



## OPEN Efficient quantization of painting images by relevant colors

Zeinab Tirandaz<sup>1✉</sup>, David H. Foster<sup>1✉</sup>, Javier Romero<sup>2</sup> & Juan Luis Nieves<sup>2</sup>

Realistic images often contain complex variations in color, which can make economical descriptions difficult. Yet human observers can readily reduce the number of colors in paintings to a small proportion they judge as relevant. These relevant colors provide a way to simplify images by effectively quantizing them. The aim here was to estimate the information captured by this process and to compare it with algorithmic estimates of the maximum information possible by colorimetric and general optimization methods. The images tested were of 20 conventionally representational paintings. Information was quantified by Shannon's mutual information. It was found that the estimated mutual information in observers' choices reached about 90% of the algorithmic maxima. For comparison, JPEG compression delivered somewhat less. Observers seem to be efficient at effectively quantizing colored images, an ability that may have applications in the real world.

Images of the real world and of the objects within it reveal the spatial and spectral complexity of natural surfaces<sup>1</sup> and their illumination<sup>2,3</sup>. But the impression of chromatic detail is likely to be founded on partial information, limited by our peripheral color awareness<sup>4</sup>, pattern of gaze<sup>5</sup>, attention<sup>6</sup>, memory<sup>7</sup>, and the emotional and semantic content of the scene<sup>8,9</sup>, and, more fundamentally, by the relative abundances or frequencies of the different colors present<sup>10,11</sup>. Of course, not all the colors in a scene are needed in order to describe or remember it, and different methods have been used to estimate which colors are relevant.

The most direct approach is to ask human observers to make the required judgments. In a psychophysical experiment<sup>12</sup>, observers selected those pixels in an image of a painting they considered to belong to a "relevant chromatic area". The images were of 20 paintings in the Prado Museum<sup>13</sup>, Madrid, and of 20 artworks from the database of Khan et al.<sup>14</sup>. Observers could choose as many colors as they wished. This procedure yielded a mean of 21 relevant colors for each image, with the number and identity of the colors varying with both the image and the observer.

An alternative theoretical approach is to use colorimetric methods<sup>15</sup>. In such an analysis<sup>16</sup>, the approximately uniform color space CIELAB<sup>17</sup> was divided into cubic cells whose side length was a multiple of the smallest discriminable step, and colorimetric arguments were then used to assign the colors of a scene to those cells. The images were of 4266 artworks from the database of Khan et al.<sup>14</sup>. The analysis yielded a mean of 18 relevant colors per image, the number and identity again varying with the image. Further details of this colorimetric method and of observers' experimental judgments are given in Methods.

Whatever the method, the use of relevant colors offers a way to simplify an image by quantizing it, that is, by reducing a large, essentially continuous range of colors to a much smaller discrete set. It is unclear, however, whether this process is efficient. Does it capture the most information about the image for a given number of relevant colors?

The present analysis addresses this question. Information was quantified by Shannon's mutual information<sup>18</sup>, though other formulations are possible<sup>19</sup>. Estimates were made of the mutual information between images and the quantized representations implied from observers' relevant colors and from the colorimetric method and then compared with estimates of the maximum mutual information by general optimization methods. These methods used clustering to partition the set of colors in each image into  $n$  subsets according to some cost or objective function. Although values of  $n$  were matched to the number of relevant colors for each image obtained by each observer and by the colorimetric method, methods for objectively determining  $n$  were also considered. In all, five clustering algorithms were considered, and were based on  $k$ -means<sup>20</sup>, maximum entropy clustering (MEC)<sup>21,22</sup>, expectation maximization with a Gaussian mixture model (GMM)<sup>21,23</sup>, minimum conditional entropy (minCEntropy)<sup>24–26</sup>, and Graph-Cut<sup>27–30</sup>. Each is explained in Methods, but minCEntropy has the notional advantage that it is designed to maximize the mutual information between the data and the clustering.

<sup>1</sup>Department of Electrical and Electronic Engineering, University of Manchester, Manchester M13 9PL, UK. <sup>2</sup>Department of Optics, University of Granada, 18071 Granada, Spain. ✉email: zeinab.tirandaz@manchester.ac.uk; d.h.foster@manchester.ac.uk

The test images are illustrated in Fig. 1. They were chosen for their complex spatio-chromatic content and realistic character and are a subset of conventionally representational painting images used earlier<sup>16</sup>, for which observer estimates of relevant colors were already available<sup>12</sup>. The color gamuts of each of these images are illustrated in the Methods section. An example of a quantized representation is shown later. Other examples are available elsewhere<sup>12,16</sup>.

For clarity, this analysis should be distinguished from more abstract approaches<sup>31</sup> in which the efficiency of color naming itself is evaluated with uniform color palettes such as the Munsell set<sup>32</sup>. It should also be distinguished from color categorization with a fixed set of basic or salient color terms<sup>33</sup>, and from image classification with a set of color descriptors for database retrieval<sup>34</sup>. The relevant colors used here were distributed non-uniformly in each image, they were not drawn from fixed categories, and observers were instructed not to name them.

It was found that the estimated mutual information in observer judgments of relevant colors was close to the estimates from all five clustering methods, including minCEntropy, and the colorimetric method. Observers seem to be efficient at effectively quantizing colored images of paintings.

