

**This paper was published in *Journal of the Optical Society of America A* and is made available as an electronic reprint with the permission of OSA. The paper can be found at the following URL on the OSA website:**

**<http://www.opticsinfobase.org/josaa/abstract.cfm?URI=josaa-26-11-B14>.**

**Systematic or multiple reproduction or distribution to multiple locations via electronic or other means is prohibited and is subject to penalties under law.**

# Approaching ideal observer efficiency in using color to retrieve information from natural scenes

David H. Foster,<sup>1,\*</sup> Iván Marín-Franch,<sup>1,3</sup> Kinjiro Amano,<sup>1</sup> and Sérgio M. C. Nascimento<sup>2</sup>

<sup>1</sup>*School of Electrical and Electronic Engineering, University of Manchester, Manchester M60 1QD, United Kingdom*

<sup>2</sup>*Department of Physics, Gualtar Campus, University of Minho, 4710-057 Braga, Portugal*

<sup>3</sup>*Current address: Department of Biomedical Sciences and Human Oncology, University of Turin, Turin 10126, Italy*

\*Corresponding author: *d.h.foster@manchester.ac.uk*

Received February 2, 2009; revised June 17, 2009; accepted July 12, 2009;  
posted July 15, 2009 (Doc. ID 107121); published August 31, 2009

Variations in illumination on a scene and trichromatic sampling by the eye limit inferences about scene content. The aim of this work was to elucidate these limits in relation to an ideal observer using color signals alone. Simulations were based on 50 hyperspectral images of natural scenes and daylight illuminants with correlated color temperatures 4000 K, 6500 K, and 25,000 K. Estimates were made of the (Shannon) information available from each scene, the redundancies in receptor and postreceptor coding, and the information retrieved by an observer identifying corresponding points across image pairs. For the largest illuminant difference, between 25,000 K and 4000 K, a postreceptor transformation providing minimum redundancy yielded an efficiency of about 80% in the information retrieved. This increased to about 89% when the transformation was optimized directly for information retrieved, corresponding to an equivalent Gaussian noise amplitude of 3.0% or to a mean of  $3.6 \times 10^4$  distinct identifiable points per scene. Using color signals to retrieve information from natural scenes can approach ideal observer efficiency levels. © 2009 Optical Society of America

OCIS codes: 330.1690, 330.1715, 330.1720, 330.1880, 330.4060, 110.3055.

## 1. INTRODUCTION

“All vision is colour vision, for it is only by observing differences of colour that we distinguish the forms of objects.” James Clerk Maxwell (1871)

Although color may be a compelling cue to the identity of objects, it is not entirely reliable. This is because the light reflected from the same surfaces changes with changes in the spectrum of the illumination on the scene, and these changes cannot be fully discounted by adaptation in the eye. Conversely, under the same illumination, light reflected from different surfaces can produce the same responses in the receptors of the eye, a phenomenon known as metamerism [1], which is a consequence of the nature of trichromatic sampling. In short, there is no certain one-to-one correspondence between surfaces and receptor responses.

This physical uncertainty sets limits on the information that can be retrieved by an observer using color signals alone to identify the constituent elements of a scene, independent of accidents of position. A previous study [2] estimated possible upper bounds on the information available but did not establish a least upper bound, and another study [3] estimated the approximate information retrieved from small samples of varying size, but not its asymptotic value. The aim of this work was to elucidate both of these limits in a computational analysis of observer performance. Natural scenes and daylight illuminants were used to generate image data because of their ubiquity, complexity, and influence in the development of the eye [4,5].

Information was interpreted in the sense of Shannon [6–8], i.e., as a measure of the reduction in uncertainty or entropy. Previous work (e.g., [9]) has explored the issue of optimal coding of color information in relation to the decorrelation of receptor signals. An advantage of an information-theoretic approach is that it deals with signal dependencies that are more general than linear ones. Although independence implies the absence of correlation, the converse need not.

The organization of this work was as follows. The first step was to define the entropy associated with an image of a scene represented by triplets of long-, medium-, and short-wavelength-sensitive cone responses and then define the information available from the scene itself. Under very general conditions, the transformation of cone responses by adaptation and by postreceptor cone-opponent interactions leaves information unaltered, but this transformation does affect the redundancy of the coding [4]. The second step was to make explicit the sources of this redundancy. The third step was to define an independent estimate of the information retrieved by an observer making the best trichromatic matches possible in order to identify spatially corresponding points across images of a scene under different illuminants. The assumptions underlying this performance are set out later. Unlike the information available, the information retrieved, defined for finite samples of points drawn from the image, depends on the nature of the postreceptor coding. The fourth step was to make numerical estimates of the required quantities: the information available, the redundancies in coding at receptor and postreceptor levels, and the information retrieved as a function of sample

size. The final step was to optimize the postreceptoral transformation to minimize redundancy, and, independently, to maximize the information retrieved, and then to compare the results with the information available.

For the largest illuminant difference, between a correlated color temperature (CCT) of 25,000 K and one of 4000 K, the transformation providing minimum redundancy allowed about 14 bits to be retrieved of the 17 bits available, equivalent to an efficiency of about 80%. When the transformation was optimized for maximum information retrieved, the efficiency increased to about 89%, corresponding to an equivalent Gaussian noise amplitude of 3.0% or to a mean of  $3.6 \times 10^4$  distinct identifiable points per scene. Using color signals to retrieve information from natural scenes can thus approach ideal observer efficiency levels.

Limits on observer performance need not, of course, be quantified only with information-theoretic measures, whether in the sense of Shannon or otherwise. Indeed, historically, ideal observer performance in discrimination and identification tasks has been expressed largely within a signal-detection framework [10], where information-theoretic considerations have been uncommon. One rationale has been that the information transmitted is unchanged by a systematic reassignment of responses ([11], p. 132), but in the present work, this property is actually exploited in the calculation of the information available.

## 2. INFORMATION MEASURES FOR TRICHROMATIC REPRESENTATIONS

### A. Probability Densities

Suppose that at each point  $(x,y)$  in a scene, the reflected light at the eye has spectral radiance  $c(\lambda;x,y)$ , where wavelength  $\lambda$  ranges over the visible spectrum. This spectral radiance  $c(\lambda;x,y)$  may be treated as the product of a spatially uniform global illumination on the scene with incident spectral radiance  $e(\lambda)$  and an effective spectral reflectance  $r(\lambda;x,y)$  at each  $(x,y)$ . Such a representation does not distinguish between variations in spectral reflectance and variations in surface orientation, occlusion, and mutual illumination, but this is unimportant here; see comments in Subsection 3.A and discussion in [12], Appendix A.

For this spectral radiance  $c(\lambda;x,y)$ , the respective responses,  $l$ ,  $m$ , and  $s$ , of the long-, medium-, and short-wavelength-sensitive (L, M, and S) cones, with corneal spectral sensitivities  $\bar{l}(\lambda)$ ,  $\bar{m}(\lambda)$ , and  $\bar{s}(\lambda)$ , are therefore given by

$$\begin{aligned} l(x,y) &= \int \bar{l}(\lambda)c(\lambda;x,y)d\lambda, \\ m(x,y) &= \int \bar{m}(\lambda)c(\lambda;x,y)d\lambda, \\ s(x,y) &= \int \bar{s}(\lambda)c(\lambda;x,y)d\lambda, \end{aligned} \quad (1)$$

where the integrals are evaluated over the visible spectrum. For the purposes of this analysis, the actual positions  $(x,y)$  in Eq. (1) are irrelevant, and the triplet of cone responses  $(l,m,s)$  may be treated as the values of a trivariate continuous random variable  $(L,M,S)$  [9]. Let the probability density function (pdf) of  $(L,M,S)$  be  $f$ . Then the (marginal) pdf of the random variable  $L$  is given by  $\int f(l,m,s)dm ds$ , and analogously for the random variables  $M$  and  $S$ . Figure 1(a) shows an image of a natural scene under a daylight illuminant with CCT 6500 K, and Figs. 1(b)–1(d) show the corresponding naïve estimates of the pdfs of  $L$ ,  $M$ , and  $S$ .

