

# Color Naming Models for Color Selection, Image Editing and Palette Design



Jeffrey Heer

Computer Science Department  
Stanford University  
jheer@cs.stanford.edu

Maureen Stone

Tableau Software  
Seattle, WA  
mstone@tableausoftware.com

## ABSTRACT

Our ability to reliably name colors provides a link between visual perception and symbolic cognition. In this paper, we investigate how a statistical model of color naming can enable user interfaces to meaningfully mimic this link and support novel interactions. We present a method for constructing a probabilistic model of color naming from a large, unconstrained set of human color name judgments. We describe how the model can be used to map between colors and names and define metrics for color saliency (how reliably a color is named) and color name distance (the similarity between colors based on naming patterns). We then present a series of applications that demonstrate how color naming models can enhance graphical interfaces: a color dictionary & thesaurus, name-based pixel selection methods for image editing, and evaluation aids for color palette design.

## Author Keywords

Color names; color modeling; palette design; image editing; visualization; perception; XKCD

## ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: UI

## INTRODUCTION

To reason and communicate about the world, people continually—and often effortlessly—categorize elements of their sensory experience. We perceive a world populated by objects that we label with named types, colors, shapes, tastes, odors, and functions. Our capacity for categorization links sensing of the physical world to language and cognition, and underlies our ability to communicate and reference objects in the world [16]. This observation suggests that user interfaces that model human category judgments might enable more compelling forms of reference and selection.

In this paper, we explore this idea in the domain of *color names*—the linguistic labels used to describe colors. Interest in color names as a means to investigate links between perception, language, and cognition was spurred by the studies of Berlin & Kay [5] in the 1960s. They asked speakers of various languages to name a set of color stimuli and then select the most representative stimulus for each provided color

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI'12, May 5–10, 2012, Austin, Texas, USA.

Copyright 2012 ACM 978-1-4503-1015-4/12/05...\$10.00.

term. Berlin & Kay noted striking regularities among the use of color terms across cultures, leading them to posit 11 universal basic color terms (in English: blue, brown, green, orange, pink, purple, red, yellow, black, grey, and white). They observed that these terms were added to a language in a similar progression across cultures. Subsequent decades of research (e.g., [10, 14, 29, 39]) have challenged and extended this understanding. For example, more than 11 basic color terms may exist (e.g., Russian contains two basic terms for blue [39]), and language can have a relative effect on categorical color perception and memory [29].

The importance of color names to perception and object classification has led researchers to formulate a variety of models describing how people associate names and colors [2, 8, 20, 22, 24, 26]. Here, we extend prior models—in particular, the non-parametric approach of Chuang et al. [8]—to propose a model-construction method for a collection of human color-name judgments. We then demonstrate how the resulting model can be applied to enhance user interfaces. Our research contributions fall into two categories:

First, we contribute a **method for constructing a probabilistic model of color naming** from a large, unconstrained set of human color name judgments. We model color-name associations using multinomial probability distributions that describe the likelihood of a color value given a color name, or vice versa. We then define useful operations in terms of this model. We apply our method to results from a large web-based survey [25] containing over 3 million entries and over 100,000 unique color name responses. Our method includes a scaling technique that reduces this survey data to a small set of maximally information-preserving color terms. We then describe how the model can be used to map between colors and names and we define metrics for *color saliency* (how reliably a color is named) and *color name distance* (the similarity between colors based on naming patterns).

Second, we present **novel interfaces enabled by our color naming model**. These applications help users express color in new ways and offer designers new means for evaluating their designs. We describe a *color dictionary & thesaurus* tool for name-based color selection and browsing of color synonyms and antonyms. We then describe techniques for *name-based pixel selection for image editing*. We introduce methods to (1) suggest highly-probable color names to describe image pixels, (2) select image regions by color name, and (3) provide a new “magic wand” tool based on color name similarity. Finally, we show how color name overlap and saliency statistics supported by our model can aid the *design and evaluation of color palettes for visualization*.

## RELATED WORK

Our research extends two streams of prior work: models of color naming and user interfaces for color design.

### Models of Color Naming

Researchers have devised multiple approaches for modeling human color naming, both to aid scientific understanding and to improve applications such as gamut mapping [24] and image analysis [20, 22]. One approach is to simply partition color space, e.g., by uniform subdivision of hue in HLS color space [17]. Another is to create a color dictionary that maps color names (basic color terms plus modifiers such as “light”, “dark”, and “vivid”) to a single color value. This approach is used by the ISCC-NBS standard [15] and derived variants [4, 11, 20]. These approaches map colors to names in a deterministic, disjoint fashion: they do not model association strength or overlap among color names.

Instead, color scientists use statistical models fitted to a corpus of human color-name judgments. Some researchers employ parametric models using a mixture of Gaussian [24, 26] or Gaussian-Sigmoid distributions [2]. While parametric models can suppress noise and be described with a small set of parameters, they make assumptions about the shape of named color regions that may not match empirical distributions. For example, a Gaussian ellipsoid may assign probability mass to color values outside the gamut of interest.

In response to these shortcomings, others have advocated the use of non-parametric models that avoid assumptions regarding the shape of color name distributions. Moroney [22] models color naming using histograms over a binned color space. Chuang et al. [8] model color-name association using multinomial conditional probability distributions. Chuang et al. also define a statistic for *color salience* —the uniqueness of a color name—in terms of model entropy. Using the 330 colors of the World Color Survey [10], they find good agreement among high salience scores and basic color name “foci” identified in previous work [5, 35].

One issue affecting model construction is the size and granularity of training data. Early work uses results from controlled settings, often limited to a few hundred color stimuli (e.g., [2]). More recent work (e.g., [8, 21, 25, 26]) employs web-based surveys to create larger corpora (though limited to the sRGB gamut of computer monitors). While web-based surveys sacrifice the controlled display and lighting environment of a laboratory, they offer access to a greater variety of people and display types. Researchers have noted consistent results when comparing such crowdsourced data sets with data gathered under controlled conditions [21, 26].

An alternative approach used in computer vision is to construct probabilistic naming models in an automated fashion [30, 36]. Given color names as input, a system can query a search engine for images associated with that color term; the image pixels can be used to fit a statistical model of color-name association. This approach has the advantage of automated construction, and has led to better performance on classification and retrieval benchmarks for photographic im-

ages. However, this approach requires that the color vocabulary be known *a priori*, and provides little insight into the relative likelihood of color terms. Here, we seek to learn the kind and frequency of color terms used by people.

In this paper, we build on the approach of Chuang et al. [8] to construct a non-parametric probabilistic model of color naming. As we will describe, we modify their definition of color salience to improve comparability among colors and introduce name-based color similarity measures. We demonstrate how to construct such models from a large, unconstrained set of color-name judgments. In particular, we use over 3 million color-name responses collected from readers of the web comic XKCD. While prior work [22] has required that the desired number of color names be given as a model parameter, we introduce a scaling routine that uses a measure of information loss to determine the number of color names.