## Results

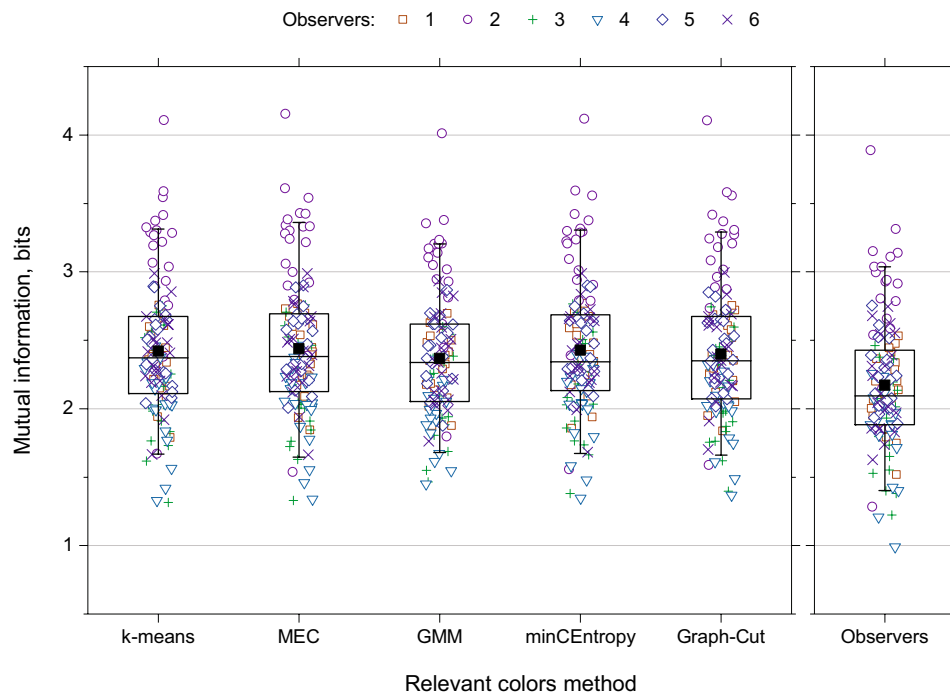
**Efficiencies of representations by relevant colors.** Figure 2 shows the estimated mutual information between each image in Fig. 1 and its quantized representation by relevant colors for each of the five clustering methods, *k*-means++, MEC, GMM, minCEntropy, and Graph-Cut, and by individual observers. The 120 data points for each method derive from the 20 images and the six observer's choices of numbers of relevant colors for each image<sup>12</sup>. The data are summarized by the superimposed boxplots.

The five clustering methods produced closely similar mean levels of estimated mutual information of about 2.4 bits, and observers slightly lower levels, about 2.2 bits. The variance in the plotted data, however, is potentially misleading in that the estimated mutual information for an image and the number of relevant colors *n* covary across methods and observers. This confound was circumvented by calculating the efficiency of one method relative to another for each image and for the same number of relevant colors *n*. Values of *n* were drawn from observers' choices, except for the colorimetric method, which determined *n* automatically for each image. That is, for *k*-means++, MEC, GMM, minCEntropy, and Graph-Cut, multiple values of *n* were available from observers for each image whereas, for the colorimetric method, only one value was available for each image.

The choice of reference method is not critical, and the minCEntropy method was chosen for the property mentioned earlier, that is, maximizing the mutual information between the data and the clustering. Thus, suppose



**Figure 1.** Montage of images of conventionally representational paintings in the Prado Museum<sup>13</sup>, adapted from Ref.<sup>12</sup>, Fig. 1, licensed under CC BY 4.0. The color gamuts of each of these images are illustrated in Fig. 4 in Methods.



**Figure 2.** Information from relevant colors. Estimated mutual information between each of the 20 images in Fig. 1 and their quantized representations by relevant colors is shown for five clustering methods and six human observers, indicated by different symbols. Horizontal jitter has been added to reduce overlap between symbols<sup>35</sup>. Estimates for individual images are not distinguished. Boxplots represent median (center line), mean (solid square), interquartile range (rectangle), and the 5th and 95th percentiles (whiskers).

that  $I$  is the mutual information between an image and its quantized representation by relevant colors estimated by a particular method and that  $I_{\text{ref}}$  is the corresponding quantity estimated by the minCENTropy method. Then the efficiency  $\eta$  of the method for this image and value of  $n$  is

$$\eta = \frac{I}{I_{\text{ref}}},$$

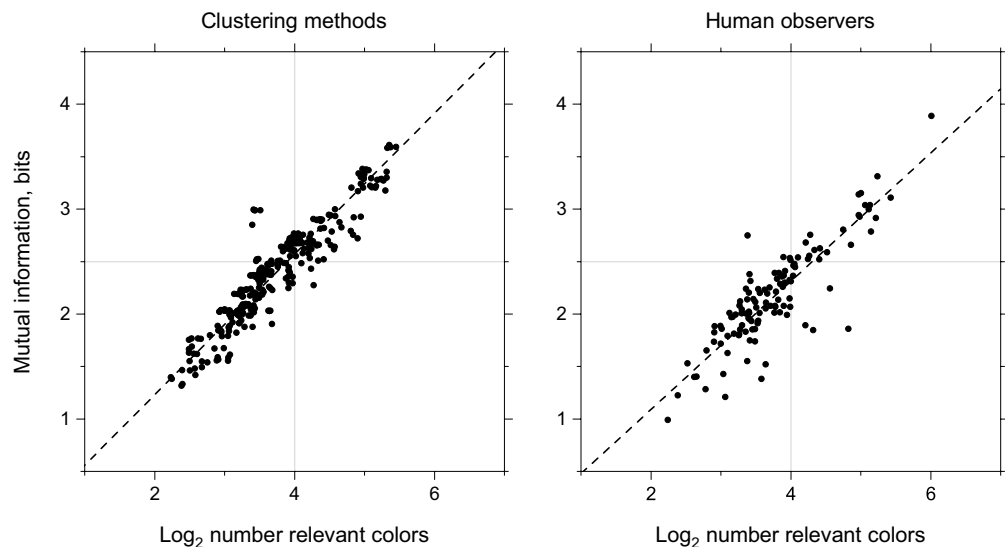
which can be averaged over all images and all  $n$ . Table 1 shows mean efficiencies  $\eta$  calculated with respect to the minCENTropy estimates, along with 95% confidence limits.

Estimates by observers are close to but less than the optimum method, with a mean efficiency of 89%. In the light of the 95% confidence limits, this difference is reliable. Estimates with the other clustering methods did not differ reliably from the minCENTropy estimate.

