Lexical Image Processing

Nathan Moroney, Pere Obrador and Giordano Beretta
Hewlett-Packard Laboratories, Palo Alto, California

Abstract

Current color image encoding and processing generally considers the color value of a pixel based on a quantized set of coordinates in an n-dimensional color space. While this approach is useful and practical for numerous reasons, it is informative to consider how these representations and processing techniques might make use of the highest level cognitive description of color, the color name or color category. This paper proposes how a lexical or name-based representation of an image might be generated. Color naming algorithms based on classical logic, prototype theory and fuzzy logic are presented and compared. This paper also describes specific applications of this lexical quantization such as morpholexical analysis, lexical harmony and non-photorealistic rendering.

Keywords: Color naming, perceptual categorization, machine learning, lexical quantization, color picker, web-based experiments, fuzzy logic

Introduction

Lexical image processing is a combination of computational linguistics, machine learning and digital image processing where the fundamental task is the design and implementation of general-purpose, robust and scalable machine color naming algorithms for the color naming of individual image pixels. This paper explores various methods for machine color naming and explores possible applications of the capability to supplement existing color space encodings with high-level cognitive categories.[1] The foundation of this work is a large-scale, task specific database derived using the World Wide Web.[2] There are many applications of this technology and this paper will discuss several of them. A summary flowchart of these techniques and processes are shown in Figure 1. There are many prior publications and much related work[7-35] in the area of color naming but these efforts tends to be focused on a specific aspect of naming or application. The intention of this paper is a more general view of lexical image processing.

A large-scale task-specific database, in this case unconstrained color naming database from thousands of volunteers is shown to the top left. Branching to the top right of figure 1 is then the derivation of color vocabularies which then be used for language-based color pickers or the computation of a color value, say in a specific RGB space, given a color name. A specific example of this type of color picker is the online color thesaurus.[3] Branching to the lower right is then the machine color naming and applications or the assignment of a color names to an input color values, again in a specific RGB space. Given a sizeable color naming database it is then possible to test a range of machine learning algorithms, such as prototype theory or fuzzy logic.

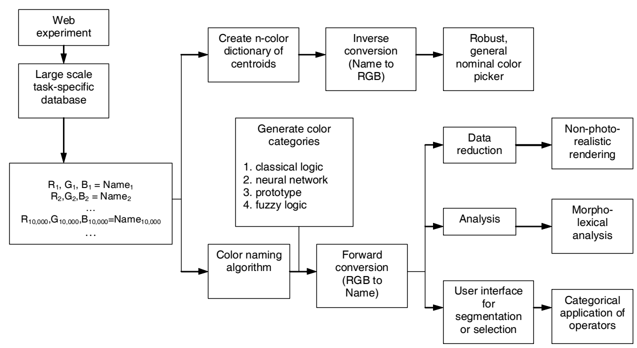


Fig. 1. Overview of techniques relating to lexical imaging.

As a key enabler for this technology, the collection and validation of the color naming database will be described in summary. Starting in 2002 and continuing on until the present over 3,000 participants have provided over 21,000 color names to a color naming database. Participants were asked to provide unconstrained color names for seven colored squares randomly selected from a uniform six by six by six sampling of the RGB cube. Each participant was provided with a random sampling of seven uniform patches using a JavaScript program. Optional comments were captured along with the originating IP address. Multiple submissions per IP address were less than 5% and participants were registered from around the world, although the majority was from North America and Europe. The experiment was conducted in phases beginning with a pool of participants from within the authors' organization. This pool of participants provided a baseline with the assumption that the rates of disruptive observers would be lower than for the external web. This initial phase yielded a scoring algorithm against which new submissions could be graded and rejected in the case of a questionable submission, like providing the hex values for the patches or the first names of women. Roughly five percent of submissions were not used due to likelihood of being a disruptive observation. The scoring algorithm weighted overall mismatch of a given set of color names to the entire corpus more than the use of novel or unfamiliar names. The impact of disruptive participants was further reduced by the distributed design of providing only a few color patches for any given page view. Finally, the correlations with previously published laboratory studies for the basic eleven color names were excellent. For instance the correlation coefficients for the predicted CIELAB hue angle for these hues was 0.99 and 0.98 relative to the results published by Boynton and Olson[34], and Sturges and Whitfield.[35] Similar correlations were also shown for object hues as measured directly or from the web-based data. [2]