### B. Entropy and Mutual Information

Entropy may be thought of generally as a measure of uncertainty of a random variable, whether discrete or continuous, but its use with continuous variables requires some care [8]. For the continuous trivariate random variable  $(L,M,S)$ , the Shannon differential entropy is defined [8] by

$$h(L,M,S) = - \int f(l,m,s) \log f(l,m,s) dl dm ds, \quad (2)$$

where, here and elsewhere, the integral is assumed to exist and is taken over the set for which  $f(l,m,s) > 0$ . If the logarithm is to the base 2, then the differential entropy is measured in bits. The differential entropies of the individual variables  $L$ ,  $M$  and  $S$ , and of their combinations, are defined analogously. Although the value of the differential entropy depends on the space in which the vari-

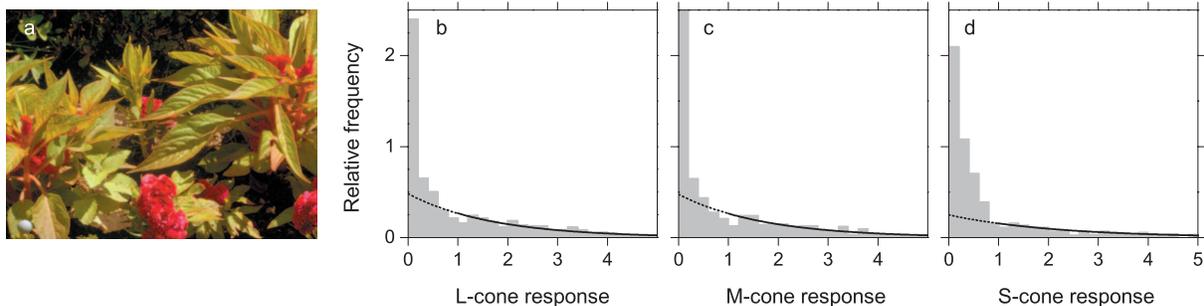


Fig. 1. Image statistics. The image in panel a is of a scene under a daylight illuminant with correlated color temperature 6500 K. The histograms in panels b, c, and d are naïve estimates of the response probability density functions of L, M, and S cones based on a sample of 1000 points drawn randomly from the image a. The small gray sphere at the bottom left of the image was excluded from the analysis by a mask. Responses were normalized to a mean of 1.0 (equivalent to von Kries scaling; abscissas and ordinates have been truncated). The continuous smooth curves are exponential model density functions fitted by maximum likelihood to responses  $> 1.0$ , with dotted lines indicating the fit over the remainder of the range.

ables  $L$ ,  $M$ , and  $S$  range, this dependence disappears when differences in differential entropies are taken [8], as in the calculation of mutual information, whose definition follows.

Two different illuminants, with incident spectral radiances  $e_1$  and  $e_2$ , say, will usually give rise to different random variables  $(L_1, M_1, S_1)$  and  $(L_2, M_2, S_2)$ , with pdfs  $f_1$  and  $f_2$ , say. Each of these variables provides a different representation of the scene, not merely because of a difference in mean reflected spectrum but also because of the uncertainties in a trichromatic representation, as noted in the Introduction. The information that the representations have in common is essentially the overlap of the separate entropies. This overlap is the mutual information.

More formally, the mutual information  $I(L_1, M_1, S_1; L_2, M_2, S_2)$  between  $(L_1, M_1, S_1)$  and  $(L_2, M_2, S_2)$  is given [8] by

$$I(L_1, M_1, S_1; L_2, M_2, S_2) = h(L_1, M_1, S_1) + h(L_2, M_2, S_2) - h(L_1, M_1, S_1, L_2, M_2, S_2). \quad (3)$$

This quantity  $I(L_1, M_1, S_1; L_2, M_2, S_2)$  may be interpreted as the reduction in uncertainty about  $(L_1, M_1, S_1)$  given  $(L_2, M_2, S_2)$ , and vice versa, i.e.,  $I(L_1, M_1, S_1; L_2, M_2, S_2) = I(L_2, M_2, S_2; L_1, M_1, S_1)$ . More importantly here,  $I(L_1, M_1, S_1; L_2, M_2, S_2)$  represents the invariant elements of the scene, for it quantifies the mean number of points in the scene that, in principle, are available for reliable identification across the two illuminants. If  $I(L_1, M_1, S_1; L_2, M_2, S_2) = a$ , say, in bits, then this number is  $2^a$  [8]. Any additional noise introduced at either receptor or postreceptor levels [13,14] serves only to diminish this value. If  $(L_1, M_1, S_1)$  and  $(L_2, M_2, S_2)$  are independent of each other, then  $I(L_1, M_1, S_1; L_2, M_2, S_2) = 0$ , and no points in the scene are reliably identified.

Another way of interpreting the value of  $I(L_1, M_1, S_1; L_2, M_2, S_2)$  is in terms of equivalent Gaussian signals and noise. Suppose that  $L_2 = L_1 + Z$ , where the signal  $L_1$  and noise  $Z$  are distributed independently and normally with variances  $\sigma^2$  and  $\sigma_z^2$ , respectively, and analogously for  $M_1, M_2$  and  $S_1, S_2$ . For these Gaussian signals (the actual distributions are strongly non-Gaussian),  $I(L_1, M_1, S_1; L_2, M_2, S_2) = (3/2) \log(1 + \sigma^2/\sigma_z^2)$  [8]. If  $I(L_1, M_1, S_1; L_2, M_2, S_2) = a$ , again in bits, then the percentage noise amplitude  $100\sigma_N/\sigma$  is given by  $100(2^{2a/3} - 1)^{-1/2} \approx 100 \times 2^{-a/3}$ .

These estimates of mutual information, which concern the stable reflecting properties of scenes independent of illuminant, should be distinguished from the optimal representation of colors in images of a scene under a fixed illuminant (e.g., [14]).

### C. Postreceptor Interactions

As a result of postreceptor cone-opponent interactions, each of the triplets of cone responses  $(l, m, s)$  yields a triplet of postreceptor responses, denoted by  $(l^\#, m^\#, s^\#)$ , to anticipate the fact that their optimum spectral sensitivities are spectrally sharpened forms of the L-, M-, and S-cone corneal spectral sensitivities [15,16]. The effect of these interactions may be represented by a linear invertible transformation  $T$  [17]; thus  $(l_1^\#, m_1^\#, s_1^\#) = T(l_1, m_1, s_1)$

for illuminant  $e_1$  and  $(l_2^\#, m_2^\#, s_2^\#) = T(l_2, m_2, s_2)$  for illuminant  $e_2$ . Examples are given later. The random variables  $(L_1, M_1, S_1)$  and  $(L_2, M_2, S_2)$  are therefore transformed to postreceptor random variables

$$\begin{aligned} (L_1^\#, M_1^\#, S_1^\#) &= T(L_1, M_1, S_1), \\ (L_2^\#, M_2^\#, S_2^\#) &= T(L_2, M_2, S_2). \end{aligned} \quad (4)$$

Responses were assumed to be subject to perfect von Kries scaling [18,19] (compare [20]), and variables were therefore normalized to a mean of 1.0. Provided that the transformation  $T$  is smooth (differentiable) and invertible (necessarily true by definition), the mutual information between  $(L_1^\#, M_1^\#, S_1^\#)$  and  $(L_2^\#, M_2^\#, S_2^\#)$  is the same as that at a receptor level (Appendix A); that is,

$$I(L_1^\#, M_1^\#, S_1^\#; L_2^\#, M_2^\#, S_2^\#) = I(L_1, M_1, S_1; L_2, M_2, S_2). \quad (5)$$

For brevity, let  $I(L_1, M_1, S_1; L_2, M_2, S_2)$  be denoted by  $I_{LMS}$  and  $I(L_1^\#, M_1^\#, S_1^\#; L_2^\#, M_2^\#, S_2^\#)$  by  $I_{LMS}^\#$ .

This invariance (5) is crucial in that the information estimate (3) is independent of whether it is computed before or after postreceptor interactions; in particular, it does not depend on the specific values of receptor or postreceptor signals, just their pdfs (Appendix A). In principle,  $I_{LMS}$  provides a least upper bound on the information retrieved, which does depend on these values.

### D. Redundancies

Although the information available may be invariant, the redundancies in the coding at receptor and postreceptor levels need not be so, for in general there will be different dependencies between the variables. The dependency between more than two random variables can be quantified by the multi-information [21], and this is used here to measure the redundancy. Thus, with differential entropies defined as in Eq. (2), the redundancy  $R_1 = R(L_1; M_1; S_1)$  in the random variables  $L_1, M_1$ , and  $S_1$  with illuminant  $e_1$  and redundancy  $R_2 = R(L_2; M_2; S_2)$  in the random variables  $L_2, M_2$ , and  $S_2$  with illuminant  $e_2$  are given by

$$\begin{aligned} R_1 &= h(L_1) + h(M_1) + h(S_1) - h(L_1, M_1, S_1), \\ R_2 &= h(L_2) + h(M_2) + h(S_2) - h(L_2, M_2, S_2). \end{aligned} \quad (6)$$

A higher-order redundancy  $R_{12} = R(L_1, L_2; M_1, M_2; S_1, S_2)$  in the pairs of random variables  $(L_1, L_2)$ ,  $(M_1, M_2)$ , and  $(S_1, S_2)$  is given by

$$\begin{aligned} R_{12} &= h(L_1, L_2) + h(M_1, M_2) + h(S_1, S_2) \\ &\quad - h(L_1, L_2, M_1, M_2, S_1, S_2). \end{aligned} \quad (7)$$

There are analogous expressions for the redundancies  $R_1^\#$ ,  $R_2^\#$ , and  $R_{12}^\#$  in the postreceptor variables  $(L_1^\#, M_1^\#, S_1^\#)$  and  $(L_2^\#, M_2^\#, S_2^\#)$ .