**Variation in numbers of relevant colors.** Observers were free to choose as many relevant colors as they wished<sup>12</sup>, and, as indicated earlier, they chose different numbers  $n$  with each scene. This is evident in Fig. 2 and was confirmed statistically (Kruskall-Wallis test,  $p < 0.001$ ). Crucially, though, observers' efficiencies  $\eta$  in making those choices did not differ significantly (Kruskall-Wallis test,  $p = 0.8$ ).

Estimation method	Mean efficiency $\eta$ (%)
k-means++	99.7 (99.3, 100.0)
MEC	100.3 (100.0, 100.7)
GMM	97.8 (96.9, 98.8)
Graph-cut	98.8 (98.5, 99.2)
Colorimetric	81.4 (78.0, 83.6)
Human observers	89.4 (88.1, 90.6)

**Table 1.** Efficiencies of methods for estimating information from relevant colors relative to minCENTropy clustering. Mean efficiencies  $\eta$  are given for three clustering methods, the colorimetric method, and six human observers. Values of  $\eta$  were averaged over 20 images and numbers of relevant colors  $n$ . Estimated 95% BCa confidence limits<sup>36</sup> are shown in parentheses, based on 1000 bootstrap replications.



**Figure 3.** Information and number of relevant colors. Estimated mutual information is plotted against logarithm to the base 2 of the number of relevant colors  $n$  for the five clustering methods and six human observers. Horizontal jitter has been added to reduce overlap between symbols<sup>35</sup>. The dashed lines represent linear regressions on theunjittered data.

With different numbers of relevant colors for each image, does mutual information increase predictably as  $n$  increases?

Figure 3 shows estimated mutual information plotted against the logarithm to the base 2 of the number of relevant colors for the five clustering methods and six human observers. Both the clustering methods and observers behaved broadly appropriately, that is, as the number of relevant colors increased, so did the estimated mutual information. The slopes of the regression fits were similar (95% confidence limits in parentheses): 0.67 (0.65, 0.69) and 0.61 (0.53, 0.66) for the clustering methods and observers, respectively.

**Objective estimates of numbers of relevant colors.** Rather than depend on observer estimates of the number of relevant colors  $n$ , is there an objective way of deciding? To this end, three algorithmic methods for estimating  $n$  for each image were tested. These were the Caliński-Harabasz<sup>37</sup> method with upper limits of  $n=40$  and  $n=80$ ; the Davies-Bouldin<sup>38</sup> method with upper limit  $n=80$ ; and an adaptive  $k$ -means method<sup>39</sup>, which does not require initialization. All are described more fully in Methods.

Table 2 shows the optimum numbers of clusters averaged over the 20 images, along with corresponding averages for observers.
















The problems of consistency are evident. The three methods deliver mean optimum numbers that differ reliably from each other, and which fall reliably below those selected by human observers.

**Comparison with JPEG coding.** The quantized representation of an image by relevant colors can be thought of as a lossy compression. As a demonstration, mutual information estimates by the minCEntropy and colorimetric methods and from observers were compared with those from the popular JPEG compression algorithm<sup>40</sup>, although accurate color quantization is not its primary goal. The JPEG implementation in MATLAB (version 9.10.0.1602886 (R2021a), The MathWorks, Inc., Natick, MA) was used with a quality setting of zero to minimize the normally large number of colors in the representation (e.g., for the image in the second row, third column of Fig. 1, a quality setting of 5 gave 6040 colors and a setting of zero gave 2611 colors).

Table 3 shows estimates of mutual information and numbers of colors  $n$  with three of the images of Fig. 1 that produced the largest visual differences across methods. Unlike Table 1 where each method was compared with the minCEntropy method for the same  $n$ , this normalization was not possible here. Since the colorimetric

Caliński-Harabasz, upper limit 40	Caliński-Harabasz, upper limit 80	Davies-Bouldin, upper limit 80	Adaptive $k$ -means	Human observers
8.2 (6.2, 11.4)	8.9 (6.5, 12.6)	2.5 (2.2, 3.1)	4.0 (3.6, 4.4)	15.4 (14.0, 17.1)

**Table 2.** Estimated optimum numbers of clusters, averaged over 20 images, according to three methods, and different initial values. Data for human observers are included for comparison. Estimated 95% BCa confidence limits<sup>36</sup> are shown in parentheses, based on 1000 bootstrap replications.

Original image	minCEntropy	Colorimetric method	Human observers	JPEG
				
	2.82 bits, $n = 20$	2.23 bits, $n = 20$	2.61 bits, $n = 21$	1.76 bits, $n = 2611$
				
	3.01 bits, $n = 23$	2.53 bits, $n = 23$	3.00 bits, $n = 35$	2.04 bits, $n = 3267$
				
	2.59 bits, $n = 19$	2.03 bits, $n = 19$	2.08 bits, $n = 14$	1.63 bits, $n = 2492$

**Table 3.** Quantized representations of an image by relevant colors and JPEG compression. The original image from Fig. 1 is shown in the first column. Subsequent columns show the image quantized by relevant colors estimated by the minCEntropy method, the colorimetric method, a human observer, and JPEG with the fewest colors. The estimated mutual information in bits and number of relevant colors  $n$  are shown beneath the corresponding image.

Estimation method	Mean CIECAM02-UCS color difference	Mean S-CIELAB color difference
minCEntropy	3.7 (3.5, 4.0)	4.1 (3.9, 4.4)
Colorimetric method	7.9 (7.7, 8.2)	9.0 (8.9, 9.3)
Human observers	5.5 (5.1, 6.0)	6.2 (5.7, 6.6)
JPEG	9.1 (8.6, 9.5)	12.7 (12.0, 13.3)

**Table 4.** Color rendering by relevant colors in two color spaces<sup>17,41,42</sup>. Color differences between the original image and its quantized representation are shown for minCEntropy clustering, the colorimetric method, six human observers, and JPEG, with values averaged over the 20 images. Estimated 95% BCa confidence limits<sup>36</sup> are shown in parentheses based on 1000 bootstrap replications.

method determined  $n$  automatically for each image, this value was used for the minCEntropy method. The observer representation was then chosen as the one with the value of  $n$  closest to that of the colorimetric method.