### Tools for Color Design

Researchers in HCI, Computer Graphics, and Visualization have devised myriad tools and algorithms for assisting colorists. One prominent class of applications is interfaces for color selection or palette design [1, 3, 12, 19, 28, 37]. For example, PRAVDAColor [3] and ColorBrewer [12] suggest color palettes for encoding data in visualizations based on data types and/or cardinality. Meier et al. [19] apply theories of artistic color harmony [13, 18] to assist interactive color selection. Cohen-Or et al. [9] use these same theories to formulate automated color harmonization methods for image composition. To our knowledge, none of these tools incorporate nuanced models of color naming to aid color selection or image analysis. In this paper, we explore ways in which color naming models can augment color picking, pixel selection for image editing, and the evaluation of color palettes.

## CONSTRUCTING COLOR NAMING MODELS

To construct a color naming model, we first process the input data: a collection of color-name pairs comprising naming judgments by human subjects. We present a method for reducing unconstrained text responses describing millions of unique colors into a compact table of color-name correspondences. We then construct a non-parametric probabilistic model of color naming that supports saliency and similarity metrics based on color-name association.

### Data Collection

We start with a *color naming data set*: a list of color-name pairs provided by human subjects. Here a “color” (or “color value”) is a stimulus shown to a respondent and a “color name” (or “color term”) is a text label provided by the respondent to describe that stimulus. Prior work has elicited naming judgments using physical color chips or calibrated monitors. Recent work [8, 21, 26] has applied crowdsourcing on the web to collect color naming data and verified that the results are consistent with those from controlled settings.

In this paper, we use a publicly-accessible English-language color naming data set [25] compiled by Randall Munroe, the author of the popular web comic XKCD. The data was collected through a web survey advertised on the XKCD site.

Respondents were asked to first provide basic demographic information: chromosomal sex, native language, color blindness, and optional information regarding monitor type, temperature and gamma. Participants then named color swatches shown against a white background. Each color value was expressed as a coordinate in sRGB color space; values were uniformly sampled from the full RGB cube. The text responses were unconstrained and respondents were free to continue naming new swatches for as long as they wished.

The data set contains over 3.4 million responses from 152,401 sessions (103,430 self-reported males, 41,464 females, and 7,507 declined to state). To our knowledge this is the largest color naming data set in existence. The top ten native languages are English 74.6%, Not stated 12.2%, German 2.7%, Spanish 1.3%, French 1.2%, Dutch 1.1%, Swedish 0.8%, Portuguese 0.6%, Polish 0.5%, and Russian 0.5%. The number of responses per participant ranges from 1 to 2,345 (median 18, inter-quartile range 10–30 responses). To combat malicious responses, the data set includes a per-user “spam score” that penalizes (a) responses that are not used by anyone else and (b) the same response applied across high variations in hue. We keep only the responses from users with normalized spam scores  $\leq 0.5$ . The spam-filtered data set consists of 3,252,134 color-name pairs spanning 2,956,183 unique RGB triples and 132,259 unique color names.

### The Color-Term Count Matrix

The starting point for our model is a table ( $T$ ) of color-term counts in which rows represent colors, columns represent color terms, and each cell contains the count of responses that use a color term to describe a corresponding color. Unsurprisingly, a naïve tabulation yields a very sparse, high-dimensional matrix. In this section, we describe our method for reducing the data to a compact and usable form. To limit the number of colors (table rows), we bin color values within a perceptual color space (CIE L\*a\*b\*). To reduce the number of terms (table columns), we apply a dimensionality reduction method. Finally, we smooth and filter the data.

### Representing Color Values

To reduce the number of unique colors, we need to bin them within a suitable color space. Uniform binning in sRGB is undesirable, as sRGB coordinates model color output devices, not human perception: distances in sRGB space are not generally consistent with perceptual judgments of color difference. We want to group colors in a perceptually uniform fashion, preferably using a simple grid-based scheme.

Accordingly, we bin colors within the standard CIE L\*a\*b\* color space using a D65 reference white point<sup>1</sup>. CIE L\*a\*b\* is a 3-dimensional perceptual color space based on opponent process theory [27]. The L\* dimension represents lightness and ranges from black ( $L^*=0$ ) to white ( $L^*=100$ ). The a\* and b\* dimensions both range from roughly -100 to 100 and correspond to green-red and blue-yellow opponent channels, respectively. Euclidean distances within CIE L\*a\*b\* color

<sup>1</sup>D65 is the standard reference white point used by sRGB, in which the source color values are defined.

space approximate color judgments made by human subjects: a distance of 2.3 is roughly equal to one Just Noticeable Difference (JND) [31]. Measurements made within a local patch of L\*a\*b\* space tend to correlate well with human judgments; however, global measurements across the color space can exhibit significant discrepancies. In response, color scientists have devised more sophisticated metrics for L\*a\*b\* space that provide a closer fit to perceptual judgments. We use the current standard, CIEDE2000 [32], as our primary color distance metric.

For the applications described in this paper, we construct color models using a bin size of 5 units within L\*a\*b\* space. Thus each bin has a radius of  $\sim 1$  JND and so subdivides color space near the theoretical resolution of human color differentiation. Using 5-unit bins, the number of colors (table rows) reduces from 2.3 million down to 8,325. For comparison, using 10-unit bins results in 1,291 colors. To subsequently look up an input color in the table, we map it to its matching bin (interpolation methods could also be used).