The redundancies  $R_1$  and  $R_2$  decrease as the three types of cone responses become more independent of one another, and, for a given level of response dependence, the higher-order redundancy  $R_{12}$  decreases as the changes in cone responses with illuminant changes become more independent of one another [22]. All three redundancies are necessarily nonnegative [21].

Although redundancies at the receptor level are set by the spectral tuning of the L, M, and S cones, at a postreceptor level they depend on the transformation  $T$  [Eq. (4)]. Unfortunately, it is not possible to find a single transformation that simultaneously minimizes the postreceptor redundancies  $R_1^\#$  and  $R_2^\#$ , but it is possible to find one that minimizes the higher-order redundancy  $R_{12}^\#$ , which has a useful side effect.

### E. Decomposing the Information Available

The information available  $I_{LMS}=I(L_1, M_1, S_1; L_2, M_2, S_2)$  can be thought of as having components from individual cone classes, i.e.,  $I_L=I(L_1; L_2)$  between L-cone responses,  $I_M=I(M_1; M_2)$  between M-cone responses, and  $I_S=I(S_1; S_2)$  between S-cone responses, where

$$\begin{aligned} I_L &= h(L_1) + h(L_2) - h(L_1, L_2), \\ I_M &= h(M_1) + h(M_2) - h(M_1, M_2), \\ I_S &= h(S_1) + h(S_2) - h(S_1, S_2). \end{aligned} \quad (8)$$

But their sum  $I_L + I_M + I_S$  does not necessarily coincide with  $I_{LMS}$ . As shown in Appendix B, the difference can be traced to the redundancies defined in Subsection 2.D. That is,

$$I_{LMS} = I_L + I_M + I_S - R_1 - R_2 + R_{12}. \quad (9)$$

An analogous decomposition holds for the postreceptor variables  $(L_1^\#, M_1^\#, S_1^\#)$  and  $(L_2^\#, M_2^\#, S_2^\#)$ ; that is,

$$I_{LMS}^\# = I_L^\# + I_M^\# + I_S^\# - R_1^\# - R_2^\# + R_{12}^\#. \quad (10)$$

All the quantities on the right-hand side of Eq. (10) depend on the transformation  $T$  [Eq. (4)], but, as already noted from Eq. (5), the quantity on the left-hand side of Eq. (10) does not.

### F. Information by Identification

The estimates of mutual information given by Eq. (3) or Eq. (9) depend directly or indirectly on the pdfs  $f_1$  and  $f_2$  of the random variables  $(L_1, M_1, S_1)$  and  $(L_2, M_2, S_2)$ , and on the joint pdf  $f_{12}$  of  $(L_1, M_1, S_1, L_2, M_2, S_2)$ . There is, however, a different approach to calculating information that reflects performance of an observer undertaking an identification task based on trichromatic matching [2].

Given the images in Figs. 2(a) and 2(b) of the scene in Fig. 1(a) under two different illuminants  $e_1$  and  $e_2$ , here daylight illuminants with CCTs of 25,000 K and 4000 K, respectively, draw  $N$  points randomly from the image in Fig. 2(b) (only one such point is shown) and consider the task of finding the spatially corresponding points in the image in Fig. 2(a) solely on the basis of the triplets of cone responses generated, i.e., with a nearest-neighbor criterion in the space of responses. It is emphasized that trichromatic matching at the level of the receptors (or postreceptorally) is not assumed to be necessarily optimum for identifying spatially corresponding points. As noted earlier (Subsection 2.C), responses were assumed to be fully adapted, but to illustrate the effect of illuminant, the images in Figs. 2(a) and 2(b) have not been normalized. Being able to establish a one-to-one spatial correspondence between points independent of illuminant changes is closely related to being able to make color-constancy judgments [23].

In more detail, if the point  $j=1$  in the image in Fig. 2(b) is one of the  $N$  points in the sample, the task is to find the point  $i$  in the image in Fig. 2(a) whose triplet of responses  $(l_1(i), m_1(i), s_1(i))$  is closest to the triplet of responses  $(l_2(j), m_2(j), s_2(j))$ . There are several candidate trichromatic matches in Fig. 2(a), namely, points  $i=1, 2, 3, 4$  (not all are distinguishable), with  $i=1$  being the closest,  $i=2$  the next closest, and so on. The correct identification is the fourth closest. How these incorrect identifications occur is quantified colorimetrically in [3].

In performing this task, the observer is assumed not to be affected by receptor or postreceptor noise or by distractions from signals elsewhere in the scene, or by attention and memory, including previous performance and scene familiarity. The extent to which nearest-neighbor matching is optimal in this task determines how close performance approaches that of an ideal observer [24].

A relevant way to quantify the error in performance is to measure how far the incorrect identification  $i$  is from the correct identification  $j$  in terms of the number  $k$  of other identifications closer to  $j$  than  $i$  [2]. This number obviously depends on the number of alternatives available, that is, on the size  $N$  of the sample. To accommodate the variance-covariance structure of the sample, matching was determined not by the Euclidean distance between signals but by the Mahalanobis distance. The identifica-



Fig. 2. Identification errors across images of the scene in Fig. 1(a) under a daylight illuminant with correlated color temperature a, 25,000 K and b, 4000 K. The points marked 1, 2, 3, 4 (not all are distinguishable) in a are candidate trichromatic matches to point 1 in b, after cone responses were scaled to a mean of 1.0 (von Kries scaling) for the sample (see text). For the purposes of illustration, the images in a and b have not been normalized. The histogram in c is the naive estimate of the probability mass function for identification errors based on a sample of 1000 points drawn randomly from the image. The smooth curve is an exponential model density function fitted to identification errors  $>5.0$ , with a dotted line indicating the fit over the remainder of the range.

tion errors  $k$  may be treated as the values of a discrete random variable, say,  $K$ , with probability mass function (pmf)  $p_k$ . The discrete entropy  $H_N(K)$  of  $K$  is then given [8] by

$$H_N(K) = - \sum p_k \log p_k, \quad (11)$$

where the sum is taken over all  $k$  for which  $p_k > 0$ . Apart from a reordering, which does not affect the value of  $H_N(K)$ , the probabilities  $p_k$  define a transition matrix of a symmetric channel [8]. Figure 2(c) shows the naïve estimate of  $p_k$ , i.e., relative frequency as a function of  $k$ , for the two images in Figs. 2(a) and 2(b). As the  $N$  points are sampled uniformly, the *a priori* probability of any one in particular being chosen is  $1/N$ . The entropy of the sample is therefore the entropy of the constant pmf with value  $1/N$  at each point; that is,  $H(1/N) = \log N$ . The mutual information afforded by the sample is given [8] by the difference between  $H(1/N)$  and  $H_N(K)$ , denoted here by  $I_K(N)$ ; that is,

$$I_K(N) = \log N - H_N(K). \quad (12)$$

This quantity provides a measure of the information retrieved, but unlike the information available [Eq. (3)], it depends on the postreceptoral transformation  $T$  [Eq. (4)], and it does not have a straightforward decomposition of the kind given in Eqs. (9) and (10).

When the random variables  $(L_1, M_1, S_1)$  and  $(L_2, M_2, S_2)$  are identical,  $I_K(N) = \log N$ , since  $H_N(K) = 0$ . When they are independent,  $I_K(N) = 0$ , since  $p_k = 1/N$ , and  $H_N(K) = \log N$ . But  $I_K(N)$  does not grow indefinitely with  $N$ . When  $N$  is small, the risk of mismatch is small, so  $I_K(N)$  is close to  $\log N$ , but as  $N$  increases, the risk of a mismatch increases, with the result that  $I_K(N)$  increases less and less rapidly. The limit of  $I_K(N)$  as  $N$  tends to infinity represents the maximum information retrieved by this procedure.

At a postreceptoral level, the task is analogous: that is, given point  $j$  in the image in Fig. 2(b), to find the point  $i$  in the image in Fig. 2(a) whose triplet of responses  $(l_1^\#(i), m_1^\#(i), s_1^\#(i))$  is closest to the triplet of responses  $(l_2^\#(j), m_2^\#(j), s_2^\#(j))$ . An expression for the discrete entropy analogous to Eq. (11) holds at a postreceptoral level

$$H_N(K^\#) = - \sum p_k^\# \log p_k^\#, \quad (13)$$

and analogous to Eq. (12) for the information retrieved  $I_K^\#(N)$ , that is,

$$I_K^\#(N) = \log N - H_N(K^\#). \quad (14)$$

The asymptotic values of the information retrieved [Eqs. (12) and (14)] may be expressed as a proportions of the information available [Eqs. (9) and (10)], providing a measure of the efficiency of the nearest-neighbor matching procedure [6,25]. Thus, at a receptoral level, the efficiency is given by

$$\lim_{N \rightarrow \infty} I_K(N)/I_{LMS} \quad (15)$$

and, at a postreceptoral level, by

$$\lim_{N \rightarrow \infty} I_K^\#(N)/I_{LMS}^\#. \quad (16)$$

The relationship between information available and information retrieved for a known distribution is considered in Appendix C. Some general results concerning nearest-neighbor decoding are described in [24].