Even with a JPEG quality setting of zero, the mean number of JPEG unique colors was  $3.9 \times 10^3$ , more than two orders of magnitude larger than with the other three methods. Yet the minCEntropy method, the colorimetric method, and individual observers were able to preserve 2.6 bits of information on average over all 20 images, more than the 2.1 bits preserved on average with JPEG. For the calculation of the efficiency of JPEG relative to the minCEntropy method, it was impracticable to use the same  $n$  with minCEntropy, and the value of  $n$  from the colorimetric method was used instead. The mean JPEG efficiency (with 95% confidence limits) was 72.3% (68.6%, 76.1%).

The effects of information loss on appearance are obvious. For all three images in Table 3, JPEG rendered light brown as light green, most noticeably the shirt of the figure in the top row on the left, presumably a side-effect of downscaling color in JPEG. The colorimetric method showed a similar but smaller bias. Both the minCEntropy method and individual observers rendered all the colors well.

**Color rendering.** To extend the comparisons of color rendering across all the images, mean color differences were estimated between each painting image and its quantized representation. The RGB values at each pixel were converted to CIE 1931 XYZ tristimulus value values ( $2^\circ$  observer) and color differences evaluated in the approximately uniform color space CIECAM02-UCS<sup>17</sup> with respect to a 6500 K illuminant. These differences were then averaged over the whole image. For comparison, color differences were also evaluated in the somewhat less uniform color space S-CIELAB<sup>11,42</sup>, which takes into account the spatial-frequency filtering of the whole image by the eye. The same illuminant was assumed. Table 4 shows these color differences averaged over all 20 images for both color spaces. As with Table 3,  $n$  was set by the colorimetric method.

In principle, a just perceptible color difference of about 0.5 in CIECAM02-UCS is around half that value in CIELAB space<sup>43</sup>, but larger values of about 1.5 in CIECAM02-UCS and 2.2 in CIELAB space have also been used with whole images of natural scenes<sup>44</sup>. With any of these thresholds, the color differences in Table 4 manifestly represent detectable effects in CIECAM02-UCS and CIELAB space.

That there should be detectable differences is not unreasonable. Quantization by relevant colors entails a reduction in the number of unique image colors of about 5000 to 1, on average. While many of the original colors coincided with their quantized values, there were inevitably many other colors that did not.

## Discussion

Representing images in terms of a limited number of relevant colors reduces their spatial and spectral complexity, but with an inevitable loss of information. The present analysis showed that with representational painting images, the estimated information captured by human observers in their choices of relevant colors<sup>12</sup> reached about 90% of that possible by algorithmic clustering methods, all of which maximized the estimated mutual information between the image and the clustering. Observers seem to be efficient at effectively quantizing these images.

In making their choices, observers were not limited to any particular number of relevant colors<sup>12</sup>. Reassuringly, whether observers settled on a few or many relevant colors (numbers ranged from 5 to 68 across scenes), the colors they chose approached the optimum in each case. As Fig. 3 shows, mutual information estimates increased approximately linearly with the logarithm of the number of colors, though not quite as rapidly as with algorithmic clustering methods. With this variation, it seems likely that establishing an objective estimate of an optimum number of relevant colors for a given image will require explicit observer modeling, along with additional constraints. As was clear from Table 2 algorithmic clustering methods produced estimates of optimum numbers that were all much lower than for observers.

The focus of this analysis has been on information processing, not on color rendering per se. Even so, there is a parallel in the ordering of the outcomes in the two approaches for a given number of colors. Thus, in Table 1, the mutual information estimates between original and quantized images progressively decreased across clustering methods, observers, and the colorimetric method. Conversely, in Table 4, the mean color differences between original and quantized images progressively increased across the same range. But the two kinds of measure are not equivalent. The key distinction is that mutual information depends on the relative frequencies of image colors whereas color differences depend on the metrical properties of the space used to represent colors. Mutual information is not necessarily maximized by minimizing color differences<sup>45,46</sup>.

There are several caveats to this analysis. First, in practice, the relevant colors characterizing an image need not be the same as the set of colors significant for a particular task. Some colors may additionally have a subjective salience<sup>47–49</sup>, affecting measures such as gaze direction<sup>5</sup> and search performance<sup>50,51</sup>. Second, although the five clustering methods invoked different criteria to achieve a solution, the possibility remains that other methods could yield higher estimates. Third, the comparison with quantization by JPEG was informative but intended only for illustration, since JPEG is designed for overall image compression, not efficient color quantization. A range of alternative color spaces has been considered for JPEG<sup>52</sup>, though an adaptable color coding using relevant colors may offer still better performance<sup>53,54</sup>. Fourth and last, all the information estimates were based on images of conventionally representational paintings rather than on images of actual scenes, objects, or individuals. That said, paintings and real scenes are known to have overlapping gamuts, elongated in the yellow–blue direction, albeit with the gamuts of paintings tilted in the red direction<sup>55–57</sup>.