Color Naming Algorithms

Given a color naming database, the next major consideration is an algorithm for color naming. A range of algorithms has been published[7-15] but these models are generally based on a fixed hierarchy of limited colors, a limited number of participants or explicit assumptions about the process of color naming. In addition, imaging applications have not been considered for many of these models. Broad categories of algorithms to be considered are classical logic, neural networks, Bayesian learning, prototype and fuzzy models. It is beyond the scope of this paper to consider all of these in detail but it is informative to consider classical logic, prototype and fuzzy naming models in some additional detail. Broadly speaking classical logic models emphasize the locations of the category boundaries, prototype models emphasize the location of the foci or centroids of the categories and fuzzy model emphasize the concept of graded memberships.

The published model[10] of Lin et al. provides a useful baseline for comparing general features of the prototype and fuzzy models. Shown in Figure 2 are four equal lightness slices through a gamut limited CIELAB color space going from darker to lighter from left to right. The slices vary smoothly but after processing with the Lin et al. model the colors are assigned to one of 11 color names. The corresponding color name is shown with a test label. For each slice the colors transition from green to red horizontally and blue to yellow vertically. The center of each slice is an achromatic color such as black, gray or white. The Lin et al. model is based on fixed thresholds for lightness, chroma and hue that were determined through a combination of experimental data and heuristic modeling. In this respect it is a classical logic model in which the data is confined by crisp thresholds. The model also considers only the boundaries and not the location of the focal colors.

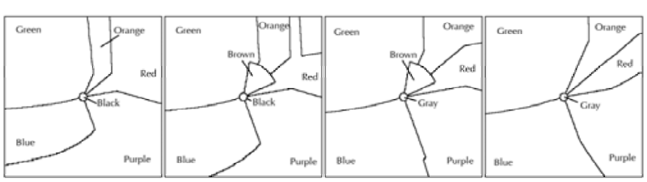


Fig. 2. Classical logic eleven color categorization.

Furthermore, it does not provide any direct means to systematically modify the location of the boundaries for an individual color or provide a clear means to add additional colors. Note that the boundaries shown in the figure show some distortions that are due to the data sampling scheme and not to the model. Specifically, the original data was a uniformly sampled CIELAB grid that was then gamut clipped to sRGB. However the basic qualitative features are evident: a circular achromatic region, straight radial hue boundaries, the nesting of brown and the relative sequence of the hues around the origin.

Prototype theory is a commonly cited model[18,19] for the cognitive process of categorization and color naming. This theory hypothesizes that people learn a single, best focal color for a given color name. The relative similarities to a set of learned prototypes are then used to categorize a given example color. This corresponds to a nearest neighbor calculation and effectively results in a Voronoi partitioning[20] of a color space. For example, given the objective to have a 27 name color vocabulary the color naming database can then be searched for all instances of the specified color names. The 27 names can be selected based on overall frequency analysis, custom considerations or a combination of the two approaches. The prototype for each name can be computed using basic statistics, in this case the arithmetic mean in RGB space. An example partitioning using this approach is shown in Figure 3.

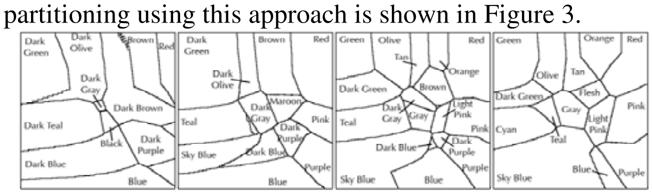


Fig. 3. Prototype naming model using 27 color categories.