*A priori*, efficiency might be expected to be reasonably good. It is known that von Kries scaling is effective, with or without cone-opponent interactions [16,26], and may be improved by being replaced by a general linear transformation [27] or by a nonlinear von Kries transformation [28].

### 3. EXPERIMENTAL METHODS

#### A. Image Data

Calculations were based on data generated from a set of 50 hyperspectral images of natural rural and urban scenes [12,29]. Each hyperspectral image had dimensions  $\leq 1344 \times 1024$  pixels and spectral range 400–720 nm in 10 nm steps, providing an effective spectral reflectance  $r(\lambda; x, y)$  at each wavelength  $\lambda$  and point  $(x, y)$  (Subsection 2.A).

The raw spectral radiance images, acquired under daylight with CCTs of approximately 4400 K to 8200 K, were corrected for dark noise, spatial nonuniformities, stray light, and any wavelength-dependent variations in magnification or translation [12]. The effective spectral reflectances  $r(\lambda; x, y)$  were obtained by dividing the spectral radiance of the image by the spectral radiance of a small neutral (Munsell N5 or N7) reference surface embedded in the scene and then multiplying by the known spectral reflectance of the surface. The effect of a different daylight was simulated by multiplying  $r(\lambda; x, y)$  at each point  $(x, y)$  by the chosen illuminant spectrum. This procedure is not intended to model geometric changes in the scene as the sun moves across the sky, merely the effects of a change in incident spectrum at any particular surface element. A detailed rationale for this approach and estimates of metamerism with and without masking shadowed regions are given in [12].

No reflectance spectrum coincided numerically with any other within a hyperspectral image, and therefore each spectrum corresponded to a unique element in the scene (within the spatial resolution of the image). As the present approach is neutral with respect to spatial structure, subsets of different reflectances were not grouped according to whether they might have been drawn from different parts of the “same” surface, an issue that is addressed more fully in [12], Subsection 2.D.

In simulations, scenes were illuminated by different daylight illuminants  $e_1(\lambda)$  and  $e_2(\lambda)$ : an extreme pair, as in Fig. 2, with CCTs of 25,000 K and 4000 K, and a more closely spaced pair with CCTs of 4000 K and 6500 K. Still larger illuminant differences could have been considered, but 25,000 K and 4000 K define useful practical limits in the CIE system [30], and represent the sun and sky at different times of the day [31]. As a control, daylight and skylight spectra from the Iberian peninsula [32] were also tested. Differences in estimates of the information available for the CCTs available were  $< 6\%$ .

Triplets of cone responses  $(l, m, s)$  at each point in each image of each scene were calculated according to Eq. (1) from the Stockman *et al.* [33] and Stockman and Sharpe

[34] estimates of corneal cone spectral sensitivities  $\bar{l}(\lambda)$ ,  $\bar{m}(\lambda)$ , and  $\bar{s}(\lambda)$ . Because of the approximately 1.3 pixel line spread function of the camera system [12], images were spatially subsampled, with only alternate pixels being used. Any similarity between spectral reflectances at points in the scene is automatically accounted for in the estimate of the mutual information. Integrations were performed numerically over 400–720 nm at 10 nm intervals (Subsections 2.A and 2.B).

## B. Estimation Algorithms

As already illustrated in Fig. 1, naïve estimates of the pdfs  $f_1$ ,  $f_2$ , and  $f_{12}$  of the random variables  $(L_1, M_1, S_1)$ ,  $(L_2, M_2, S_2)$ , and  $(L_1, M_1, S_1, L_2, M_2, S_2)$  (Subsections 2.A and 2.B) may be obtained by binning, i.e., partitioning the cone-response spaces into finite cells and counting the frequencies of responses in each cell [35]. But systematic errors in estimating the information available then occur [36]; examples of how this bias arises are given in [37,38]. These errors may be minimized by introducing correction terms, but, instead,  $k$ -nearest-neighbor statistics were used here [39,40], the results of which are efficient and adaptive and have minimal bias [36].

Estimates of the information available (3) and its decomposition (9) based on differential entropies (2) were obtained from a modified version of the  $k$ -nearest-neighbor estimator due to Kozachenko and Leonenko [39]. In this modification [22], the entropy estimate was partitioned into a Gaussian component, which has known entropy, and a non-Gaussian component to which the Kozachenko–Leonenko estimator was applied. This offset estimator had the advantage of converging more rapidly and accurately [41] than when the Kozachenko–Leonenko estimator was applied directly to the response variables. Thus, if  $\text{Var}(L, M, S)$  denotes the variance–covariance matrix of  $(L, M, S)$  and  $(L^*, M^*, S^*)$  denotes the product  $[\text{Var}(L, M, S)]^{-1/2} (L, M, S)$ , then the differential entropy  $h(L, M, S)$  is given [8] by

$$h(L, M, S) = h(L^*, M^*, S^*) + \frac{1}{2} \log |\text{Var}(L, M, S)|,$$

where  $|\cdot|$  denotes the determinant. The log-variance term on the right-hand side of this equation is the differential entropy of a trivariate Gaussian random variable with variance–covariance matrix  $\text{Var}(L, M, S)$  [8] less a constant factor. The offset Kozachenko–Leonenko estimator of  $h(L, M, S)$  is the sum of the Kozachenko–Leonenko estimate of  $h(L^*, M^*, S^*)$  and the log-variance term (see Appendix C). Notice that although this estimator is based on a finite sample (the maximum size of which was, in principle,  $1344 \times 1024 = 1,376,256 = 2^{20.39}$ ), it is shown in Appendix C that this did not limit the size of the estimated differential entropy or, in turn, estimates of the number of distinct, identifiable points in a scene.

Systematic errors also occur in estimating the information retrieved from the pmf  $p_k$  of the random variable  $K$  (Subsection 2.F) from each finite sample of size  $N$ . The naïve estimators of the probabilities  $p_k$  in Eq. (11) and  $p_k^\#$  in Eq. (13), obtained by counting the observed frequencies at each value of  $k$ , as illustrated in Fig. 2, are negatively biased. The Kozachenko–Leonenko estimator cannot be used since it applies to an infinite population, and, in-

stead, a bias-corrected version of the naïve estimator due to Grassberger ([42], Eqs. (23) and (27)) was used to estimate the discrete entropies receptorally [Eq. (11)] and postreceptorally [Eq. (13)]. Estimates of the information retrieved receptorally [Eq. (12)] and postreceptorally [Eq. (14)] were then obtained as a function of increasing sample size  $N$ .

Both the offset Kozachenko–Leonenko estimator and the Grassberger estimator were validated against a known distribution (Appendix C).

The coefficients of the  $3 \times 3$  matrix representing the postreceptor transformation  $T$  [Eq. (4)] were optimized according to two cost functions: first, as explained in Subsection 2.D, by minimizing the magnitude of the higher-order redundancy  $R_{12}^\#$  in the information available and, second, in order to test whether it was optimal for matching, by maximizing the limiting value of the information retrieved  $I_K^\#(N)$ . Optimizations were performed with a simplex algorithm with multiple initializations.

## 4. RESULTS AND COMMENT

### A. Information Available

Figure 3(a) shows the decomposition (9) of the information available  $I_{LMS}$  at the level of the receptors. Estimates were for daylight illuminants with CCTs 25,000 K and 4000 K (solid symbols) and 4000 K and 6500 K (open symbols). Means and sample standard deviations (SDs) over the 50 natural scenes are shown for the information available  $I_{LMS}$ ; the individual contributions  $I_L$ ,  $I_M$ , and  $I_S$  of L, M, and S cones; and the redundancies  $R_1$ ,  $R_2$ , and  $R_{12}$ . Figure 3(b) shows the corresponding estimates for the decomposition (10) of the information available  $I_{LMS}^\#$  at a postreceptor level. The postreceptor transformation  $T$  (4) was chosen to minimize the higher-order redundancy  $R_{12}^\#$ , but the other redundancies  $R_1^\#$  and  $R_2^\#$  were also reduced as a consequence.

Other daylight illuminants were also tested. The information available from illuminants with CCTs of 4000 K and 25,000 K was always less than that from any other pair drawn from within the same range.