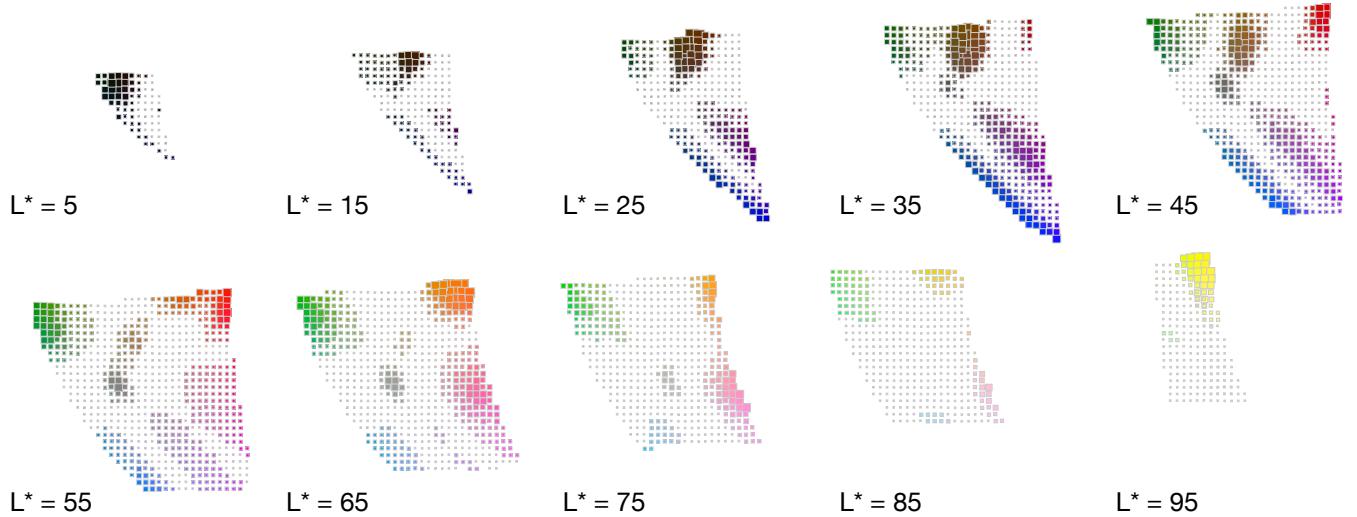
### Processing Color Terms

Unconstrained text responses result in a large and at times amusing variety of color terms. While most responses use common one or two word phrases (“blue”, “lavender”, “dark grey”, “hot pink”), there is a long tail of responses including rare (“butter yellow”) and bizarre (“velociraptor cloaca”) descriptions. Entity resolution is also a concern, as the same color name can have multiple variants (“blue green”, “blue-green”, “blue/green”) and include misspellings (“fuchsia” vs. “fuschia” & “fushia”). We would like to group variants representing the same color term, eliminate noise, and reduce the dimensionality from 132,259 color terms to a more manageable yet representative set.

To handle variants of punctuation and spacing, we first strip all non-alphabetical characters (numbers, punctuation, whitespace, etc) from the response. This process maps many variants to the same term (e.g., the previous blue-green examples map to “bluegreen”), removing about 20,000 variants and reducing the distinct term count to 114,860. We then tabulate a color-term count matrix  $T$  using the resulting terms.

Next, we compute a lower-rank approximation of the color-term matrix to reduce dimensionality and remove noise. The idea is to remove the color terms (columns) that contribute the least to the “information” contained within the data set. We quantify this information using the Frobenius (element-wise) matrix norm  $|T| = \sqrt{\sum_{i,j} (T_{ij})^2}$ .

We first compute the sum of squares of each matrix column and sort the columns in ascending order. We then incrementally subtract these values from  $|T|$  to determine the matrix norm  $|T_k|$  for the color-term count matrix with the  $k$ -lowest columns removed. Given a threshold percentage value  $p$ , we can now find the maximal value of  $k$  such that  $|T_k|/|T| \geq p$ . Retaining 99% of the information ( $p = 0.01$ ) reduces the number of color terms to 66,526 – still a large a number. The model used in this paper retains 95% of the information ( $p = 0.05$ ), resulting in a reasonable set of 179 color names.



**Figure 1.** Color Saliency Map. Each point represents a  $5 \times 5 \times 5$  bin in CIE  $L^*a^*b^*$  color space; slices for every other  $L^*$  value are shown. The area of each point is proportional to its *saliency*: large points indicate consistently named colors, small points indicate colors that exhibit high naming confusion. Interestingly, visible clusters correspond to basic color terms identified by Berlin & Kay (blue, brown, green, grey, pink, purple, etc).

Our reduction technique optimizes the same metric as Singular Value Decomposition (SVD), the method underlying dimensionality reduction techniques such as Latent Semantic Analysis (LSA). Whereas SVD finds a new set of basis vectors to describe the data, our method simply zeroes-out the columns that make the smallest contribution to the matrix norm, preserving each color term as a separate dimension.

While the above method successfully reduces color term dimensionality, it does not handle misspellings. Fortunately, the reduced color space is now small enough that manual correction takes only a few minutes. We add a hand-crafted lookup table to the first stage of our processing pipeline to correct identified misspellings. We then re-tabulate  $T$  on the spell-corrected set of terms and perform dimensionality reduction again to produce a final set of 153 color names.

#### Smoothing & Simplification

To arrive at our final table, we smooth the data and filter isolated responses. To perform smoothing, we first represent the color-term count matrix as a three-dimensional grid in  $L^*a^*b^*$  space. We convolve the grid with a  $3 \times 3 \times 3$  kernel with a value of 4 in the center and 1 in all other positions. For each grid cell, we average over the adjacent non-zero grid cells weighted according to the kernel. We then round the smoothed cell values to the nearest integer count. We find that this approach adequately smooths the space without over-blurring. Next, we zero-out any table cells with a value of 1. We do this for two reasons: (1) if a color-term pair has only a single “vote”, there is no corroboration of the judgment, and (2) dropping these cells significantly reduces the size of the model with no discernible detriment in subsequent applications.

#### A Probabilistic Model of Color Names

Given a color-term count matrix, we can model the probability distributions of colors and names. We then use this model

to compute additional metrics to analyze the effects of color naming. Here we describe our modeling approach and the model-based metrics used in our subsequent applications.

Following Chuang et al. [8], we model the likelihoods of colors and names as multinomial probability distributions. We are concerned with the random variables  $C$  (which takes on color values) and  $W$  (which takes on color names). We use the symbols  $c$  and  $w$  to denote specific values taken by these variables; the symbols also serve as indices for the rows ( $c$ ) and columns ( $w$ ) of the color-term count matrix  $T$ .

We can now express the likelihood of a color value given a specific color name as the conditional probability  $p(C|w)$ . For each color  $c$  taken by  $C$ , we compute the probability by normalizing the rows of  $T$ :

$$p(c|w) = T_{c,w} / \sum_c T_{c,w} \quad (1)$$

Similarly, the probability  $p(W|c)$  of a name given a color is:

$$p(w|c) = T_{c,w} / \sum_w T_{c,w} \quad (2)$$

These distributions give us the likelihood of a color conditioned on a specific term or vice versa. We can also model the association between two colors or between two terms. Given a color  $c$ , the probability of other colors that have been labeled with matching names (the *categorical association* of a color [8]) is given by

$$p(C|c) = \sum_w p(C|w)p(w|c) \quad (3)$$

The categorical association between words  $p(W|w)$  is expressed in a symmetric manner:

$$p(W|w) = \sum_c p(W|c)p(c|w) \quad (4)$$

### Color Saliency

Given a probabilistic setting, we can use common measures to quantify other aspects of color naming. We define *color saliency*—the degree to which a color value is uniquely named—in terms of the entropy of the conditional probability  $p(W|c)$ . Entropy ( $H$ ) is a standard information-theoretic measure of the “randomness” of a distribution. Specifically, it measures the number of bits necessary to encode a random variable. If a color is uniquely named by all respondents, there is no randomness and  $p(W|c)$  will have an entropy of zero. For colors with a high degree of naming disagreement,  $p(W|c)$  will have a correspondingly higher entropy. We thus express saliency in terms of the negative entropy:

$$\text{Saliency}(c) = -H(p(W|c)) = \sum_w p(w|c) \log p(w|c) \quad (5)$$

Figure 2 shows the distribution of saliency scores for the XKCD data. To create a normalized saliency measure, we rescale the values from  $[-4.5, 0]$  to the interval  $[0, 1]$ .

