper's invariants. First the distribution is manipulated such that its statistics are standardized in some sense. Second, features which are independent of the position of the standardized distribution are extracted.

Standardizing the distribution's statistics is best understood by example. Let I_1 and I_2 be $M \times 3$ matrices denoting the color images of some scene viewed under a pair of illuminants (the M rgb triplets in each image are placed in the rows of the matrices) where first the mean image color has been subtracted in both cases. Thus, $\underline{\mu}(I_1) = (0,0,0)^T$ and $\underline{\mu}(I_2) = (0,0,0)^T$. So long as reflectances are approximately 3-dimensional the two images are related by a 3×3 matrix $\mathcal{M}$

$$I_2 \approx I_1\mathcal{M} \tag{12}$$

I_1 and I_2 are standardized by transforming them by the matrices $\mathcal{O}_1$ and $\mathcal{O}_2$ such that their column spaces are orthonormal,

$$\mathcal{O}_1^T I_1^T I_1 \mathcal{O}_1 = \mathcal{O}_2^T I_2^T I_2 \mathcal{O}_2 = \mathcal{I} \tag{13}$$

Since $I_1\mathcal{O}_1$ and $I_2\mathcal{O}_2$ are orthonormal they differ only by a rotation and represent the same color distributions with respect to different coordinate axes. Thus the basic *shape* of the distributions is the same.

The second step in Taubin and Cooper's method is to extract features from the standardized distributions $I_1\mathcal{O}_1$ and $I_2\mathcal{O}_2$ which are independent of the coordinate frame. The precise details of their method do not concern us—it suffices that the invariants exist. However, we must ask whether these invariants are expressive; that is, do they convey useful information?

To explore this question let us examine the matrices in equation (13) more closely. From Equation (6b) it follows that:

$$\Sigma(I_1\mathcal{O}_1) = \frac{\mathcal{O}_1^T I_1^T I_1 \mathcal{O}_1}{M} = \frac{\mathcal{I}}{M} \tag{14}$$

where $\mathcal{I}$ denotes the identity matrix. That is the covariance matrix of all standardized images is equal to the scaled identity matrix. This is always the case regardless of the starting image statistics. Thus all the low-order statistics—those which convey the most useful information about the distribution—have been lost through the need to discount the effect of the illuminant. It follows that only high order moments can be extracted from $I_1\mathcal{O}_1$ and $I_2\mathcal{O}_2$. As discussed above, we expect that these will not suffice for reliable recognition and this prediction is borne out by experiment in section 4.

3 Distributions angles

While finite-dimensional models are a useful tool for investigating colors under a changing illuminant they do not tell the whole story. Indeed it turns out that a restricted subset of the possible linear transforms correspond to plausible illuminant changes. This observation allows us to extract useful illuminant-invariant statistics from color distributions.

Suppose that the sensor sensitivities of the color camera are delta functions, $R_k(\lambda) = \delta(\lambda - \lambda_k)$. In this case, the camera responses p_k and q_k generated by an arbitrary surface $S_j(\lambda)$ viewed under illuminants $E_1(\lambda)$ and $E_2(\lambda)$ are:

$$p_k = S_j(\lambda_k)E_1(\lambda_k) \ , \ q_k = S_j(\lambda_k)E_2(\lambda_k) \tag{15}$$

It is immediate that

$$q_k = \frac{E_2(\lambda_k)}{E_1(\lambda_k)} p_k \tag{16}$$

Since (16) no longer involves the reflectance function $S_j(\lambda)$ the camera responses induced by any surface are related by the same scaling factor $\frac{E_2(\lambda_k)}{E_1(\lambda_k)}$. Combining the scalings for each sensor class into a diagonal matrix, (16) can be expressed as:

$$\underline{q} = \mathcal{D}\underline{p} \quad (\mathcal{D}_{kk} = \frac{E_2(\lambda_k)}{E_1(\lambda_k)} \ k = 1,2,3) \tag{17}$$

Thus for narrow-band sensors illuminant change is exactly modelled by a diagonal matrix and the full generality of a 3×3 linear model is not required.