Both receptor and postreceptor estimates of  $I_{LMS}$  and  $I_{LMS}^\#$  were obtained as the sums of the corresponding components and redundancies (9) and (10), respectively, and, as expected, their mean values were indistinguishable, at 17.1 bits (SD 1.2 bits) for CCTs of 25,000 K and 4000 K and 20.0 bits (SD 1.2 bits) for CCTs of 4000 K and 6500 K. These values also coincided with the result of direct estimation (3), providing a control on the independent estimates on the right-hand sides of Eqs. (9) and (10). The particular values of 17.1 and 20.0 bits correspond to  $1.37 \times 10^5 (=2^{17.1})$  and  $1.03 \times 10^6 (=2^{20.0})$  distinct points per scene that, in principle, are available for identification through the images (Subsection 2.B). The equivalent Gaussian percentage noise amplitudes for these two information estimates were 1.9% and 1.0%, respectively (Subsection 2.B).

At a receptor level, the mean estimates of the redundancies  $R_1 + R_2$  and  $R_{12}$ , as proportions of the information available  $I_{LMS}$ , were 92% and 96% for daylight illuminants with CCTs of 25,000 K and 4000 K, and 78% and 80% for CCTs of 4000 K and 6500 K, respectively. Postre-

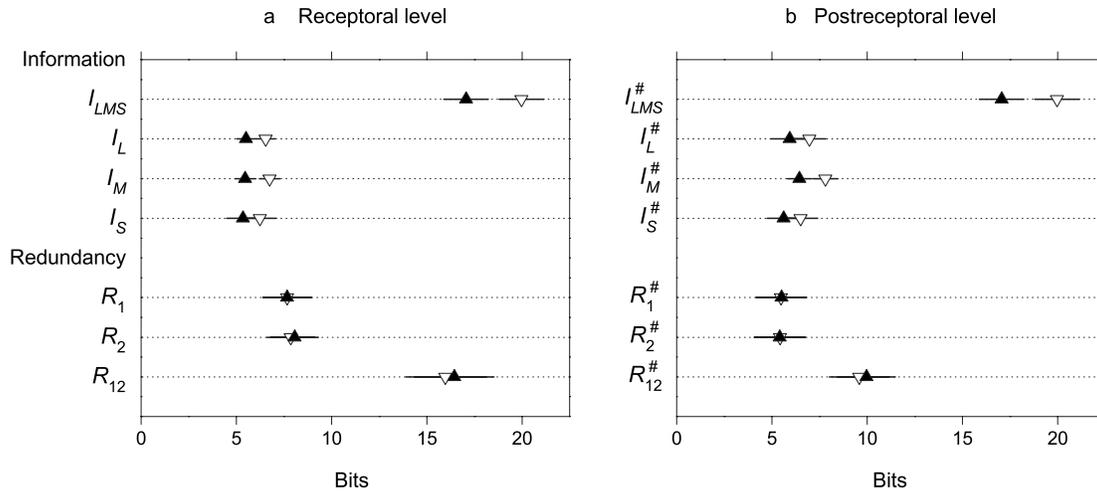


Fig. 3. Decomposition of the mean estimated information available into individual information components and redundancies at a, receptor and b, postreceptor level. The postreceptor transformation  $T$  [Eq. (4)] was optimized for minimum higher-order redundancy  $R_{12}^{\#}$  (b). Means were taken over 50 scenes and for daylight illuminants with correlated color temperatures of 25,000 K and 4000 K (solid symbols) and 4000 K and 6500 K (open symbols). Horizontal bars show  $\pm 1$  sample SD.

ceptorally, with the transformation  $T$  [Eq. (4)] optimized for minimum  $R_{12}^{\#}$ , the mean estimates of  $R_1^{\#} + R_2^{\#}$  and  $R_{12}^{\#}$ , as proportions of the information available, fell to 64% and 58% and to 55% and 48%, respectively.

The mean transformation  $T$  (with sample SDs in parentheses) for the 50 scenes for daylight illuminants with CCTs of 25,000 K and 4000 K is given, in matrix notation, by

$$T = \begin{pmatrix} 1 & -0.931(0.060) & 0.066(0.025) \\ -0.259(0.096) & 1 & -0.156(0.032) \\ 0.003(0.054) & -0.035(0.087) & 1 \end{pmatrix}. \quad (17)$$

For CCTs of 4000 K and 6500 K, the result was similar; that is,

$$T = \begin{pmatrix} 1 & -0.971(0.048) & 0.093(0.028) \\ -0.249(0.056) & 1 & -0.173(0.034) \\ 0.005(0.041) & -0.034(0.070) & 1 \end{pmatrix}. \quad (18)$$

The small SDs indicate the stability of  $T$  over the 50 scenes. The effect of  $T$  on the corneal spectral sensitivities of the  $L$ ,  $M$ , and  $S$  cones is to shift their peaks from 570, 543, and 442 nm at a receptor level to 604, 541, and 442 nm at a postreceptor level [15,16].

### B. Information Retrieved

Figure 4(a) shows, for daylight illuminants with CCTs of 25,000 K and 4000 K, mean estimates of the information retrieved receptorally  $I_K(N)$  (circles) and postreceptorally  $I_K^{\#}(N)$  (squares) as a function of the logarithm of the size  $N$  of the sample. The mean estimate of the information available  $I_{LMS}$  [Eq. (3)] is indicated by the horizontal line. The postreceptor transformation  $T$  was optimized for minimum higher-order redundancy  $R_{12}^{\#}$  (Subsection 3.B). Figure 4(b) shows the corresponding estimates for CCTs of 4000 K and 6500 K.

The mean estimate of  $I_K(N)$  with CCTs of 25,000 K and 4000 K (Fig. 4(a)) reached its asymptotic value of about 7.7 bits (SD 1.0 bits), and with CCTs of 4000 K and 6500 K (Fig. 4(b)), it was still increasing slightly at the largest sample size feasible  $N=2^{18}$ , reaching about 10.4 bits (SD 1.2 bits). The mean estimate of  $I_K^{\#}(N)$  clearly failed to asymptote at  $N=2^{18}$  with both pairs of illuminants.

In the absence of an analytic expression for the asymptotes, an empirical estimate was constructed. The observed dependence of the mean estimate of  $I_K(N)$  on  $N$  with CCTs of 25,000 K and 4000 K was first fitted by a locally weighted linear regression  $g$  (loess; see [43]). The function  $g$  was then scaled and shifted to provide the best fit  $g'$  to the observed dependence of the mean estimate of  $I_K(N)$  on  $N$  with CCTs of 4000 K and 6500 K and of  $I_K^{\#}(N)$  on  $N$  with both illuminant pairs. The bandwidth [43] of  $g$  was adjusted iteratively for minimum mean square error in the fit of  $g'$ , to obtain a common template. With the inclusion of a small ( $\leq 3\%$ ) quadratic correction to the fit of  $g'$ , the root mean square error was 0.05 bit with residual degrees of freedom 11 per data set. The results are shown by the solid curves in Fig. 4, with arrows indicating the estimated asymptotic values. With CCTs of 25,000 K and 4000 K, the estimated asymptote for  $I_K(N)$  was 10.4 bits (SD 1.3 bits), differing from the value at  $N=2^{18}$  by  $< 0.1$  bit. For  $I_K^{\#}(N)$  the estimated asymptote was 13.6 bits (SD 1.7 bits) with CCTs of 25,000 K and 4000 K and 17.0 bits (SD 2.0 bits) with CCTs of 4000 K and 6500 K, each less than the corresponding mean estimate of the information available  $I_{LMS}$ , namely, 17.1 and 20.0 bits, respectively.

These estimated postreceptor asymptotes of 13.6 and 17.0 bits with the two pairs of CCTs correspond to means of  $1.28 \times 10^4$  and  $1.29 \times 10^5$  distinct points per scene that can be identified reliably by postreceptor trichromatic matching across images (Subsection 2.B). The corresponding efficiencies were about 80% and 85% respectively [Eq. (16)], which may be compared with those at the receptors of 45% and 52% respectively [Eq. (15)]. The equivalent

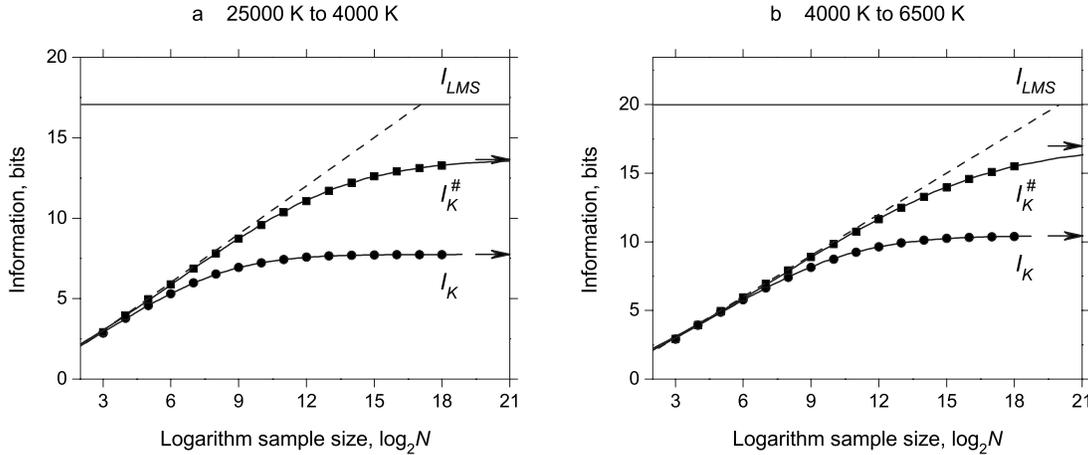


Fig. 4. Information retrieved. The mean estimate of the information retrieved receptorally  $I_K(N)$  (circles) and postreceptorally  $I_K^\#(N)$  (squares) is shown as a function of the logarithm of the size  $N$  of the sample for two daylight illuminants with correlated color temperatures of a, 25,000 K and 4000 K and b, 4000 K and 6500 K. The postreceptoral transformation was optimized for minimum higher-order redundancy  $R_{12}^\#$ . The bounding value of the discrete entropy of a sample of size  $N$  is shown by the dashed line and the mean estimate of the information available  $I_{LMS}$  by the horizontal line (different ordinates in a and b). The smooth curves are locally weighted linear regressions derived from a best-fitting template from which asymptotic values were estimated, shown arrowed. Means were taken over 50 scenes.