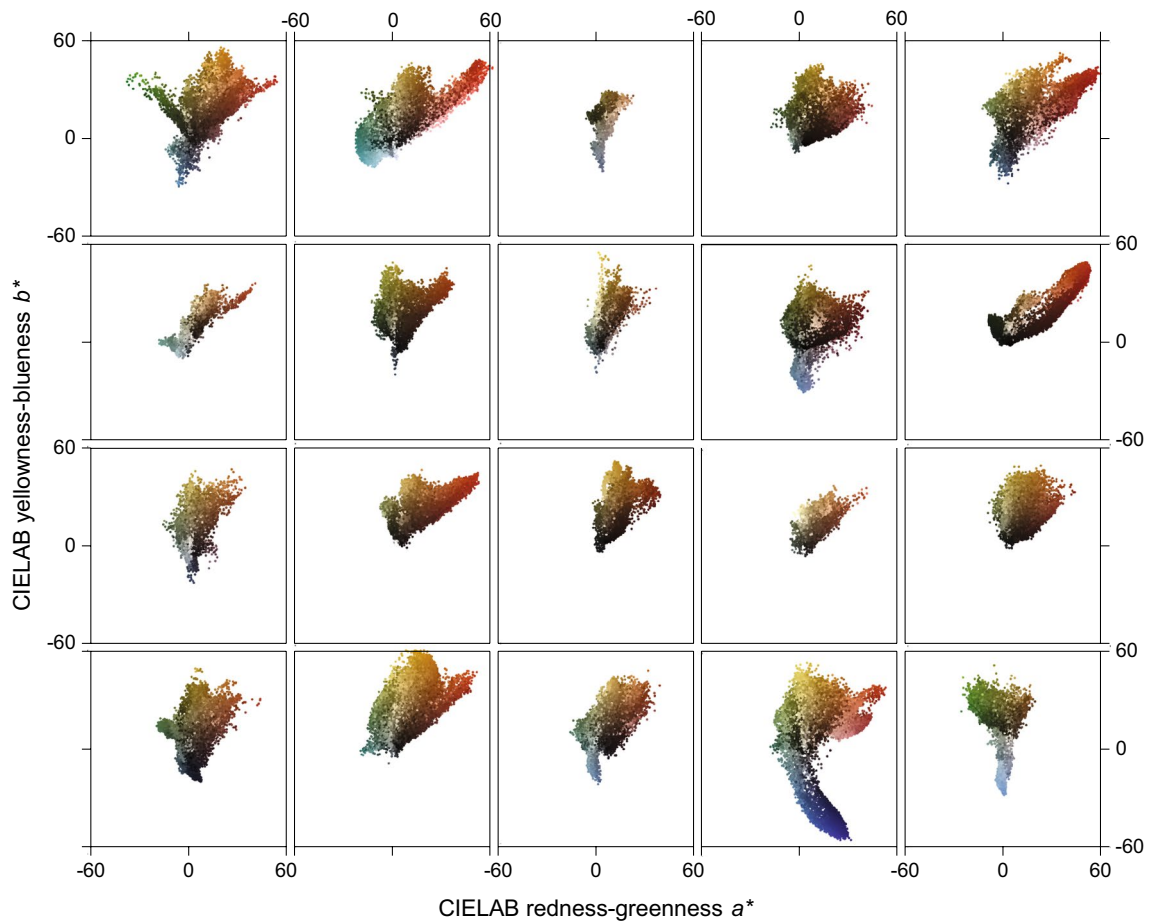
Given the levels of performance estimated here, if observers are equally efficient in judging relevant colors in the real world, the resulting quantized representations may find applications in a variety of tasks, including the description and memorization of scenes, and their eventual recall.

## Methods

**Image data.** The 20 test images of conventionally representational paintings from the Trecento to the Romantic era were the same as those used in a psychophysical study<sup>12</sup> to obtain observer estimates of relevant colors. They represent rural landscapes, indoor scenes, still lifes, portraits, and historical events. The images, made available by the Prado Museum<sup>13</sup>, were coded as 24-bit RGB with 6,520,320 to 9,682,560 pixels per image. The mean number of unique colors in each image was  $1.3 \times 10^5$  within the bit resolution of the dataset.

Figure 4 illustrates the gamuts of the colors in each of the painting images in Fig. 1. Pixels were drawn uniformly from each image and plotted, for consistency with earlier data, in the chromaticity plane of the approximately uniform color space CIELAB<sup>17</sup>, where  $a^*$  correlates with redness–greenness and  $b^*$  with yellowness–blueness. Notice the small gamut for the painting in the top row, column 3 of Fig. 1 and the extended blue lobe for the painting in the bottom row, column 4 of Fig. 1. The locations of relevant colors with these image gamuts are shown in Ref.<sup>12</sup>.

**Observer data.** Data on relevant colors chosen by human observers were taken from a previously reported psychophysical experiment<sup>12</sup> with six participants, three men and three women. The number of observers was of the same order as in similar works<sup>58–60</sup> and the homogeneity of their responses is considered in the Results section. The observers sat in front of a monitor with head position stabilized by a chinrest. Images of paintings were displayed with PsychToolbox-3<sup>61</sup>, a package consisting of MATLAB and GNU Octave functions and Python toolkits that provided an interactive visual environment. Observers used a mouse to select within the image representative locations they considered “valid instances of relevant colors”. The recorded relevant colors were defined by the average of 25 pixels around each of the selected locations. The quantized images were created after the experiment and so did not influence observers’ choices. The study was conducted according to the



**Figure 4.** Color gamuts of painting images in Fig. 1. Each plot shows the CIELAB<sup>17</sup> chromaticity coordinates ( $a^*$ ,  $b^*$ ) of pixels drawn uniformly from the corresponding image in Fig. 1, which has the same row-column layout. Only about 1% of pixels are plotted to more clearly reveal density variations.

guidelines of the Declaration of Helsinki and approved by the Institutional Review Board (Ethics Committee) of the University of Granada, Spain (protocol code 1746/CEIH/2020). Informed consent was obtained from all participants in that study. Additional detail about the choice of observers is available in Ref.<sup>12</sup>.

**Algorithmic clustering methods.** The colorimetric method used here<sup>16</sup> mapped RGB pixel values of the original image into CIELAB color space<sup>17</sup>, which was then divided into small cubic cells with fixed sides. Based on the number of pixels in each cube and whether their luminance and chroma values were within a certain range, each cube was designated as relevant or not. Pixels with values in the same relevant cell were assigned the average for that cell, and pixels in non-relevant cells were assigned the closest relevant cell average.

The MEC method<sup>21,22</sup> incorporates Shannon entropy<sup>18</sup> in the objective function. Initial clusters were obtained by  $k$ -means++, and the entropy-based objective function was then maximized, equivalent to minimizing negative entropy. Clusters were updated iteratively until the objective function value converged.

With the GMM method<sup>21,23</sup>, pixel values are assumed to be generated by  $k$  Gaussian mixture components. The built-in MATLAB  $k$ -means++ clustering function `kmeans` was used to obtain an initial clustering with the mean, covariance, and membership weights forming the objective function. Iterative expectation maximization was applied until there was no significant change in the objective function value.