Let us consider the problem of extracting invariant features from a color distribution under a diagonal model of illuminant change. We follow the basic approach of Taubin and Cooper in that we first standardize the statistics of the color distribution and then extract the statistical features. Under the diagonal model Equation (17), the relationship between a pair of images can be rewritten as:

$$I_2 = I_1\mathcal{D} \tag{18}$$

where $\mathcal{D}$ is a diagonal matrix. By (18) the corresponding columns of I_1 and I_2 are vectors in the same direction but of different length so the distributions can be standardized by normalizing the lengths of the columns of I_1 and I_2. We define a function $N()$ for carrying out the column normalization:

$$N(I_1) = I_1\mathcal{D}_N \tag{19}$$

where the ith diagonal entry of the diagonal matrix $\mathcal{D}_N$ is equal to the reciporical of the length of the ith column of I_1.

$$[\mathcal{D}_N]_{ii} = \frac{1}{|[I_1]_i|} \tag{20}$$

$[]_i$ and $[]_{ij}$ denote the ith column and ijth element of a matrix respectively. The normalized distributions of I_1 and I_2 are equal:

$$N(I_1) = N(I_2) \tag{21}$$

The covariance matrix of $N(I_1)$ equals:

$$\Sigma(N(I_1)) = \frac{[N(I_1)]^T N(I_1)}{M} = \frac{1}{M}\begin{bmatrix} 1 & M\sigma_r\sigma_g & M\sigma_r\sigma_b \\ M\sigma_r\sigma_g & 1 & M\sigma_g\sigma_b \\ M\sigma_r\sigma_b & M\sigma_g\sigma_b & 1 \end{bmatrix} \tag{22}$$

Note the off-diagonal terms are non-zero, so under a diagonal model of illuminant change the color distributions can be standardized while preserving 3 of the 6 second-order moments, namely the covariances $\sigma_r\sigma_g$, $\sigma_r\sigma_b$ and $\sigma_g\sigma_b$. This contrasts favourably with standardization under a linear model of illuminant where all second-order moments are lost. Note the covariances will not be the same as those defined for the pre-standardized distribution i.e. the covariance terms in (22) are not equal to those in (6a).

Consider the geometric meaning of the covariance terms. The ijth entry in $M\Sigma(N(I_1))$ equals the dot-product of the i and jth th columns of $N(I_1)$. Because each column of $N(I_1)$ is unit length it follows that each dot-product equals the cosine of the angle between the i and jth columns. The cosine function is non-linear which is inappropriate for indexing. Thus we calculate the inverse cosine of the covariance terms in (24) effectively linearizing the feature giving us the *angles of a color distribution.*

$$\phi_{ij}(N(I_1)) = \cos^{-1}([M\Sigma(N(I_1)]_{ij}) \quad (i \neq j) \tag{23}$$

A distribution is represented by its three angles $\phi_{12}(N(I_1))$, $\phi_{13}(N(I_1))$ and $\phi_{23}(N(I_1))$. The 3-tuple of the 3 distribution angles for a color distribution I is denoted $\underline{\phi}(I)$. The distance between distributions I and J is calculated by:

$$||\underline{\phi}(I) - \underline{\phi}(J)||_F \tag{24}$$

where $_F$ denotes the Frobenius norm; that is, the distance between distributions is calculated as the root-mean square error between the respective vectors of color angles.

3.1 Relaxing the Narrow-band assumption

If the camera sensors are not narrow-band then the analysis (15) through (17) does not hold and a single diagonal matrix will not relate sensor responses across a change in illumination. However, Finlayson et al.[2, 1] have shown that in this case a generalized diagonal matrix can be used instead. A generalized diagonal matrix is defined as $\mathcal{T}\mathcal{D}\mathcal{T}^{-1}$, where $\mathcal{T}$ is fixed and $\mathcal{D}$ varies with illumination. Under the generalized scheme, images under different illuminants are related by

$$I_1 \approx I_2\mathcal{T}\mathcal{D}\mathcal{T}^{-1} \tag{25}$$

The relationship in (25) holds exactly if illumination and reflectances are well described by 2- and 3-dimensional linear models[2]. Because *2–3* conditions roughly hold in practice the generalized diagonal relationship describes illuminant change for all sensor sets. Equation (25) can be rewritten making the role of the diagonal matrix explicit:

$$I_1\mathcal{T} \approx I_2\mathcal{T}\mathcal{D} \tag{26}$$

It follows that the angles $\underline{\phi}(N(I_1\mathcal{T}))$ are approximately illuminant-invariant features of color distributions. Since the cameras used in the experiments reported later do in fact have quite narrow-band sensor sensor sensitivities, $\mathcal{T}$ is set to the identity matrix.