The format of the graphs in Figure 3 is the same as that shown in Figure 2. In this case the much larger database of colors enables a simple extension to an arbitrary number of colors. In this case, 27 names were chosen to be comparable with a uniform 3 by 3 by 3 quantization presented later in this paper. The nearest neighbors were computed using the Euclidean distance in the CIELAB color space. Other color spaces and distance metrics are possible, but the results in Figure 4 demonstrate the general properties of prototype models. Specifically, these models will tend to have distinct 'corners' or large straight edges between categories. Furthermore careful determination of the prototypes does not necessarily provide a simple method to systematically expand and contract specific color categories. For example, an expansion of the range of the gray region would require a distance to a line, curve, region or volume and while algorithms exist for this calculation, it is not trivial to parameterize this process. Finally, the calculation of the inter-point distances can be computed in the O(n log(n)) worst-case optimal time in O(n) space.

Fuzzy logic, which has been introduced by Zadeh in 1965, and has since been used in many different applications, is related to the fuzziness in the human thought and reasoning process, whose logic goes beyond traditional two-valued or even multi-valued logic.[21,22] Fuzzy sets, the basic elements, are extensions of crisp sets, which allow only full or no membership. Fuzzy set theory allows the concept of partial membership defined via a membership function $\mu_A(x)$ that can take on values in the interval [0,1]. Per definition, fuzzy sets represent linguistic labels or terms such as slow, fast, low, medium and high. As a consequence, it is very easy to see that color names can be seen as a fuzzy set. A point in n-dimensional color space can be a member of several color names (like orange, red or brown) at the same time in different strengths. Figure 4 shows the resulting partitioning CIELAB color space slices using 27-name fuzzy model.

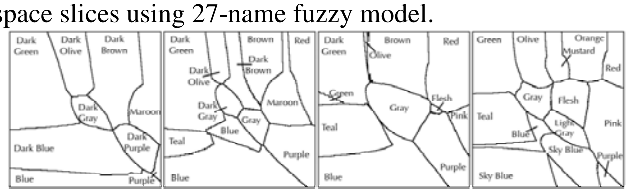


Fig. 4. Fuzzy naming model using twenty-seven categories.

The format used in Fig. 4 is the same as that used for Figures 2 and 3 in that each of the four boxes corresponds to a different approximately constant lightness plane. The results shown for the fuzzy naming model can be compared to those for the prototype naming model. The general trends are similar and it should be noted that there were some small differences in the exact color terms used for each of these figures. Comparing Figure 4 with Figures 2 and 3, the most significant difference is the curvature of name boundaries that results from using the membership functions derived from the aggregate data. It is an open question whether they represent the true nature of cognitive categories but these plots highlight differences between various implementations. From an implementation standpoint it is worth noting that the membership model can be efficiently computed with n times 3 multiplications where n is the number of names. In addition the application of an exponent to the corresponding color membership can be used to easily expand or contract an individual membership or range of a color name. However a fuzzy naming model requires additional memory for the storage of a complete membership function for each color name. For instance an 8-bit function would require 256 time 3 data points to be stored in memory.

Lexical Quantization

Determining the optimal color name for a given pixel can be considered a form of data reduction. For example, the over 16 million individual colors possible within a 24 bit RGB image could be assigned to a much smaller set of color names. This is comparable to various quantization and color palette creation algorithms. However, uniform quantization as one extreme generally does not yield very satisfactory results and palettes generated specifically for one image or optimized for that image but used on another image generally don't yield very good results. Lexical quantization however reduces the higher bit depth color information to highly salient cognitive nodes and as a result yields a considerably improved result relative to other quantization schemes and yet it is more broadly applicable than an image specific color palette. An example comparing lexical quantization to uniform device quantization and the application of the Lin et al color-naming model are shown in Figure 5.

