Using colour as a label in natural scenes

David H. Foster,[1,*] Iván Marín-Franch, [1] Kinjiro Amano,[1] and Sérgio M. C. Nascimento[2]

[1]School of Electrical and Electronic Engineering, University of Manchester, Manchester M60 1QD, UK
[2]Department of Physics, Gualtar Campus, University of Minho, 4710-057 Braga, Portugal

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

Trichromatic sampling by the eye and variations in illumination both limit the effectiveness of colour as a label for identifying the elements of a scene. The aim of this work was to quantify these physical limits for an ideal observer who uses colour signals to represent each point in an image. Simulations were based on 50 hyperspectral images of natural scenes under various daylight illuminants with correlated colour temperatures 4000 K, 6500 K, and 25000 K. Estimates were made of the information, in the sense of Shannon, available from each scene under different illuminants and the information retrieved with optimum, linear, cone-opponent processing. For the largest illuminant difference, between 25000 K and 4000 K, the information available was 17 bits, equivalent to an average of 1.4×10^5 distinct identifiable points per scene. The maximum information retrievable with optimum cone-opponent coding was ca. 15.5 bits, equivalent to an average of 4.5×10^4 distinct identifiable points per scene, although estimates were obtained by extrapolation. Even without taking into account receptoral and postreceptoral noise, these estimates were much smaller than the estimated average number of discriminable colours, ca. 2.7×10^5, available in single images of natural scenes (Linhares, et al., J. Opt. Soc. Am. A, 2008, 25, 2918-2924). This difference suggests a possible physical basis for the observation that colour naming is coarser than colour discrimination.