3.2 Color-edge distribution angles

Let us define a color edge map as an image convolved with a Laplacian of Gaussian filter in which the usual two-dimensional filter is replicated for each of the three image bands. Denoting the convolution filter $\nabla^2 G$ the edge map of the image I is written as $\nabla^2 G \star I$ where $\star$ represents convolution. Where, as before $\nabla^2 G \star I$ can be thought of as an $M \times 3$ matrix. Because convolution is a linear operator the edge maps of the same scene viewed under two illuminants are related by a diagonal matrix:

$$\nabla^2 G \star I_2 \;=\; \nabla^2 G \star I_1\mathcal{D} \tag{27}$$

It follows that the angles $\underline{\phi}(N(\nabla^2 G \star I))$ encode second-order moment information about the color-edge distribution and are illuminant invariant. If color and color-edge distribution angles encode distinct information then we can expect that used together they will out perform recognition using either alone.

3.3 Properties of distribution angles

Distribution angles (either of colors or color edges) do not depend on the spatial characteristics of an image. In particular they do not depend on the order of the rows in I_1 or $\nabla^2 G \star I_1$. That this is so is clear from the definition of a moment in Equation (7) since a moment is a sum of terms, with each term calculated on a per-pixel basis. Distribution angles are also independent of scale

since $N(I_1) = N(I_1 k)$ for any non-zero k. Because distribution angles are independent of image spatial characteristics and image scale, we can expect angular indexing to recognize an object in different contexts such as when it is placed at different viewing distances or is rotated about the optical axis.

4 Results

The invariants described in section 3 were used as cues for object recognition. They were tested on two published sets of color images[5, 13]. Results are presented for color angle invariants, color-edge angle invariants and their combination. Three existing distribution-based techniques—color indexing, color constant color indexing and Healey and Slater's moment approach—are applied to the same data sets for comparison.

4.1 Swain's Database

Swain's model database consists of 66 images of objects. However, because ratio invariants are ill-defined for images containing saturated pixels, eleven of the images with saturated pixels were pruned from the data set leaving 55 images. The same whitish illumination was used for all objects. A set of an additional 24 images of the same objects but viewed in different poses and with small amounts of deformation (e.g. a rumpled T-shirt) is then used to test the recognition algorithm. The test images are shown (in black and white) in Figure 1. The recognition rankings for color indexing, color constant color indexing (denoted CCCI in Tables 1 and 2), and Healey and Slater's moment-based method are tabulated along with that of color angular indexing in Table 1. Rank is defined to be the position of the correct match in the sorted list of match values. Thus, a match rank of 1 indicates correct recognition, a rank of 2 means that the correct answer was the second best and so on.

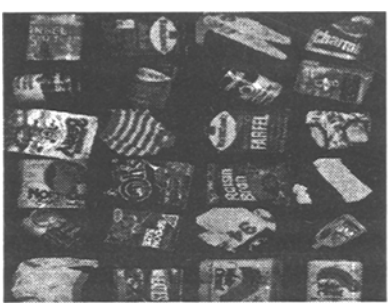

Fig. 1. 24 of Swain's images.

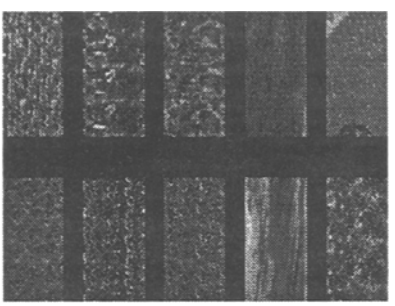

Fig. 2. Healey's texture images.

Algorithm	*Rankings*			
	1	*2*	*3*	*> 3*
Color angles	16	5	2	1
Edge angles	17	3	3	1
Color and edge angles	21	2	1	0
Color Indexing	23	1	0	0
CCCI	22	2	0	0
Healey's moments	7	7	3	7

Table 1. Object database performance.

Algorithm	*Rankings*			
	1	*2*	*3*	*> 3*
Color angles	124	45	29	32
Edge angles	222	8	0	0
Color and edge angles	224	6	0	0
Color indexing	74	21	27	108
CCCI	120	37	21	52
Healey's moments	121	40	20	49

Table 2. Texture database performance.