The MinEntropy method<sup>24–26</sup> is a hill-climbing method. It maximized the mutual information between the original pixel values and their quantized representation by clustering. Initialization was again by  $k$ -means++ and clustering was updated iteratively until the clusters stabilized.

The Graph-Cut method<sup>27–30</sup> is an energy-minimization method based on graph cuts. Initialization was by  $k$ -means. The graph was then constructed with pixels as its nodes and the probability of pixels belonging to the same cluster as its weighted edges. The algorithm then cut through the weak edges to yield a final clustering.

**Algorithmic estimates of cluster numbers.** Three algorithms for estimating the number of clusters in each image were evaluated and compared with results from human observers. The Caliński-Harabasz index<sup>37</sup>, also known as the variance ratio criterion (VRC), determines the ratio of the sum of between-clusters dispersion and inter-cluster dispersion for all clusters, where the dispersion is the sum of squared distances. A range was set for the number of clusters and the number within the range with the highest VRC chosen as the optimum

number of clusters. The Davies-Bouldin index<sup>38</sup> is based on the ratio of within-cluster distances to between-cluster distances. The measure is unrelated to VRC. The index values were measured for a range of numbers of clusters, and the number with the lowest Davies-Bouldin index was selected as the best solution. Adaptive  $k$ -means clustering<sup>39</sup> partitions the pixel values without a predetermined number of clusters and is different from the built-in  $k$ -means++ mentioned elsewhere. The clusters were initialized with the mean values of the RGB image planes and then updated with the Euclidean distances of pixel values from these centers until they stabilized.

**Estimating mutual information.** For a pixel chosen randomly from an image, its RGB values  $\mathbf{a} = (R, G, B)$  may be treated as an instance of a trivariate discrete random variable  $\mathbf{A}$  whose probability mass function (pmf) is  $p$  say. The entropy  $H(\mathbf{A})$  of  $\mathbf{A}$  is defined<sup>18</sup> by

$$H(\mathbf{A}) = - \sum_{\mathbf{a}} p(\mathbf{a}) \log p(\mathbf{a})$$

where  $\mathbf{a}$  ranges over the gamut of pixel values and  $H(\mathbf{A})$  is in bits if, as here, the logarithm is to the base 2. For two images represented by random variables  $\mathbf{A}_1$  and  $\mathbf{A}_2$  with respective pmfs  $p_1$  and  $p_2$ , the mutual information  $I(\mathbf{A}_1; \mathbf{A}_2)$  between  $\mathbf{A}_1$  and  $\mathbf{A}_2$  can be defined as

$$I(\mathbf{A}_1; \mathbf{A}_2) = H(\mathbf{A}_1) + H(\mathbf{A}_2) - H(\mathbf{A}_1, \mathbf{A}_2),$$

where  $H(\mathbf{A}_1, \mathbf{A}_2)$  is the entropy of  $\mathbf{A}_1$  and  $\mathbf{A}_2$  taken jointly<sup>18</sup>. Naïve estimates of  $p$  can be obtained by binning the space of color values  $\mathbf{a}$  into a finite number of cells and counting the frequency of occurrences in each cell. This procedure can, though, lead to bias when the number of samples is small<sup>62</sup>. As elsewhere<sup>63</sup>, a bias-corrected estimator  $\hat{H}_G(\mathbf{A})$  due to Grassberger<sup>64</sup> was used instead. Estimates were made of the mutual information  $I(\mathbf{A}; \mathbf{A}_q)$  between random variables  $\mathbf{A}$  and  $\mathbf{A}_q$  representing, respectively, an original RGB image and its quantized representation by relevant colors.

### Data availability

The images of paintings analysed in this study are available at <https://www.museodelprado.es/en/the-collection>. Software for estimating mutual information is available at <https://github.com/imarinfr/klo>. Software for implementing the clustering methods is available at [https://github.com/kailugaji/Color\\_Image\\_Segmentation](https://github.com/kailugaji/Color_Image_Segmentation), <https://github.com/shaibagon/GCMex>, and [https://github.com/ZT-HT/Clustering\\_minEntropy](https://github.com/ZT-HT/Clustering_minEntropy). Software for converting images to S-CIELAB color space is available at <https://github.com/wandell/SCIELAB-1996/>.