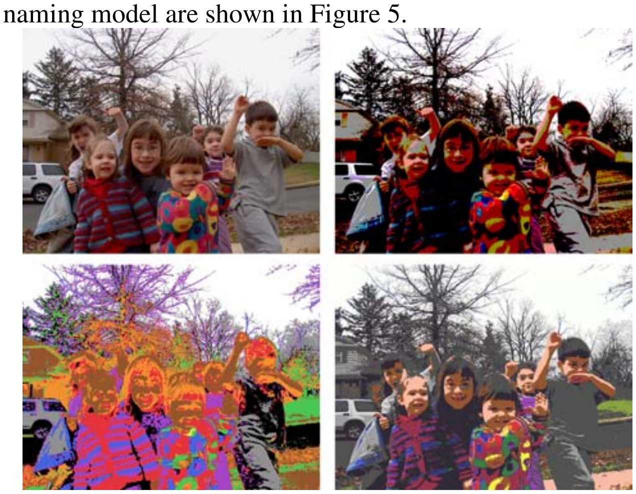


Fig. 5. Original 24 bit RGB image in upper left reduced to 27 colors using uniform device quantization in the upper right. The lower left is the original reduced to 11 color names using the algorithm of Lin et al and the lower right is the image reduced to 27 color names using lexical quantization described in this paper.

The lexical quantization result shown in Figure 5 is based on using 27 color names and the fuzzy color naming model. These names were selected by both considering the most frequently used color terms of the color naming database and by adding a skin tone node, given that this color name was not in the top 27 names. Centroids for these terms were computed by averaging corresponding data in the full experimental data set. The fuzzy memberships for flesh and white were expanded using a power of 0.3 while that of gray was contracted using a power of 3. It should be noted that the Lin et al. model[2] was not optimized for use with images but it is included as an example of a classical logic based model. The uniform quantization is also not expected to work very well and in fact has colors, like the faces, that are both too brownish and too greenish. However the lexical quantization with the 'greedy' flesh category has been applied to the sidewalk as a result of the expanded skin tone memberships.

Lexical Histograms

The lexically quantized image also provides an intuitive way to visualize the summary statistics for an image. One widely used summary statistic for an image is a channel by channel histogram of the colors. While RGB histograms are a powerful and straightforward metric to analyze and edit an image they are not always the most intuitive abstraction of the image color properties. For instance if we consider the two images shown in the top of Figure 10 the RGB histograms are in the middle and the lexical histograms are on the bottom. Visualizing the relative pixel count as "olive" or "brown" on the lower right provides a highly abstracted view of the image color statistics. Using lexical quantization schemes with large and smaller color vocabularies provides a scalable way to adjust this histogram representation.

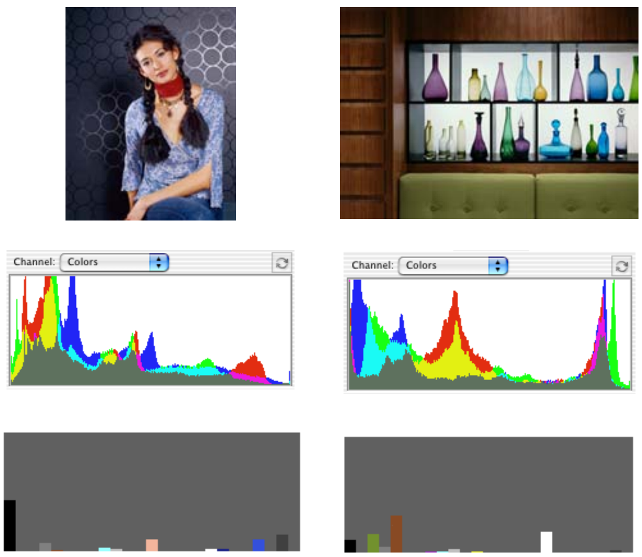


Fig. 6. Two sample images on top, two RGB histograms in the middle and two lexical histograms on the bottom.

Lexical Interfaces

As has been demonstrated previously,[7] a lexical quantization scheme also provides an intuitive user interface for selecting and editing color. It should be emphasized that even the most basic naming scheme provides considerable utility for experts and naïve users alike. Consider for example the image shown in Figure 7.

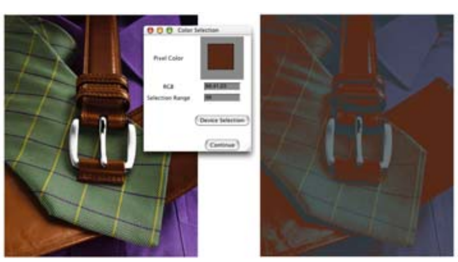


Fig. 7. Selection of pixels using color thresholds in device space.