Gaussian percentage noise amplitudes (Subsection 2.B) for these two information estimates were 4.3% and 2.0%, respectively.

Figures 5(a) and 5(b) show the corresponding results when the postreceptoral transformation  $T$  was optimized for maximum information retrieved  $I_K^\#(N)$  rather than minimum redundancy  $R_{12}^\#$ . The estimated asymptote for  $I_K^\#(N)$  was larger, at 15.2 bits (SD 1.4 bits) with CCTs of 25,000 K and 4000 K and 18.3 bits (SD 1.7 bits) with CCTs of 4000 K and 6500 K. These two estimates correspond to means of  $3.6 \times 10^4$  and  $3.2 \times 10^5$  distinct points per scene that can be identified reliably. The corresponding efficiencies were about 89% and 92% [Eq. (16)], respectively. The equivalent Gaussian percentage noise amplitudes were 3.0% and 1.5%, respectively.

Replacing postreceptoral von Kries scaling by a general linear transformation [27] increased the estimated efficiencies to about 91% and 93%, but given their dependence on extrapolation, these and the preceding larger estimates should be treated with some caution.

Whether optimized for minimum redundancy  $R_{12}^\#$  or maximum information retrieved  $I_K^\#(N)$ , the mean transformation  $T$  was almost identical. Root-mean-square differences in the entries in the matrix expressions (17) and (18) were about 0.02 and 0.01, respectively. The expressions (17) and (18) do, however, differ somewhat from sensor transformations designed to achieve optimal spectral sharpening based on Munsell reflectance spectra [16].

## 5. GENERAL DISCUSSION

The information available to the eye about the constituent elements of a scene depends both on the scene and on the spectrum of the illumination. For the 50 rural and urban scenes considered here, the mean estimated information available was about 17 bits for daylight illuminants with correlated color temperatures of 25,000 K and 4000 K, and this increased to about 20 bits with CCTs of 4000 K and 6500 K. These values of information available

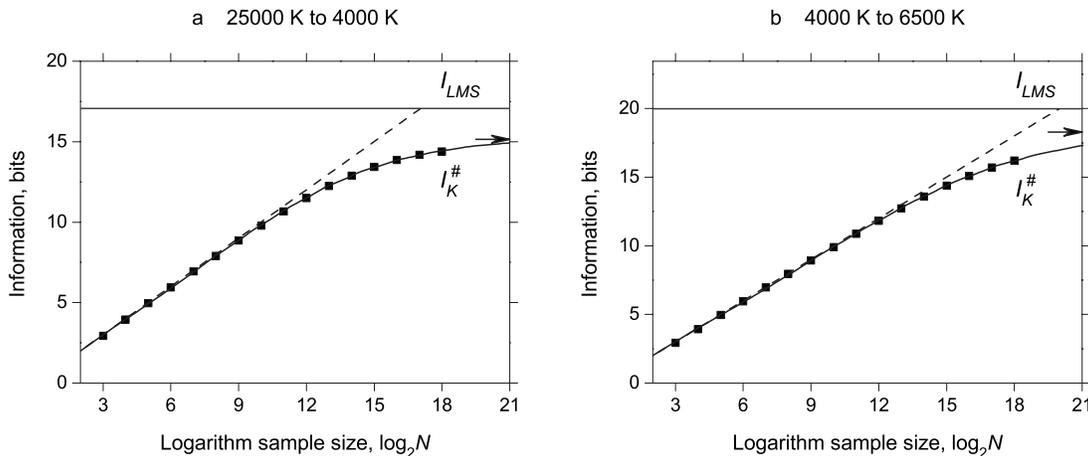


Fig. 5. Information retrieved. Details as for Fig. 4, except that the postreceptoral transformation was optimized for maximum information retrieved  $I_K^\#(N)$ . Data for information retrieved receptorally were identical to those in Fig. 4 and have been omitted.

correspond to means of about  $1.4 \times 10^5$  and  $1.0 \times 10^6$  distinct points in each scene that, in principle, are available for reliable identification.

From Fig. 3, it is clear that coding at the receptor level is highly redundant. Not only are there large interdependencies in the responses of L, M, and S cones to a scene under a particular illuminant as revealed in the redundancies  $R_1$  and  $R_2$ , there are even greater interdependencies in the way that these responses change with changes in that illuminant, as revealed by the higher-order redundancy  $R_{12}$ , despite full adaptation to the mean by von Kries scaling. Unsurprisingly, these interdependencies were reduced with an appropriate postreceptoral transformation  $T$ . When  $T$  was optimized for a minimum in the higher-order postreceptoral redundancy  $R_{12}^\#$ , the value of the latter declined to about 60% of  $R_{12}$ , and the postreceptoral redundancies  $R_1^\#$  and  $R_2^\#$  also declined, with means about 70% of their receptor counterparts. The choice of the pair of daylight illuminants had little effect on these reductions.

With this transformation  $T$ , the estimated information retrieved by postreceptoral trichromatic matching across two images of a scene yielded asymptotic means of about 14 bits for CCTs of 25,000 K and 4000 K and about 17 bits for CCTs of 4000 K and 6500 K. These values represent efficiencies of about 80% and 85%, respectively, and correspond to means of about  $1.3 \times 10^4$  and  $1.3 \times 10^5$  distinct points in each scene that can be identified reliably by trichromatic matching. The equivalent Gaussian percentage noise amplitudes were about 4.3% and 2.0%, respectively.

The limit on the information retrieved is set by the transformation  $T$ , assumed to be linear [17], and the matching procedure. In fact, only with Gaussian distributed signals is nearest-neighbor matching in the space of responses certain to be optimal [24] (Appendix C). As illustrated in Figs. 1 and 2, the distributions of both receptor responses and identification errors have tails much longer than Gaussian, falling off approximately exponentially. Even so, with  $T$  optimized directly for maximum information retrieved, rather than for minimum redundancy, the estimated efficiencies increased to about 89% and 92%, respectively, approaching ideal observer efficiency levels, although the extrapolated values were more uncertain. These asymptotes correspond to means of  $3.6 \times 10^4$  and  $3.2 \times 10^5$  distinct points per scene that can be identified reliably. By allowing the scaling of postreceptoral responses by a general linear transformation rather than by von Kries scaling, the estimated efficiencies could be increased slightly further still, but precise estimates were difficult.

This analysis was developed with reference to an ideal observer. For a real observer, in addition to the noise present at receptor and higher levels in the visual pathway, there are other uncertainties to do with attention, memory, and the nature of the experimental task. Nevertheless, there is evidence from observers' judgments of whether materials are constant across differently illuminated scenes that suggests that identification is probably as good as it can be considering the theoretical limitations of the task [44]. With some natural scenes where there is little metamerism, performance can actually approach

the ideal [45]. In neither [44] nor [45] was the spatial structure of the stimuli apparently critical.