It is evident that Healey and Slater's higher-order moments based approach delivers poor performance. Note that 7 objects are matched with a greater than 3 rank. The color distribution angles and edge angles, used independently, give reasonable performance with 16 and 17 objects recognized correctly in each case. The combination of both, however, performs very well, comparable with the almost flawless recognition provided by color indexing and color constant color indexing. However, the latter two methods represent objects using a 4096 element feature vector (histogram bin counts).

4.2 Healey and Wang's texture database

Will color angular indexing successfully recognize the colored textures in Healey and Wang's texture data set[5]? This model data set contains ten images of natural textures viewed under white light; they are shown in (black and white) Figure 2. In addition to the model base set of 10 images, 30 other images were taken of the same textures but through 3 separate colored filters placed in front of the camera. This is equivalent to placing the filters in front of the illuminant so it models illumination change. The filters used had narrow pass bands in the blue, green and red regions of the spectrum. Such filters represent quite extreme illuminants and provide a stringent test for the illuminant invariance of the angular index.

Each of the 40 images (10 model and 30 test) was then rotated by 30°, 45°, 60°, 90° and 110° resulting in 240 images in total. Note the angle invariants of rotated textures are not trivially invariant because they are calculated with respect to a square image window so there is a windowing effect. The total test database consists of 230 images: the 30 test images in all 6 orientations and the model base in 5 orientations (all orientations except 0°). Results for the various algorithms are shown in Table 2.

Once again, recognition rates for color angle distributions alone are poor with almost half the textures not being recognized. Color angular indexing with the color and edge angle distributions yields the best results, with all but six of the textures being correctly identified. Note also that, color edge angles by themselves deliver excellent recognition. All the other methods, color indexing, color constant color indexing and Healey and Slater's moment based method, perform very poorly.

5 Conclusion

We have described a new color-based approach to object recognition called *color angular indexing* in which objects are represented by just 6 features: three color distribution angles and three color-edge distribution angles. Our experiments with real images on data bases of several hundred images show that colour angular indexing provides excellent recognition rates for a wide variety of objects and textures even under modest change in orientation and substantial change in illumination conditions.

References

1. G.D. Finlayson, M.S. Drew, and B.V. Funt. Color constancy: Generalized diagonal transforms suffice. *J. Opt. Soc. Am. A*, 11:3011–3020, 1994.
2. G.D. Finlayson, M.S. Drew, and B.V. Funt. Spectral sharpening: Sensor transformations for improved color constancy. *J. Opt. Soc. Am. A*, 11(5):1553–1563, May 1994.
3. B.V. Funt and G.D. Finlayson. Color constant color indexing. *IEEE transactions on Pattern analysis and Machine Intelligence*, 1995.
4. 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*, 11(11):3003–3010, November 1994.
5. G. Healey and L. Wang. The illumination-invariant recognition of texture in color images. *Journal of the optical society of America A*, 12(9):1877–1883, 1995.
6. L.T. Maloney. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *J. Opt. Soc. Am. A*, 3:1673–1683, 1986.
7. L.T. Maloney and B.A. Wandell. Color constancy: a method for recovering surface spectral reflectance. *J. Opt. Soc. Am. A*, 3:29–33, 1986.
8. D.H. Marimont and B.A. Wandell. Linear models of surface and illuminant spectra. *J. Opt. Soc. Am. A*, 9(11):1905–1913, 92.
9. H. Murase and S.K. Nayar. Visual learning and recognition of 3D objects from appearance. *International Journal of Computer Vision*, 14(1):5–24, 1995.
10. W. Niblack and R. Barber. The QIBC project: Querying images by content using color, texture and shape. In *Storeage and Retrieval for Image and Video Databases I, volume 1908 of SPIE Proceedings Series.* 1993.
11. R.J. Prokop and A.P. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processing*, 54(5):438–460, 1992.
12. M. A. Stricker and M. Orengo. Similarity of color images. In *Storage and Retrieval for Image and Video Databases III*, volume 2420 of *SPIE Proceedings Series*, pages 381–392. Feb. 1995.
13. M.J. Swain and D.H.. Ballard. Color indexing. *International Journal of Computer Vision*, 7(11):11–32, 1991.
14. G. Taubin and D. Cooper. "Object recognition based on moment (or algebraic) invariants". In J. Mundy and A. Zisserman, editors, *Geometric Invariance in Computer Vision*, pages 375–397. MIT Press, Cambridge, Mass., 1992.