On the left is a sample image and on the right is a selection of colored pixels. The intention is to select the brown belt and wallet but using regular volumes in device space, in this case RGB, results in a case where there is no threshold for the color channels that is large enough to select all the pixels in the belt without also selecting other pixels. In contrast, the example shown in Figure 8 the same image has the "brown"[27] pixels selected using lexical quantization. The default result is a much better selection of primarily the belt and wallet pixels.

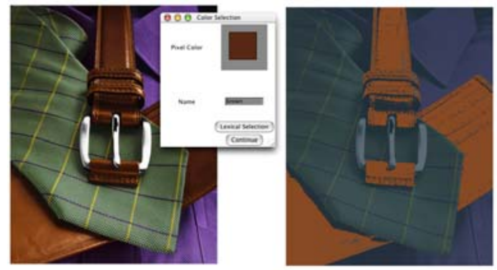


Fig. 8. Selection of pixels using lexical quantization and the identification of "brown" pixels.

Morpholexical Analysis

Morphological imaging or analysis of shape in imaging is largely focused on binary images. Color with it's higher dimensionality provides a range of challenges and considerations for morophological operators. However an image that has been lexically quantized can be subject to the standard opening and closing operators in a more straightforward manner. For example an image with a white background and a reddish figure that has been lexically quantized to red can have the red pixels smoothed in the lexical domain. In the cases where there are multiple color names under consideration in a specific region a series of alternating sequential filters can be applied. A graphical example of this is shown in Figure 9.

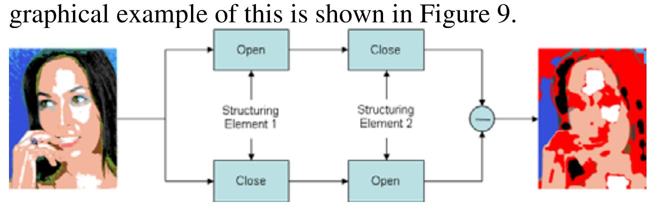


Fig. 9. Graphical example of the proposed filtering stage, with two structuring elements. Pure red at the output signifies "disagreement" between the parallel filtering branches.

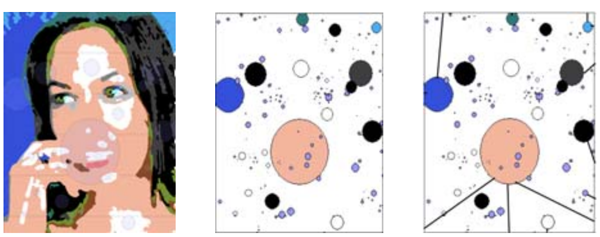
More specifically, an extraction of the underlying color patches at multiple scales is carried out with an alternating sequential filters pyramid, where the structuring elements become increasingly larger. These filter perform the opening and closing in a certain pyramid level with the same structuring element. Next, the size of the structuring element is reduced in the second filtering stage to preserve detail in high color activity regions. The areas of non-agreement have high color activity at this pyramid level and the detail is incorporated from a previous filtering stage. The areas of disagreement are filtered using a template base mode operator that takes into account only the pixels that lie within the areas of disagreement, with a small structuring element. Finally, the areas of agreement are then combined with the output of the template base mode filter. This combination of lexically quantized image and morphological operators has a number of variants but the combination of the two results in morpholexical analysis and provides a powerful means to smooth and segment and image. [5]

Lexical Harmony

Color harmony sets guidelines on how to create effective color combinations. Many attempts have been made, through many historical periods, to create recipes for color harmony. It is, however, not possible to make a list of rules to describe the harmonious or disharmonious visual image. Complementary contrast, whatever the subject, is not a requirement for a harmonious color image. "Ton-sur-ton" or analogous color scheme (where all colors are related to one color hue in slightly different shades or tints) color use doesn't guarantee harmony either. Only the human eye can judge the final artistic result.