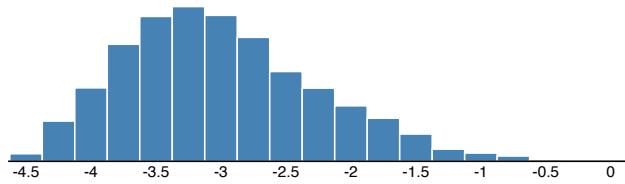


Figure 2. Histogram of XKCD Color Salience Scores.

Our formulation of saliency differs from that of prior work. Chuang et al. [8] define saliency as the negative entropy of the distribution  $p(C|c)$ , the categorical association between colors. Using  $p(C|c)$  allows one to combine (often sparse) color naming data from multiple languages into a unified “term-free” distribution. However, their formulation measures the uniqueness of color names across the entire color space, not at a single value. This occurs because the saliency values are biased by the volume of color space spanned by a name. If one shade of blue and one shade of yellow are both uniquely named (e.g., 100% of responders call color 1 “blue” and call color 2 “yellow”) the yellow shade will receive a higher saliency value. This is because “yellow” occupies a smaller volume of color space and hence has higher  $p(C|w)$  values. To ensure saliency scores are comparable across hues, we instead compute saliency using  $p(W|c)$ .

Figure 1 shows the distribution of saliency scores across  $L^*a^*b^*$  color space, clipped to the sRGB gamut. The saliency scores appear to cluster, with local maxima corresponding to the basic color terms identified by Berlin & Kay. These clusters exhibit different shapes, lending credence to the use of non-parametric models. We can also identify boundaries between color names. Note the region of low saliency between green and blue. This area exhibits high naming confusion, including the terms “teal”, “turquoise”, “green” and “blue.” Note also the narrow “valley” between orange and red at  $L^* = 55$ . This fine-grained detail is visible due to the dense sampling of responses in the XKCD data; sparse data would require larger bin sizes and hence lost detail.

### Name-Based Distance Measures

Our color naming model also enables new distance measures among color values. In addition to metrics within a color space, we can compare two colors by how similarly they are named. One option is to use a metric defined specifically for probability distributions, such as the Hellinger distance:

$$D_h(a, b) = (1 - \sum_w \sqrt{p(w|a)p(w|b)})^{0.5} \quad (6)$$

Though not grounded in probabilistic semantics, another option is to compute the cosine of the angle between two distributions. Due to normalization, this is equivalent to the cosine among rows of the color-term count matrix  $T$ . Denoting the row vector of  $T$  for the color  $c$  as  $T_c$ , we have:

$$D_c(a, b) = 1 - \cos(T_a, T_b) = 1 - \frac{T_a \cdot T_b}{\|T_a\| \|T_b\|} \quad (7)$$

Due to its simplicity and familiarity, we use the cosine metric to measure name-distance between two colors in our subsequent applications. Using other metrics, such as Hellinger distance, produces qualitatively similar results.

## APPLICATIONS OF COLOR NAMING MODELS

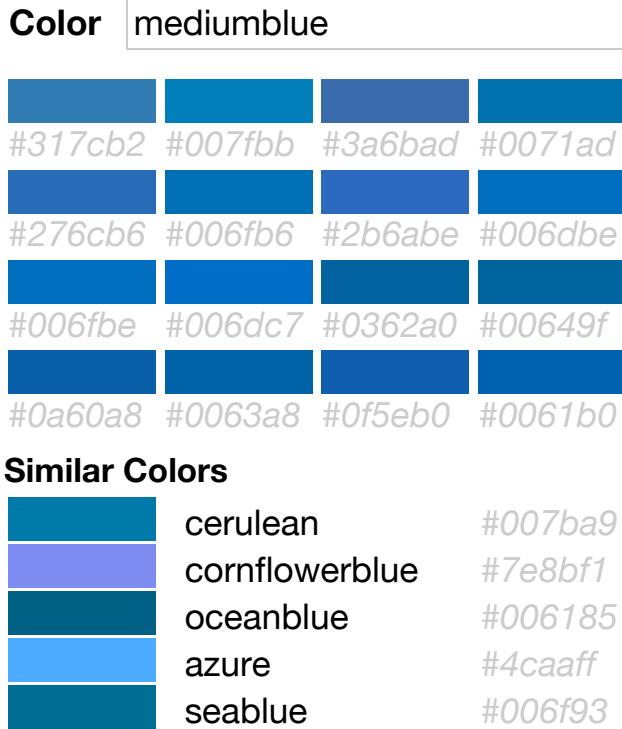
To demonstrate the utility of color naming models for graphical interfaces, we present a series of novel applications: a name-based *color dictionary & thesaurus*, name-based interaction techniques for *selecting image regions*, and an evaluation tool for *color palette design*. The primary data structure used in each of these applications is the color-term count matrix  $T$ , from which we can compute probabilities, color salience, and color distance measures. We have implemented each example as a browser-based web application, written in JavaScript using the D3 (Data-Driven Documents) [6] framework. Both the applications and our library of color modeling routines is freely available as open-source software at: <http://vis.stanford.edu/color-names>.

### Color Dictionary & Thesaurus

A simple and direct application of color naming models is to look up the color values associated with a name (a dictionary) and find other, related color terms (a thesaurus). Figure 3 shows our interface for these tasks. Users might work with the interface to select colors (e.g., for graphic design) or explore relationships among color names.

In response to a color term query, the interface displays the most probable color values matching that color name. Next, the thesaurus view lists related color names ranked by similarity. The top of the list shows highly similar color names (*synonyms*), whereas the bottom (not shown) lists opposed color names (*antonyms*).

*Dictionary.* Given a color name, retrieving representative color values is straightforward: we find the most probable color values according to  $p(C|w)$ . To select a single representative color, we average the four most probable colors (the four largest values of the  $p(C|w)$  distribution). We find that this produces more helpful results than choosing only the single most probable value, particularly for names corresponding to bright, desaturated colors.



**Figure 3.** Color dictionary showing the 16 most probable color values for the query “mediumblue.” The thesaurus lists related color names sorted according to  $p(\text{query} \mid \text{name})$ . Ties are broken by sorting representative colors by their CIEDE2000 distance.