## APPENDIX A: DIFFERENTIAL ENTROPIES OF IMAGES

Suppose that a scene under illuminants with incident spectral radiances  $e_1$  and  $e_2$  gives rise to continuous random receptor-response variables  $(L_1, M_1, S_1)$  and  $(L_2, M_2, S_2)$ , and joint random variable  $(L_1, M_1, S_1, L_2, M_2, S_2)$ , with pdfs  $f_1$ ,  $f_2$ , and  $f_{12}$ , respectively. The corresponding differential entropies  $h(L_1, M_1, S_1)$ ,  $h(L_2, M_2, S_2)$ , and  $h(L_1, M_1, S_1, L_2, M_2, S_2)$  are given [8] by

$$h(L_1, M_1, S_1) = - \int f(l_1, m_1, s_1) \log f(l_1, m_1, s_1) dl_1 dm_1 ds_1,$$

$$h(L_2, M_2, S_2) = - \int f(l_2, m_2, s_2) \log f(l_2, m_2, s_2) dl_2 dm_2 ds_2,$$

$$\begin{aligned} & h(L_1, M_1, S_1, L_2, M_2, S_2) \\ &= - \int f(l_1, m_1, s_1, l_2, m_2, s_2) \\ & \quad \times \log f(l_1, m_1, s_1, l_2, m_2, s_2) dl_1 dm_1 ds_1 dl_2 dm_2 ds_2, \end{aligned} \quad (\text{A1})$$

where the integrals are assumed to exist and are taken over the support sets of  $(L_1, M_1, S_1)$ ,  $(L_2, M_2, S_2)$ , and  $(L_1, M_1, S_1, L_2, M_2, S_2)$ .

The mutual information  $I(L_1, M_1, S_1; L_2, M_2, S_2)$  between  $(L_1, M_1, S_1)$  and  $(L_2, M_2, S_2)$  is defined [8] by

$$\begin{aligned} & I(L_1, M_1, S_1; L_2, M_2, S_2) \\ &= \int f_{12}(l_1, m_1, s_1, l_2, m_2, s_2) \\ & \quad \times \log \frac{f_{12}(l_1, m_1, s_1, l_2, m_2, s_2)}{f_1(l_1, m_1, s_1) f_2(l_2, m_2, s_2)} dl_1 dm_1 ds_1 dl_2 dm_2 ds_2. \end{aligned} \quad (\text{A2})$$

From Eq. (A1), it follows that  $I(L_1, M_1, S_1; L_2, M_2, S_2)$  may be rewritten as a difference in differential entropies; thus

$$\begin{aligned} I(L_1, M_1, S_1; L_2, M_2, S_2) &= h(L_1, M_1, S_1) + h(L_2, M_2, S_2) \\ & \quad - h(L_1, M_1, S_1, L_2, M_2, S_2). \end{aligned} \quad (\text{A3})$$

Suppose that the postreceptoral transformation  $T$  is smooth and invertible, so that  $(L_1^\#, M_1^\#, S_1^\#) = T(L_1, M_1, S_1)$  and  $(L_2^\#, M_2^\#, S_2^\#) = T(L_2, M_2, S_2)$  with corresponding pdfs  $f_1^\#$ ,  $f_2^\#$ , and  $f_{12}^\#$ . Then

$$\begin{aligned} & I(L_1^\#, M_1^\#, S_1^\#; L_2^\#, M_2^\#, S_2^\#) \\ &= \int f_{12}^\#(l_1^\#, m_1^\#, s_1^\#, l_2^\#, m_2^\#, s_2^\#) \\ & \quad \times \log \frac{f_{12}^\#(l_1^\#, m_1^\#, s_1^\#, l_2^\#, m_2^\#, s_2^\#)}{f_1^\#(l_1^\#, m_1^\#, s_1^\#) f_2^\#(l_2^\#, m_2^\#, s_2^\#)} dl_1^\# dm_1^\# ds_1^\# dl_2^\# dm_2^\# ds_2^\#. \end{aligned} \quad (\text{A4})$$

Let  $J$  be the Jacobian of  $T$ . Then  $f_1^\#(l_1^\#, m_1^\#, s_1^\#) = f_1(l_1, m_1, s_1) |J(l_1, m_1, s_1)|^{-1}$ , where  $|\cdot|$  represents the absolute value of the determinant, and analogously for  $f_2^\#(l_2^\#, m_2^\#, s_2^\#)$  and  $f_{12}^\#(l_1^\#, m_1^\#, s_1^\#, l_2^\#, m_2^\#, s_2^\#)$ . Substituting into Eq. (A4) gives [8]

$$I(L_1^\#, M_1^\#, S_1^\#; L_2^\#, M_2^\#, S_2^\#) = I(L_1, M_1, S_1; L_2, M_2, S_2).$$

That is, the information available postreceptorally is the same as that at the receptors.

## APPENDIX B: DECOMPOSITION OF INFORMATION

With the information available  $I_{LMS}$  defined by Eq. (3) and individual cone contributions  $I_L, I_M, I_S$  defined by Eq. (8), the difference  $I_{LMS} - I_L - I_M - I_S$  is given by

$$\begin{aligned} I_{LMS} - I_L - I_M - I_S &= h(L_1, M_1, S_1) + h(L_2, M_2, S_2) \\ &\quad - h(L_1, M_1, S_1, L_2, M_2, S_2) - h(L_1) \\ &\quad - h(L_2) + h(L_1, L_2) - h(M_1) - h(M_2) \\ &\quad + h(M_1, M_2) - h(S_1) - h(S_2) + h(S_1, S_2). \end{aligned}$$

After rearrangement,

$$\begin{aligned} I_{LMS} &= I_L + I_M + I_S - [h(L_1) + h(M_1) + h(S_1) - h(L_1, M_1, S_1)] \\ &\quad - [h(L_2) + h(M_2) + h(S_2) - h(L_2, M_2, S_2)] + [h(L_1, L_2) \\ &\quad + h(M_1, M_2) + h(S_1, S_2) - h(L_1, M_1, S_1, L_2, M_2, S_2)]. \end{aligned}$$

Since  $h(L_1, M_1, S_1, L_2, M_2, S_2) = h(L_1, L_2, M_1, M_2, S_1, S_2)$ , it follows from Eqs. (6) and (7) that

$$I_{LMS} = I_L + I_M + I_S - R_1 - R_2 + R_{12},$$

as in Eq. (9). The corresponding postreceptor decomposition (10) follows in the same way.

## APPENDIX C: VALIDATION OF ESTIMATORS

To show the relationship between the information available (9) and (10) and the information retrieved (12) and (14), and to verify the offset Kozachenko–Leonenko estimator [22,39] for continuous distributions and the Grassberger estimator ([42] Eqs. (23) and (27)) for discrete distributions, these estimators were applied to trichromatic Gaussian images, which have a known entropy.

The Gaussian images were generated from trivariate normal distributions of independent continuous variables  $L, M, S$ , with correlation coefficients  $\rho(L_1, L_2) = \rho(M_1, M_2) = \rho(S_1, S_2) = \rho$ . Then information available and information retrieved coincide [24] and can be calculated analytically from the value of  $\rho$  [8]; thus

$$I_{\text{true}} = -\frac{3}{2} \log(1 - \rho^2).$$

Figure 6 shows the mean estimates of the information available  $I_{LMS}$  from the offset Kozachenko–Leonenko estimator and information retrieved by trichromatic matching  $I_K(N)$  from the Grassberger estimator ([42] Eqs. (23) and (27)) as a function of the logarithm of the size  $N$  of the sample. Means were taken over 100 samples each of size  $N$ , with  $N$  ranging from  $2^3$  to  $2^{18}$ . The value of the corre-

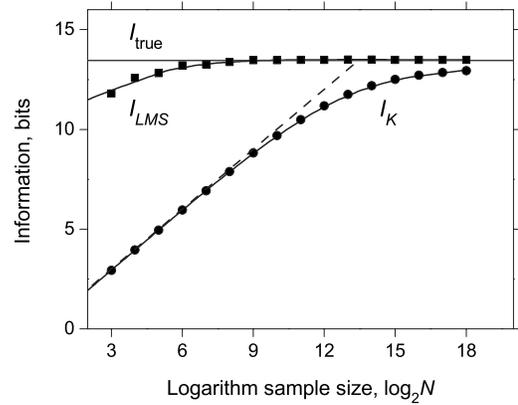


Fig. 6. Information estimates for Gaussian trichromatic images. The mean estimate of the information available  $I_{LMS}$  from the offset Kozachenko–Leonenko estimator [22,39] (squares) and the mean estimate of the information retrieved  $I_K(N)$  by trichromatic matching with a sample of size  $N$  from the Grassberger estimator ([42], Eqs. (23) and (27)) (circles) are shown as a function of the logarithm of the size  $N$  of the sample. The smooth curves are loess fits [43]. The bounding value of the discrete entropy of a sample of size  $N$  is shown by the dashed line and the true value  $I_{\text{true}}$  of the information available for Gaussian images by the horizontal line. Means were taken over 100 samples each of size  $N$ . Sample SDs at  $\log_2 N = 18$  were 0.008 bits for  $I_{LMS}$  and 0.002 bits for  $I_K(N)$ .

lation coefficient  $\rho$  was fixed at 0.999, chosen because it corresponds to a value of  $I_{\text{true}} = 13.5$  bits, which is similar to values estimated for information retrieved for natural-scene data (see Subsection 4.B). Other values are considered in [41].

It is clear that the offset Kozachenko–Leonenko estimator of  $I_{LMS}$  gives accurate estimates of the information available  $I_{\text{true}}$  with relatively small sample sizes (e.g.,  $N = 2^8$ ), much smaller than the maximum of  $N = 1344 \times 1024 = 2^{20.39}$  available.