The color schemes used the most in harmonization are: *Analogous* scheme: uses any three consecutive hues or any of their tints and shades on the color wheel, *Complementary* scheme: uses direct opposites on the color wheel, *Clashing* scheme: combines a color with the hue to the right or left of its complement on the color wheel, *Monochromatic* scheme: uses one hue in combination with any or all of its tints and shades and *Split complementary* scheme: consists of a hue and the two hues on either side of its complement

Figure 10 shows an input image with morpholexical analysis applied and a further abstraction to regions. These regions are roughly coded by side and proximity to the edge of the image. Given this information shown in Figure 7.c it is then possible to create image color mattes that follow one of the above color schemes.[6]



(a) (b) (c)

Fig. 10. a) "girl" image with superimposed color patch abstraction, b) "girl" color patch abstraction, c) larger patches touching borders abstraction.

Non-Photorealistic Rendering

Within the field of computer graphics there is the research domain of non-photographic rendering. This type of rendering ranges from emulation of illustration techniques to simulation of schools of painting. These processes often make use of some form of color reduction or segmentation.[36]

Fig. 11. Original 24 bit RGB, above, and lexically rendered image, on bottom

Lexical quantization could be a useful component in this process as the input color data can be reduced to a small palette of perceptually salient color categories with a modest amount of chroma expansion and hue preservation. An example of this type of rendering is shown in Figure 11 in which an original 24 bit RGB image is lexically quantized to yield an abstract version of the image. The lexically rendered image has an almost 'paint-by-numbers' appearance relative to the original. Another intriguing application of the lexically quantized image would be to blur the resulting image and use with mask-based spatial color reproduction algorithms to get an image with exaggerated hue-constant 'soft-focus' image renderings..

Conclusions

There has been a range of work touching on machine color naming and in some cases the applications to imaging. However this work has tended to be focused on specific fixed vocabularies and processing techniques and has generally been more focused on specific applications rather than the broadest considerations of lexical rendering. This paper has provided sample graphical results for machine color naming with classical logic, prototype theory and fuzzy logic. Next this paper has considered the applications of these techniques to quantization, histograms, interfaces, morphology analysis, color harmony and non-photorealistic rendering.