**Thesaurus.** To rank color names, we sort by the name-name association probabilities  $p(W|w)$ . We score each color name by  $p(\text{query} \mid \text{name})$ , the probability of the query term given the name. Empirically, we have found that this choice provides better results than  $p(\text{name} \mid \text{query})$ , which can privilege names with high marginal probability. For the query term “mediumblue,” the name “purple” is ranked highly according to  $p(\text{name} \mid \text{query})$  because it is a common term. Choosing terms that instead maximize the likelihood of the query favors similar shades of blue (shown in Figure 3). The sort order is determined by similarities in naming patterns, which need not be the same as perceptual similarity among representative colors. Again, a color name corresponds to a variably sized distribution across a range of color values.

To break ties, we sort color names according to the perceptual (CIEDE2000) distance between representative colors. As some color names have association probabilities of zero (e.g., the hues yellow, orange, and red were never labeled “blue”), sorting by perceptual distance ensures that color antonyms align with opponent process theory: yellow is the antonym of blue, green is the antonym of red, and so on.

### Using Color Name Models for Image Editing

People in conversation regularly use color names to refer to visual elements. Color name models can enable analogous forms of reference in user interfaces, for example, within image editors such as Adobe Photoshop or Illustrator. Here we describe two techniques for name-based selection: *color name queries* and a *name-based magic wand* selector.

### Color Name Queries

To select image pixels matching a color name  $w$ , we simply include those pixels with non-zero  $p(c|w)$  values. To control the sensitivity of the selection, users can adjust a tolerance parameter to set a minimum  $p(c|w)$  threshold. Users can either type a color term query or interact with the image to generate relevant suggestions. In response to a mouse-driven selection, we show the most probable color names according to  $p(W|c)$ . In the case of a single pixel, we use the corresponding probability distribution. In the case of image regions, we average the distributions for each pixel.

Figure 4 shows our image selection prototype. The left panel displays an image and the right panel shows a list of suggested color names. By default, the list contains the most probable names for the entire image. The bottom panel includes a slider for adjusting selection tolerance (the  $p(c|w)$  threshold). Clicking a pixel or dragging over a rectangular region updates the list to show the most probable color names for those pixels. A user can page through each name-based selection using the up and down arrow keys. A user can also type in a search query, initiating a selection and revealing related color names based on our color thesaurus.

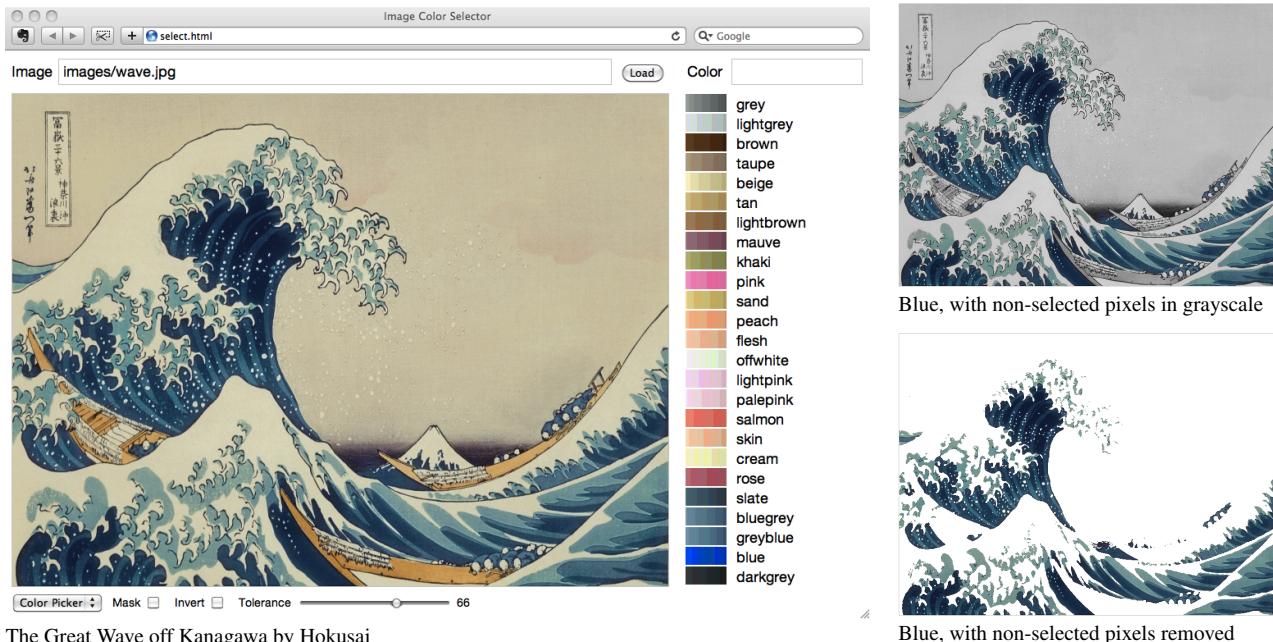
Figures 4–6 show the use of color name queries to isolate image regions of interest. Figure 5 shows selections resulting from a variety of color name queries, including the use of more specific, automatically suggested color names (“olive”, “forestgreen” and “puke”) to select sub-regions of an initial, general query (“green”). Figure 6 shows selections seeded by direct manipulation: the user can click a pixel or drag a region to view related color names and then page through the results to find a desired selection.

### A Name-Based Magic Wand Selector

A common selection mechanism within image editors is the *magic wand* tool. Using the wand tool, a user first clicks a desired pixel. The application then selects all adjacent pixels within a specified color distance using a flood-fill algorithm. Most wand tool implementations measure color distance as the maximal absolute difference in the red, green, or blue channels – the  $L_\infty$  norm in RGB space. A tolerance parameter ranging from 0–255 controls the threshold distance.