Received: 5 July 2022; Accepted: 3 February 2023

Published online: 21 February 2023

### References

- Arend, L. in *Human Vision and Electronic Imaging VI* 392–399 (SPIE 2001).
- Foster, D. H., Amano, K. & Nascimento, S. M. C. Time-lapse ratios of cone excitations in natural scenes. *Vis. Res.* **120**, 45–60. <https://doi.org/10.1016/j.visres.2015.03.012> (2016).
- Nascimento, S. M. C., Amano, K. & Foster, D. H. Spatial distributions of local illumination color in natural scenes. *Vis. Res.* **120**, 39–44. <https://doi.org/10.1016/j.visres.2015.07.005> (2016).
- Cohen, M. A., Botch, T. L. & Robertson, C. E. The limits of color awareness during active, real-world vision. *Proc. Natl. Acad. Sci.* **117**, 13821–13827. <https://doi.org/10.1073/pnas.1922294117> (2020).
- Amano, K. & Foster, D. H. Influence of local scene color on fixation position in visual search. *J. Opt. Soc. Am. A* **31**, A254–A262. <https://doi.org/10.1364/JOSAA.31.00A254> (2014).
- Sun, P., Chubb, C., Wright, C. E. & Sperling, G. Human attention filters for single colors. *Proc. Natl. Acad. Sci.* **113**, E6712–E6720. <https://doi.org/10.1073/pnas.1614062113> (2016).
- Nilsson, T. What came out of visual memory: Inferences from decay of difference-thresholds. *Atten. Percept. Psychophys.* **82**, 2963–2984. <https://doi.org/10.3758/s13414-020-02032-z> (2020).
- Pilarczyk, J., Kuniecki, M., Wołoszyn, K. & Sterna, R. Blue blood, red blood. How does the color of an emotional scene affect visual attention and pupil size?. *Vis. Res.* **171**, 36–45. <https://doi.org/10.1016/j.visres.2020.04.008> (2020).
- Hwang, A. D., Wang, H.-C. & Pomplun, M. Semantic guidance of eye movements in real-world scenes. *Vis. Res.* **51**, 1192–1205. <https://doi.org/10.1016/j.visres.2011.03.010> (2011).
- Marin-Franch, I. & Foster, D. H. Number of perceptually distinct surface colors in natural scenes. *J. Vis.* **10**, 9–9. <https://doi.org/10.1167/10.9.9> (2010).
- Foster, D. H. The Verriest Lecture: Color vision in an uncertain world. *JOSA A* **35**, B192–B201. <https://doi.org/10.1364/JOSAA.35.00B192> (2018).
- Nieves, J. L., Ojeda, J., Gómez-Robledo, L. & Romero, J. Psychophysical determination of the relevant colours that describe the colour palette of paintings. *J. Imaging* **7**, 72. <https://doi.org/10.3390/jimaging7040072> (2021).
- Masterpieces, Prado Museum, Spain <https://www.museodelprado.es/en/the-collection> (2022).
- Khan, F. S., Beigpour, S., van de Weijer, J. & Felsberg, M. Painting-91: A large scale database for computational painting categorization. *Mach. Vis. Appl.* **25**, 1385–1397. <https://doi.org/10.1007/s00138-014-0621-6> (2014).
- Hunt, R. W. G. & Pointer, M. R. *Measuring Colour* 4th edn. (Wiley, 2011).
- Nieves, J. L., Gómez-Robledo, L., Chen, Y.-J. & Romero, J. Computing the relevant colors that describe the color palette of paintings. *Appl. Opt.* **59**, 1732–1740. <https://doi.org/10.1364/AO.378659> (2020).
- CIE. Colorimetry, 4th Edition. Report No. CIE Publication 015:2018, (CIE Central Bureau, Vienna, 2018).
- Cover, T. M. & Thomas, J. A. *Elements of Information Theory* 2nd edn. (Wiley, 2006).
- Arndt, C. *Information Measures: Information and its Description in Science and Engineering* (Springer, 2004).
- Arthur, D. & Vassilvitskii, S. in *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms* 1027–1035 (2007).
- Wang, R. *Color\_Image\_Segmentation*, [https://github.com/kailugaji/Color\\_Image\\_Segmentation](https://github.com/kailugaji/Color_Image_Segmentation) (2020).
- Karayannis, N. B. in *Proceedings of 3rd IEEE International Conference on Fuzzy Systems* 630–635 (1994).
- Bonald, T. Expectation-maximization for the Gaussian mixture model, lecture notes, Telecom Paris, Institut Polytechnique de Paris (2019). <https://perso.telecom-paristech.fr/bonald/documents/gmm.pdf>.
- Vinh, N. X. & Epps, J. in *Proceedings of the 10th IEEE International Conference on Data Mining*. 521–530.