References

1. N. Moroney, "Unconstrained web-based color naming experiment", *Proc. SPIE* **5008**, pp. 36-46 (2003).
2. N. Moroney, "Uncalibrated Color", *Proc SPIE* **6058**, (2006).
3. N. Moroney, *The Color Thesaurus*, June 2008 edition, http://magcloud.com/browse/Issue/2344 (2008).
4. G. Beretta and N. Moroney, "Cognitive Aspects of Color" Hewlett-Packard Laboratories Technical Report 2008-109, http://library.hp.com/techpubs/2008/HPL-2008-109.html (2008).
5. Pere Obrador, "Automatic color scheme picker for document templates based on image analysis and dual problem", *Proc. SPIE,* **6076**, (2006).
6. Pere Obrador, "Document layout and color driven image retrieval", *Proc. of ACM SIGIR*, pp. 889-890, (2007).
7. G. Woolfe, "Making color adjustments accessible to non-experts through the use of language", *Proc. 15th IS&T/SID Color Imaging Conference*, 3-7, (2007).
8. H. Motomura, "Categorical color mapping using color-categorical weighting method-Part I: Theory: Color imaging science", *Journal of Imaging Science and Technology*, 45(2), 117-129 (2001).
9. H. Motomura, "Analysis of gamut mapping Algorithms from the Viewpoint of Color Name Matching", *J. SID*, **13**(3), pp. 247-254 (2002).
10. H. Lin, M.R.Luo, L.W.MacDonald and A.W.S. Tarrant, "A Cross Cultural Colour-Naming Study, Part III – A Colour-Naming Model", *Col. Res. Appl.* **26**(4), pp. 270-277 (2001).
11. K. Okajima, A. R. Robertson and G.H. Fielder, "A quantitative network model for color categorization", *Color Res. Appl.* **27**(4), pp. 225-232 (2002).
12. M. Dowman, "A Bayesian Model of Colour Term Acquisition and Typology". Presentated at the *Australian Linguistics Society Conference*, July 13-14, Macquarie University, Sydney (2002).
13. S. Tseng, J.-P. Hsu and T.-H. Wei, "Quantifying Qualitative Relations about Colors Through Fuzzy and Statistical Operations", *Color Res. and App*., **15**(2), pp. 93-98 (1990).
14. A. Mojsilovic, "A Computational Model for Color Naming and Describing Composition of Images", *IEEE Trans. Img,. Proc.,* **14**(5), pp. 690-699 (2005).
15. M. Seaborn and L. Hepplewhite, "Fuzzy Colour Category Map for Content Based Image Retrieval", *Proc. 10th British Machine Vision Conf.* (1999).
16. H. Lin, M.R. Luo, L.W. MacDonald and A.W.S. Tarrant, "A Cross-Cultural Colour-Naming Study. Part I: Using an Unconstrained Method", *Col. Res. Appl*. **26**(1), pp. 40-60 (2001).
17. H. Lin, M.R. Luo, L.W. MacDonald and A.W.S. Tarrant, "A Cross-Cultural Colour-Naming Study. Part II: Using n Constrained Method", *Col. Res. Appl*. **26**(3), pp. 193-208 (2001).
18. J.R. Taylor, *Linguistic Categorization: Prototypes in Linguistic Theory*, Oxford University Press, Oxford, England (1989).
19. Gardenfors and M.A. Williams, "Reasoning about categories in conceptual spaces", *Proc. 14th International Joint Conference of Artificial Intelligence*, pp. 385-392 (2001).
20. J. O'Rourke, *Computational Geometry*, Chapter 5 – Voronoi Diagrams, Cambridge University Press, (1998).
21. A.D. Kulkarni, *Computer Vision and Fuzzy –Neural Systems*, Prentice Hall, USA, (2001).
22. V. Cherkassky and F. Mulier, *Learning from Data*, Chapter 10, John Wiley & Sons, USA, (1998).
23. P. Kay and T. Reiger, "Resolving the question of color naming universals", *Proc. Nat. Acad. Sci.,* **100** (15), pp. 9085-9089 (2003).
24. D.L. Post, "Color-Name Boundaries for Equally Bright Stimuli on a CRT: Phase I", *SID Digest 86*, pp. 70-73 (1986).
25. D.L. Post, "Color-Name Boundaries for Equally Bright Stimuli on a CRT: Phase II", *SID Digest 88*, pp. 65-68 (1988).
26. D.L. Post, "Color-Name Boundaries for Equally Bright Stimuli on a CRT: Phase III", *SID Digest 89*, pp. 284-287 (1989)
27. C. J. Bartleson, "Brown", *Color Res. and App*., **1**(4), pp. 181-191 (1976).
28. R. Kaufmann and M.C. O'Neil, "Colour Names and Focal Colours on Electronic Displays", *Ergonomics*, **36**(8), pp. 881-890 (1993).
29. H. Shinoda, K. Uchikawa and M. Ikeda, "Categorized Color Space on CRT in the Aperature and the Surface Color Mode", *Col. Res. Appl*. 18(5), pp. 326-333 (1993).
30. R. Benavente, F. Tous, R. Baldrich and M. Vanrell, "Statistical Modeling of a Colour Naming Space", *Proc. First CGIV*, pp. 406-411 (2002).
31. S. Tominaga, "A Color-Naming Method for Computer Color Vision", *Proc. IEEE Int. Conf. Cybernetics and Society*, pp. 573-577 (1985).
32. T. Berk, L. Brownstone and A. Kaufman, "A Human Factors Study of Color Notation Systems for Computer Color Graphics", *Comm. ACM*, **25**(8), pp.547-550 (1982).
33. S.N. Yendrikhovskij, "Computing Color Categories", *Proc. SPIE* **3959**, pp. 356-364 (2000).
34. R.M. Boyton and C.X. Olson, "Locating Basic Colors in the OSA Space", *Col. Res. Appl.* **12**(4), pp. 94-105 (1987).
35. J. Sturges and T.W.A. Whitfield, "Locating the Basic Colours in the Munsell Space", *Col. Res. Appl*., 20(6), pp. 364-376 (1995).
36. D. DeCarlo and A. Santella, "Stylization and Abstraction of Photographs", *SIGGRAPH 2002*, pp. 769-776 (2002).