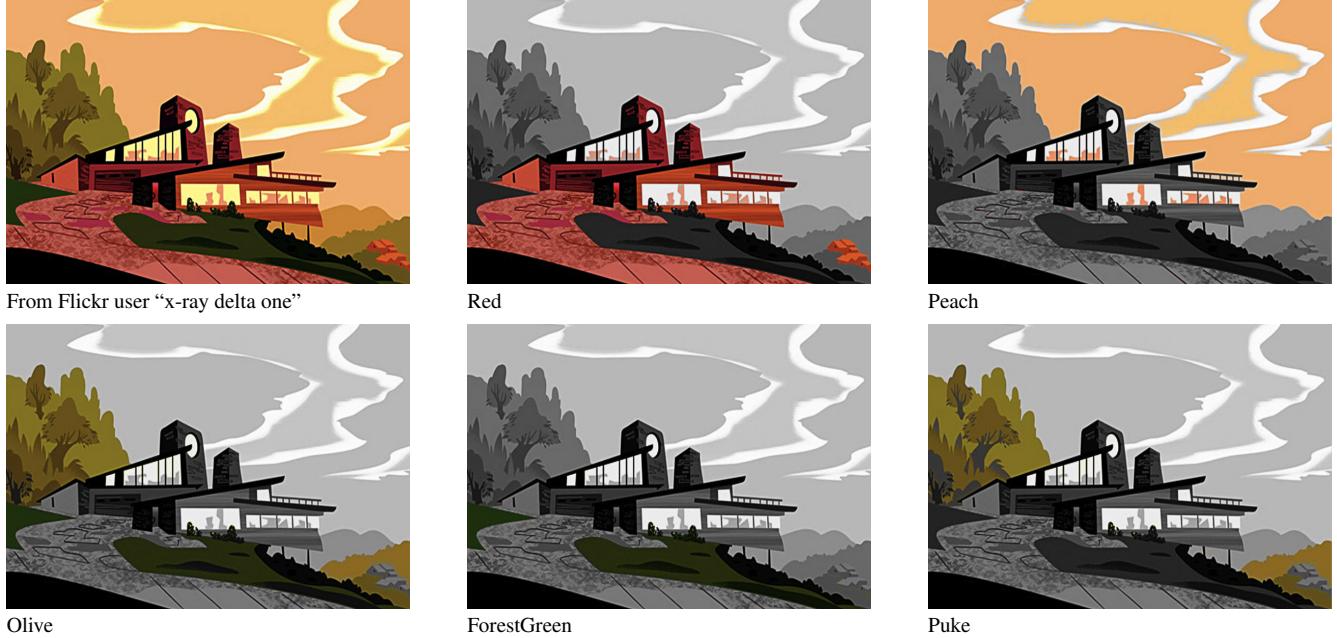
Instead, we can use color naming distance — computed using the cosine of the angle between  $p(W|c)$  vectors — to determine pixel similarity. Our implementation provides a tolerance parameter with range 0–100, which maps to name-based distances on the interval [0, 1].

Figure 7 compares selections from an RGB-based wand and a name-based wand. The name-based selection better respects color name boundaries and is more stable across a range of tolerance settings. At low tolerance settings, RGB distance fails to select a perceptually coherent brown region due to color value variation; at higher tolerances the selection bleeds across color name boundaries, selecting adjacent orange regions. In contrast, the color name wand selects the brown region and excludes the orange pixels. To ensure that this discrepancy is not simply due to the use of RGB space,

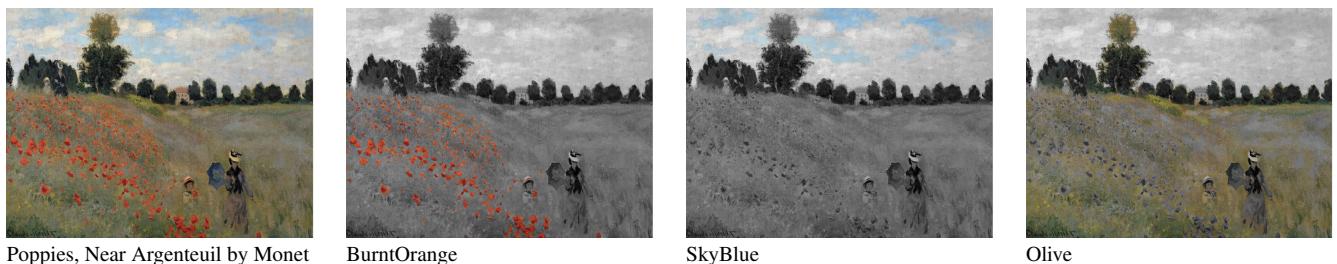


The Great Wave off Kanagawa by Hokusai

**Figure 4.** Image selection by color name query. *Left:* Our prototype selection interface, listing the most probable color names according to  $p(W|c)$ . The many light colors correspond to the sky and boats. Beneath the image are selection options, including a tolerance parameter. *Top-Right:* The selection for the color query “blue”, showing non-selected pixels in grayscale. *Bottom-Right:* The same query with non-selected pixels removed.



**Figure 5.** Image regions selected by color name query. Non-selected pixels are shown in grayscale. The bottom row illustrates selections made when moving from a basic color term (“green”) to more specific terms to isolate foliage (“olive”), dark grass (“forestgreen”) or tree tops (“puke”).



**Figure 6.** Image regions selected by color name query. Non-selected pixels are shown in grayscale. The selected query terms were chosen from a list of most probable colors for an image region selected by mouse-drag. In this case, a user can rapidly isolate flower, sky, and foliage pixels.



Suprematism (Supremus No. 58) by Malevich  
The red circle indicates the input pixel to the magic wand.

**Figure 7. Magic wand pixel selections based on RGB distance (top row) and color name similarity (bottom row) across a range of parameter values. The RGB selector misses pixels at low tolerance values ( $t=30$ ) and includes orange and red hues at higher tolerance values ( $t=60$ ). The color name selector produces more uniform results across parameter settings and is sensitive to color name boundaries.**

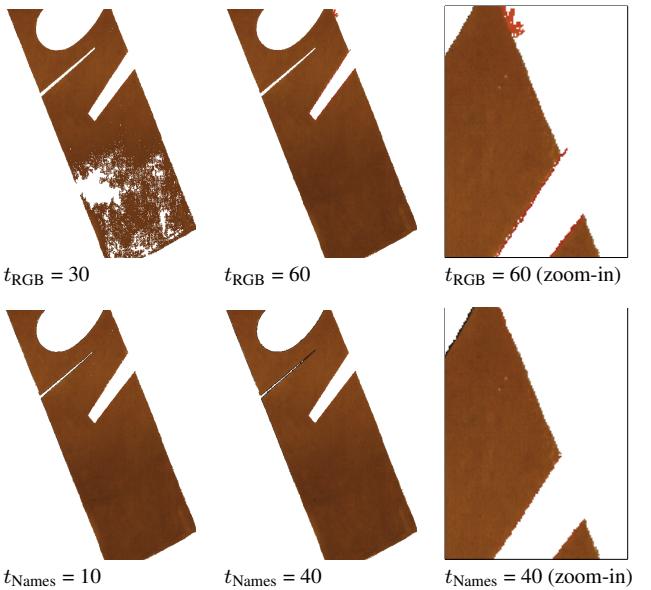
we also implemented a magic wand based on CIEDE2000 distance in  $L^*a^*b^*$  space. We observed qualitatively similar results as the RGB-based selector: low tolerances lead to under-selection and higher tolerances lead to color “bleed.”