25. Vinh, N. X. *The minEntropy Algorithm for Alternative Clustering* <https://www.mathworks.com/matlabcentral/fileexchange/32994-the-minentropy-algorithm-for-alternative-clustering> (2022).
26. Tirandaz, Z. & Croucher, M. *Modified minEntropy Clustering: Partitional Clustering Using a Minimum Conditional Entropy Objective for Image Segmentation*, [https://github.com/ZT-HT/Clustering\\_minEntropy](https://github.com/ZT-HT/Clustering_minEntropy) (2022).
27. Boykov, Y., Veksler, O. & Zabih, R. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**, 1222–1239. <https://doi.org/10.1109/34.969114> (2001).
28. Kolmogorov, V. & Zabih, R. What energy functions can be minimized via graph cuts?. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 147–159. <https://doi.org/10.1109/TPAMI.2004.1262177> (2004).
29. Boykov, Y. & Kolmogorov, V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 1124–1137. <https://doi.org/10.1109/TPAMI.2004.60> (2004).
30. Bagon, S. *Matlab Wrapper for Graph Cuts*, <https://github.com/shaibagon/GCMex> (2006).
31. Zaslavsky, N., Kemp, C., Regier, T. & Tishby, N. Efficient compression in color naming and its evolution. *Proc. Natl. Acad. Sci.* **115**, 7937–7942. <https://doi.org/10.1073/pnas.1800521115> (2018).
32. Cook, R. S., Kay, P. & Regier, T. in *Handbook of Categorization in Cognitive Science* Ch. 9, 223–241 (Elsevier, 2005).
33. Witzel, C. Misconceptions about colour categories. *Rev. Philos. Psych.* **10**, 499–540. <https://doi.org/10.1007/s13164-018-0404-5> (2019).
34. Othman, A., Wook, T. & Arif, S. M. Quantization selection of colour histogram bins to categorize the colour appearance of landscape paintings for image retrieval. *Int. J. Adv. Sci. Eng. Inf. Technol.* **6**, 930–936. <https://doi.org/10.18517/ijaseit.6.6.1381> (2016).
35. Cleveland, W. S. *The Elements of Graphing Data* (Hobart Press, 1994).
36. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap* (Chapman & Hall/CRC, 1994).
37. Caliński, T. & Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat. Theory Methods* **3**, 1–27. <https://doi.org/10.1080/03610927408827101> (1974).
38. Davies, D. L. & Bouldin, D. W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1**, 224–227. <https://doi.org/10.1109/TPAMI.1979.4766909> (1979).
39. Dixit, A. *Adaptive Kmeans Clustering for Color and Gray Image*, <https://uk.mathworks.com/matlabcentral/fileexchange/45057-adaptive-kmeans-clustering-for-color-and-gray-image> (2022).
40. Wallace, G. K. The JPEG still picture compression standard. *IEEE Trans. Consum. Electron.* **38**, xviii–xxxiv. <https://doi.org/10.1109/30.125072> (1992).
41. Zhang, X. & Wandell, B. A. A spatial extension of CIELAB for digital color-image reproduction. *J. Soc. Inf. Disp.* **5**, 61–63. <https://doi.org/10.1889/1.1985127> (1997).
42. Zhang, X. & Wandell, B. *SCIELAB-1996*, <https://github.com/wandell/SCIELAB-1996/> (1998).
43. Sun, P.-L. & Morovic, J. in *Tenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications* Vol. 10 55–60 (Society for Imaging Science and Technology, Scottsdale, AZ, 2002).
44. Foster, D. H. & Reeves, A. Colour constancy failures expected in colourful environments. *Proc. R. Soc. B* **289**, 20212483. <https://doi.org/10.1098/rspb.2021.2483> (2022).
45. Foster, D. H., Marín-Franch, I., Amano, K. & Nascimento, S. M. C. Approaching ideal observer efficiency in using color to retrieve information from natural scenes. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **26**, B14–B24. <https://doi.org/10.1364/JOSAA.26.000B14> (2009).
46. Lapidoth, A. Nearest neighbor decoding for additive non-Gaussian noise channels. *IEEE Trans. Inf. Theory* **42**, 1520–1529. <https://doi.org/10.1109/18.532892> (1996).
47. Boynton, R. M. & Olson, C. X. Saliency of chromatic basic color terms confirmed by three measures. *Vis. Res.* **30**, 1311–1317. [https://doi.org/10.1016/0042-6989\(90\)90005-6](https://doi.org/10.1016/0042-6989(90)90005-6) (1990).
48. Itti, L. & Koch, C. Computational modelling of visual attention. *Nat. Rev. Neurosci.* **2**, 194–203. <https://doi.org/10.1038/35058500> (2001).
49. Parkhurst, D., Law, K. & Niebur, E. Modeling the role of saliency in the allocation of overt visual attention. *Vis. Res.* **42**, 107–123. [https://doi.org/10.1016/S0042-6989\(01\)00250-4](https://doi.org/10.1016/S0042-6989(01)00250-4) (2002).
50. Nascimento, S. M. C. *et al.* The colors of paintings and viewers' preferences. *Vis. Res.* **130**, 76–84. <https://doi.org/10.1016/j.visres.2016.11.006> (2017).
51. Nuthmann, A. & Malcolm, G. L. Eye guidance during real-world scene search: The role color plays in central and peripheral vision. *J. Vis.* **16**, 1–16. <https://doi.org/10.1167/16.2.3> (2016).
52. Moroney, N. & Fairchild, M. D. Color space selection for JPEG image compression. *J. Electron. Imaging* **4**, 373–381. <https://doi.org/10.1117/12.217266> (1995).
53. Clark, D. The popularity algorithm. *Dr. Dobbs J.* **21**, 121–127 (1995).
54. Abbas, Y., Alsultany, K. & Shilbayeh, N. Applying popularity quantization algorithms on color satellite images. *J. Appl. Sci.* **1**, 530–533. <https://doi.org/10.3923/jas.2001.530.533> (2001).
55. Montagner, C., Linhares, J. M. M., Vilarigues, M. & Nascimento, S. M. C. Statistics of colors in paintings and natural scenes. *J. Opt. Soc. Am. A* **33**, A170–A177. <https://doi.org/10.1364/JOSAA.33.00A170> (2016).
56. Feitosa-Santana, C., Gaddi, C. M., Gomes, A. E. & Nascimento, S. M. C. Art through the colors of graffiti: From the perspective of the chromatic structure. *Sensors* **20**, 1–12. <https://doi.org/10.3390/s20092531> (2020).
57. Nakauchi, S. *et al.* Universality and superiority in preference for chromatic composition of art paintings. *Sci. Rep.* **12**, 4294. <https://doi.org/10.1038/s41598-022-08365-z> (2022).
58. To, M. P. & Tolhurst, D. J. V1-based modeling of discrimination between natural scenes within the luminance and isoluminant color planes. *J. Vis.* **19**, 1–19. <https://doi.org/10.1167/19.1.9> (2019).
59. Boehm, A. E., Bosten, J. & MacLeod, D. I. Color discrimination in anomalous trichromacy: Experiment and theory. *Vis. Res.* **188**, 85–95. <https://doi.org/10.1016/j.visres.2021.05.011> (2021).
60. Menegaz, G., Le Troter, A., Sequeira, J. & Boi, J. M. A discrete model for color naming. *EURASIP J. Adv. Signal Process.* **2006**, 1–10. <https://doi.org/10.1155/2007/29125> (2006).
61. *Psychtoolbox-3*, <http://psychtoolbox.org/> (2022).
62. Kraskov, A., Stögbauer, H. & Grassberger, P. Estimating mutual information. *Phys. Rev. E* **69**, 066138. <https://doi.org/10.1103/PhysRevE.69.066138> (2004).
63. Marín-Franch, I. & Foster, D. H. Estimating information from image colors: An application to digital cameras and natural scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 78–91. <https://doi.org/10.1109/TPAMI.2012.78> (2013).
64. Grassberger, P. *Entropy Estimates from Insufficient Samplings*. (2008). <https://arxiv.org/abs/physics/0307138>.

## Acknowledgments

We are grateful for assistance from Research IT, University of Manchester, and for the use of the Computational Shared Facility, University of Manchester. Part of this work was supported by the Engineering and Physical Sciences Research Council (EP/B000257/1, EP/E056512/1), the Junta de Andalucía (A-TIC-050-UGR18), and the

Ministerio de Ciencia, Innovación y Universidades (RTI2018-094738-B-I00). We thank J. Morovič, P. Smith, and H. Yin for advice and N. Taheri and H. Yin for critically reading the manuscript.

### Author contributions

Z.T. and D.H.F. conceived the research; Z.T. implemented the computer routines and with assistance from D.H.F. performed the analysis; J.L.N and J.R provided image data, code, and observer data from previous work; Z.T. and D.H.F. drafted the manuscript in consultation with J.L.N and J.R. All authors reviewed and approved the final version of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to Z.T. or D.H.F.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023