By contrast, the original Kozachenko–Leonenko estimator (not shown here) underestimated the information available and provided estimates that converged only with the largest sample size tested ( $N = 2^{18}$ ).

The information retrieved  $I_K(N)$  from each discrete sample  $N$  is necessarily bounded by the discrete entropy of the sample, i.e.,  $\log N$  (Fig. 6, dashed line). As explained in Subsection 2.F,  $I_K(N)$  cannot not grow indefinitely with  $N$ . When  $N$  is small, the risk of error is small, so that  $I_K(N)$  is close to  $\log N$ ; but as  $N$  increases, the risk of error increases, with the result that  $I_K(N)$  increases less and less rapidly. For Gaussian images and trichromatic matching, the limit of  $I_K(N)$  should equal the information available. For  $N = 2^{18}$ , the Grassberger estimate of  $I_K(N)$  underestimated  $I_{\text{true}}$  by 4%. For other values of  $\rho$ , the underestimation of information retrieved may be larger or smaller than 4%, but overestimation by the Grassberger estimate of  $I_K(N)$  never occurred, although it cannot always be ruled out.

## ACKNOWLEDGMENTS

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), UK (grant nos. EP/B000257/1 and EP/E056512/1), by the Fundação para

a Ciência e a Tecnologia (grant POSC/EEA-SRI/57554/2004), and by the Centro de Física da Universidade do Minho.

## REFERENCES

- G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed. (Wiley, 1982).
- D. H. Foster, S. M. C. Nascimento, and K. Amano, "Information limits on neural identification of colored surfaces in natural scenes," *Visual Neurosci.* **21**, 331–336 (2004).
- D. H. Foster, S. M. C. Nascimento, and K. Amano, "Information limits on identification of natural surfaces by apparent colour," *Perception* **34**, 1003–1008 (2005).
- W. S. Geisler, "Visual perception and the statistical properties of natural scenes," *Annu. Rev. Psychol.* **59**, 167–192 (2008).
- N. J. Dominy, J.-C. Svenning, and W.-H. Li, "Historical contingency in the evolution of primate color vision," *J. Human Evolution* **44**, 25–45 (2003).
- C. E. Shannon, "A mathematical theory of communication," [part 1] *Bell Syst. Tech. J.* **27**, 379–423 (1948).
- C. E. Shannon, "A mathematical theory of communication," [part 2] *Bell Syst. Tech. J.* **27**, 623–656 (1948).
- T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, 1991).
- G. Buchsbaum and A. Gottschalk, "Trichromacy, opponent colours coding and optimum colour information transmission in the retina," *Proc. R. Soc. London, Ser. B* **220**, 89–113 (1983).
- J. A. Swets, W. P. Tanner, Jr., and T. G. Birdsall, "Decision processes in perception," *Psychol. Rev.* **68**, 301–340 (1961).
- N. A. Macmillan and C. D. Creelman, *Detection Theory: A User's Guide*, 2nd ed. (Erlbaum, 2005).
- D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes," *J. Opt. Soc. Am. A* **23**, 2359–2372 (2006).
- M. Vorobyev and D. Osorio, "Receptor noise as a determinant of colour thresholds," *Proc. R. Soc. London, Ser. B* **265**, 351–358 (1998).
- D. I. A. MacLeod and T. von der Twier, "The pleistochrome: optimal opponent codes for natural colours," in *Colour Perception: Mind and the Physical World*, R. Mausfeld and D. Heyer, eds. (Oxford Univ. Press, 2003), pp. 155–184.
- D. H. Foster and R. S. Snelgar, "Test and field spectral sensitivities of colour mechanisms obtained on small white backgrounds: action of unitary opponent-colour processes?" *Vision Res.* **23**, 787–797 (1983).
- G. D. Finlayson, M. S. Drew, and B. V. Funt, "Spectral sharpening: sensor transformations for improved color constancy," *J. Opt. Soc. Am. A* **11**, 1553–1563 (1994).
- R. L. De Valois, N. P. Cottaris, S. D. Elfar, L. E. Mahon, and J. A. Wilson, "Some transformations of color information from lateral geniculate nucleus to striate cortex," *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4997–5002 (2000).
- J. von Kries, "Theoretische Studien über die Umstimmung des Sehorgans," in *Festschrift der Albrecht-Ludwigs-Universität* (Freiburg, 1902), pp. 145–158.
- J. von Kries, "Die Gesichtsempfindungen," in *Handbuch der Physiologie des Menschen*, W. Nagel, ed. (Vieweg, 1905), pp. 109–282.
- M. A. Webster and J. D. Mollon, "Adaptation and the color statistics of natural images," *Vision Res.* **37**, 3283–3298 (1997).
- M. Studený and J. Vejnárová, "The multi-information function as a tool for measuring stochastic dependence," in *Learning in Graphical Models*, M. I. Jordan, ed. (MIT Press, 1999), pp. 261–298.
- I. Marín-Franch and D. H. Foster, "Distribution of information within and across colour spaces," in *CGIV 2008* (Society for Imaging Science and Technology, 2008), pp. 303–306.
- K. Amano and D. H. Foster, "Colour constancy under simultaneous changes in surface position and illuminant," *Proc. R. Soc. London, Ser. B* **271**, 2319–2326 (2004).
- A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Trans. Inf. Theory* **42**, 1520–1529 (1996).
- H. Barlow, "Redundancy reduction revisited," *Network Comput. Neural Syst.* **12**, 241–253 (2001).
- G. D. Finlayson, M. S. Drew, and B. V. Funt, "Color constancy: generalized diagonal transforms suffice," *J. Opt. Soc. Am. A* **11**, 3011–3019 (1994).
- B. Funt and H. Jiang, "Non-von-Kries 3-parameter color prediction," *Proc. SPIE* **5007**, 182–189 (2003).
- D. H. Foster and K. Żychaluk, "Is there a better non-parametric alternative to von Kries scaling?" in *CGIV 2008* (Society for Imaging Science and Technology, 2008), pp. 41–44.
- S. M. C. Nascimento, F. P. Ferreira, and D. H. Foster, "Statistics of spatial cone-excitation ratios in natural scenes," *J. Opt. Soc. Am. A* **19**, 1484–1490 (2002).
- CIE, "Colorimetry, 3rd Edition," CIE Publication 15:2004 (CIE Central Bureau, Vienna, 2004).
- D. B. Judd, D. L. MacAdam, and G. Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature," *J. Opt. Soc. Am.* **54**, 1031–1040 (1964).
- J. Hernández-Andrés, J. Romero, and J. L. Nieves, "Color and spectral analysis of daylight in southern Europe," *J. Opt. Soc. Am. A* **18**, 1325–1335 (2001).
- A. Stockman, L. T. Sharpe, and C. Fach, "The spectral sensitivity of the human short-wavelength sensitive cones derived from thresholds and color matches," *Vision Res.* **39**, 2901–2927 (1999).
- A. Stockman and L. T. Sharpe, "The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype," *Vision Res.* **40**, 1711–1737 (2000).
- B. W. Silverman, *Density Estimation for Statistics and Data Analysis* (Chapman & Hall, London, 1986).
- A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E* **69**, 066138 (2004).
- R. Steuer, J. Kurths, C. O. Daub, J. Weise, and J. Selbig, "The mutual information: Detecting and evaluating dependencies between variables," *Bioinformatics* **18**, S231–S240 (2002).
- J. D. Victor, "Binless strategies for estimation of information from neural data," *Phys. Rev. E* **66**, 051903 (2002).
- L. F. Kozachenko and N. N. Leonenko, "Sample estimate of the entropy of a random vector," *Probl. Inf. Transm.* **23**, 95–101 (1987); Russian translation, *Probl. Peredachi Inf.* **23**, 9–16 (1987).
- M. N. Goria, N. N. Leonenko, V. V. Mergel, and P. L. Novi Inverardi, "A new class of random vector entropy estimators and its applications in testing statistical hypotheses," *J. Nonparametric Statistics* **17**, 277–297 (2005).
- I. Marín-Franch, "Information-theoretic analysis of trichromatic images of natural scenes under different phases of daylight," Ph.D. thesis (The University of Manchester, Manchester, 2009).
- P. Grassberger, "Entropy estimates from insufficient samplings," (2008), retrieved <http://arxiv.org/abs/physics/0307138v2>.
- W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *J. Am. Stat. Assoc.* **74**, 829–836 (1979).
- J. J. M. Granzier, E. Brenner, and J. B. J. Smeets, "Reliable identification by color under natural conditions," *J. Vision* **9**, 1–8 (2009).
- D. H. Foster, K. Amano, and S. M. C. Nascimento, "Color constancy in natural scenes explained by global image statistics," *Visual Neurosci.* **23**, 341–349 (2006).