Of course, not all name-based selections work so perfectly. We observe that desaturated colors (e.g., pastels) have higher name overlap, leading to less granular selections. Excluding achromatic names (e.g., “grey”) may improve the situation. In addition, we have combined color space and name-based distances within a hybrid wand tool that selects either the union or intersection of the two selection measures. This hybrid tool uses two tolerance parameters (one for names, one for color space) and enables more nuanced selections.

### Evaluating Color Palette Designs

Color design experts [7, 34, 33] argue that attention to color names is important in palette design, particularly for information visualization. First, nameable colors facilitate communication: it is easier to refer to graphical elements when one can name them unambiguously. Second, experimental evidence suggests that highly nameable colors are better remembered [29], perhaps due to the cue of the color name. To inform color palette design, we use our model to quantitatively characterize palettes with respect to color naming. To analyze a color palette, we examine both individual color saliency scores and color name distances. This data can help designers reason about the effects of color choices.

To optimize the presentation of categorical data, we might seek to minimize name overlap (to avoid ambiguity) and maximize salience (to avoid confusion and aid memory). Figure 8 characterizes qualitative color palettes for encoding categorical data. The palettes come from Tableau (a visualization tool with palettes designed by a color specialist [33]), the ColorBrewer selection tool [12], Microsoft Excel, and The Economist magazine. Tables show color name distances



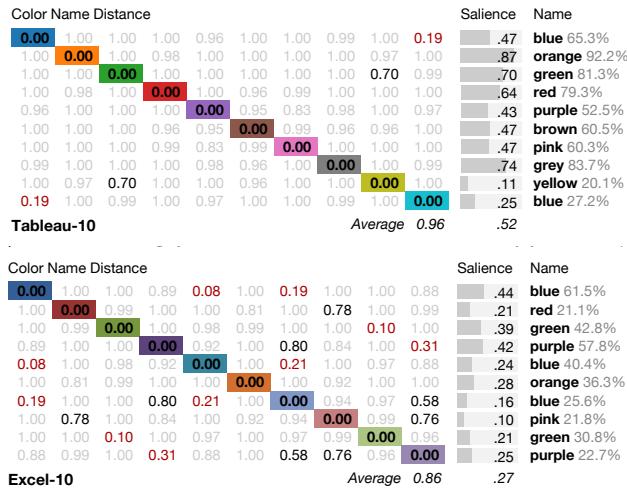
among colors, while bar charts show salience scores. These metrics enable rapid comparison of the palettes. Tableau and ColorBrewer both limit name overlap and include high-salience colors. The Excel and Economist palettes, on the other hand, exhibit high naming overlap and lower salience colors. A designer can use these displays to evaluate different color choices and assess questions such as: “If I shift the cyan more towards green, will it change names?”

Figure 9 shows diverging palettes for numerical data with a meaningful mid-point. The first palette, used in the Map of the Market [38], naively ramps from red through black to green in RGB space. The other three palettes come from ColorBrewer. We note that the Map of the Market palette has significant name overlap on each ramp, while the ColorBrewer palettes show more naming variation. In displays such as choropleth maps, simultaneous contrast can hamper subtle luminance comparisons [34]. By including small shifts in color hue and naming, the ColorBrewer palettes may improve discrimination. Yet by making the shifts subtle, the colors are still perceived as a ramp. ColorBrewer palettes also exhibit salience gradients oriented toward extreme values, emphasizing greater deviation from the midpoint through both luminance and categorical salience.

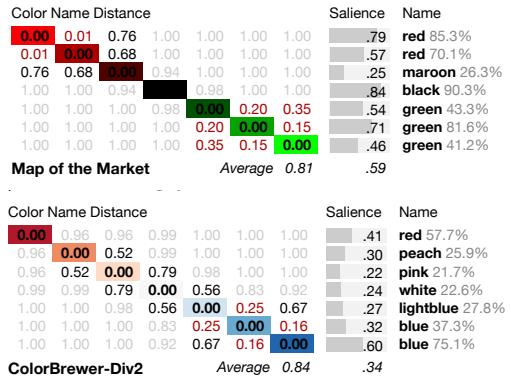
While informative for human designers, these metrics might also improve automated design tools, for example by minimizing name overlap and increasing saliency. Of course, effective color design involves many concerns, including other contrast effects and cultural associations. In future work, we plan to apply our naming model to automatically optimize color palette design and evaluate the results.

### DISCUSSION

In this paper, we presented a method of constructing a probabilistic model of color naming from large, unconstrained naming data sets. We described a model based on over 3 mil-



**Figure 8.** Name-based characterization of qualitative color palettes. Matrices show all pairwise color-name distances; bar charts show salience scores for each color. Salience scores below 0.2 indicate colors with a high degree of naming confusion. The Tableau-10 palette provides the best color salience and minimal name overlap. Palettes from Excel and The Economist exhibit higher name overlap and diminished saliency.



**Figure 9.** Name-based characterization of diverging quantitative color palettes. Naïve interpolation in RGB space (as in the Map of the Market) leads to name overlap among non-adjacent colors. ColorBrewer palettes (excluding “teal”) exhibit salience gradients towards the extreme values.

lion web survey responses — to our knowledge the largest color naming data set in existence — and showed how it can be used to map between colors and names, calculate color saliency, and measure color similarity based on naming patterns. We then introduced a set of novel applications that illustrate how color naming models can enhance graphical user interfaces: a color dictionary & thesaurus, name-based pixel selection techniques for image editing, and evaluation aids for comparing color palette designs. Through these examples, we demonstrate that color naming models enable users to express their intentions in new ways and offer designers new avenues for assessing their designs.

Though our initial results are encouraging, some limitations remain. All respondents in the XKCD survey were asked to name colors shown against a white background, a design decision that biases the results. Unsurprisingly, the region of color space named “white” is small, as respondents were presumably sensitive to background contrast for high luminance colors, naming nearby colors “offwhite”, “cream”, or “light grey.” The term “black,” on the other hand,

occupies a larger volume. Future large-scale surveys might vary the background color among black, white, and one or more shades of grey to construct color naming models more sensitive to background contrast. Alternatively, automated methods (e.g., using image search engines [30, 36]) might be used to refine color name regions.

While our approach uses CIE L\*a\*b\* color space, recent work in color science concerns *color appearance models*. These models incorporate contrast effects due to background and surround regions. Though unable to predict all the intricate effects of human color vision, a color appearance model such as CIECAM02 [23] might associate colors with names while taking contrast into account. Future research is needed to assess the potential benefits and determine if the additional modeling complexity is warranted.

Our current work uses an English-language color naming data set. In subsequent work we would like to construct color naming models across a variety of languages (c.f., [26]). Multi-lingual color naming data can enable scientific inves-

tigation of linguistic differences in color naming, for example to verify basic color terms and assess differences in color name boundaries and saliences. Cross-language models would support internationalization of color name-based applications and could potentially lead to culturally sensitive color transfer or gamut mapping methods.

We developed our applications through a design exploration of how color naming models might significantly enhance user interfaces. The goal of this process was to establish the utility of color naming models and develop new techniques. We believe that each application provides novel support for real-world tasks and initial feedback from informal usage has been positive. Users have expressed appreciation for the applications (including requests for the software) and have suggested interesting use cases. For example, one person used name-based pixel selection to analyze a series of art works and “deconstruct” the coloring choices of the artist.

However, we do not claim that these applications are “finished.” Further design iteration and end-user evaluation will undoubtedly refine these applications and inform how they could be incorporated into existing tools and workflows. To facilitate future work, our color naming model and applications are available as open-source software, downloadable from <http://vis.stanford.edu/color-names>.

## ACKNOWLEDGEMENTS

The authors thank Randall Munroe of XKCD for sharing his color survey data, and Jason Chuang and Pat Hanrahan for their helpful insights and conversation. This work was supported by NSF Grant IIS-1017745.

## REFERENCES

1. Adobe Kuler. <http://kuler.adobe.com/>.
2. R. Benavente, F. Tous, R. Baldrich, and M. Vanrell. Statistical modelling of a colour naming space. In *European Conference on Colour Graphics, Imaging, and Vision*, pages 406–411, 2002.
3. L. D. Bergman, B. E. Rogowitz, and L. A. Treinish. A rule-based tool for assisting colormap selection. In *Proc. IEEE Visualization*, pages 118–125, 1995.
4. T. Berk, L. Brownston, and A. Kaufman. A new color-naming system for graphics languages. *IEEE Comp. Graphics & App.*, 2(3), 1982.
5. B. Berlin and P. Kay. *Basic Color Terms: Their Universality and Evolution*. Univ. of California Press, 1969.
6. M. Bostock, V. Ogievetsky, and J. Heer. D3: Data-driven documents. *IEEE Trans. Visualization & Comp. Graphics*, 17(12):2301–2309, 2011.
7. C. Brewer. Color use guidelines for data representation. In *Proc. Section on Statistical Graphics, American Statistical Association*, pages 55–60, 1999.
8. J. Chuang, M. Stone, and P. Hanrahan. A probabilistic model of the categorical association between colors. In *Color Imaging Conference*, pages 6–11, 2008.
9. D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and Y.-Q. Xu. Color harmonization. In *ACM SIGGRAPH*, pages 624–630, 2006.
10. R. Cook, P. Kay, and T. Regier. *The World Color Survey database: history and use*. Elsevier, 2005.
11. M. Das, R. Manmatha, and E. M. Riseman. Indexing flowers by color names using domain knowledge-driven segmentation. In *Proc. IEEE Workshop on Applications of Computer Vision (WACV'98)*, pages 94–99, 1998.
12. M. Harrower and C. Brewer. Colorbrewer.org: an online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37, 2003.
13. J. Itten. *The Art of Color*. Van Nostrand Reinhold Company, 1960.
14. P. Kay and C. K. McDaniel. The linguistic significance of the meanings of basic color terms. *Language*, 54(3):610–646, 1978.
15. K. L. Kelly and D. B. Judd. Color, universal language and dictionary of names. National Bureau of Standards (U.S.) Publ. 440, 1976.
16. G. Lakoff. *Women, Fire, and Dangerous Things*. University Of Chicago Press, 1990.
17. Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma. Region-based image retrieval with high-level semantic color names. In *Proc. IEEE International Multimedia Modeling Conference*, pages 180–187, 2005.
18. Y. Matsuda. *Color Design*. Asakura Shoten, 1995.
19. B. Meier, A. Spalter, and D. Karelitz. Interactive color palette tools. *IEEE Comp. Graphics & App.*, 24(3):64–72, 2004.
20. A. Mojsilovic. A computational model for color naming and describing color composition of images. *IEEE Trans. Image Processing*, 14(5):690–699, 2005.
21. N. Moroney. Unconstrained web-based color naming experiment. In *Proc. SPIE Color Imaging: Device-Dependent Color, Color Hardcopy and Graphic Arts VIII*, 2003.
22. N. Moroney. US Patent Application 11/259,597: Adapative lexical classification system, 2007.
23. N. Moroney, M. Fairchild, R. Hunt, C. Li, M. R. Luo, and T. Newman. The CIECAM02 color appearance model. In *Color Imaging Conference*, pages 23–27, 2002.
24. H. Motomura. Categorical color mapping using color-categorical weighting method. Part I: Theory. *J. of Imaging Science and Technology*, 45(2):117–129, 2001.
25. R. Munroe. Color survey results. <http://blog.xkcd.com/2010/05/03/color-survey-results/>, May 2010.
26. D. Mylonas, L. MacDonald, and S. Wuergler. Towards an online color naming model. In *Color Imaging Conference*, pages 140–144, 2010.
27. S. Palmer. *Vision Science: Photons to Phenomenology*. MIT Press, 1999.
28. P. Rheingans and B. Tebbs. A tool for dynamic explorations of color mappings. *ACM SIGGRAPH Comp. Graphics*, 24(2):145–146, 1990.
29. D. Roberson, D. I., and J. Davidoff. Colour categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General*, 129:369–398, 2000.
30. B. Schauerte and G. A. Fink. Web-based learning of naturalized color models for human-machine interaction. In *Proceedings of the 2010 International Conference on Digital Image Computing: Techniques and Applications*, pages 498–503, 2010.
31. G. Sharma. *Digital Color Imaging Handbook*. CRC Press, 2002.
32. G. Sharma, W. Wu, and E. N. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Applications*, 30(1):21–30, 2005.
33. M. Stone. Color in information display. <http://www.stonesc.com/VisCourses.htm>.
34. M. Stone. *A Field Guide to Digital Color*. A. K. Peters, 2003.
35. J. Sturges and T. W. A. Whitfield. Locating basic colours in the Munsell space. *Color Research & Applications*, 20(5):364–376, 1995.
36. J. van de Weijer, C. Schmid, J. Verbeek, and D. Larlus. Learning color names for real-world applications. *IEEE Trans. Image Processing*, 18(7):1512–1523, July 2009.
37. L. Wang, J. Giesen, K. McDonnell, P. Zollinger, and K. Mueller. Color design for illustrative visualization. *IEEE Trans. Visualization & Comp. Graphics*, 14(6):1739–1754, Nov-Dec 2008.
38. M. Wattberg. Visualizing the stock market. In *Extended Abstracts, ACM Human Factors in Computing Systems (CHI)*, 1999.
39. J. Winawer, N. Witthoft, M. Frank, L. Wu, A. Wade, and L. Boroditsky. Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, 104(19):7780–7785, 2007.