A Large-Scale Multi-Lingual Color Thesaurus

***Albrecht Lindner** †, **Bryan Zhi Li** †, **Nicolas Bonnier** ‡, and **Sabine Süsstrunk** †*
† School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
‡ Océ Print Logic Technologies, Créteil, France

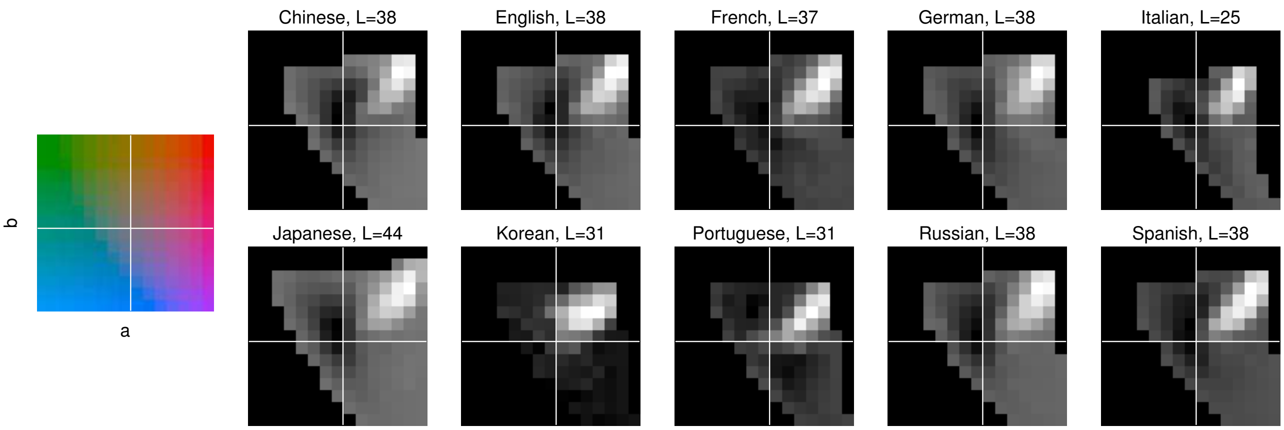


Figure 1: Distributions of the color name *red* for ten different languages (in alphabetical order) in the CIELAB *ab* plane. The left figure shows the colors of the *ab* plane for better orientation. The distributions indicate for each language to which region in color space the color *red* is related. We show a 2-dimensional cross section of the 3-dimensional CIELAB distribution at the *L* value where the distribution is maximal. The black homogeneous areas are out of gamut values. We see that *red* is slightly darker and less saturated in Italian, Portuguese and Korean with respect to other languages. The corresponding color patches for *red* and other color names are visualized in Figure 7.

Abstract

We present a color thesaurus with over 9000 color names in ten different languages. Instead of using conventional psychophysical experiments, we use a statistical framework that is based on search results from Google Image Search. For each color name we compute a significance distribution in CIELAB space whose maximum indicates the location of the color name in CIELAB. A first analysis discusses the quality of the estimations in the context of human language. Further, we conduct an advanced analysis supporting our choice to use a statistical method. Finally, we demonstrate that a color name mainly depends on the chromatic values and varies more along the lightness axis.

Introduction

Color naming is a research topic with a widespread range of related areas, such as color science, linguistics, psychology, anthropology, design, color reproduction and so forth. This paper aims at exploring color naming as a large-scale crowd-sourcing task on the internet. We present a multi-lingual color thesaurus with over 9000 color names in ten different languages that is build without using any psychophysical experiments.

We use a statistical framework that is able to relate semantic color names with color values [8, 9]. The framework is based on a large database of images with annotated keywords. For a given color name, all images with this keyword in the annotation are united in a separate subset. The framework then executes a statistical test that finds the dominant colors in this subset with respect to all remaining images in the database. This is reasonably precise and does not require excessive human labor as is the case with conventional psychophysical experiments.

We build a large-scale database of images with annotations in different languages using Google Image Search [6]. We adapt the search query parameters to obtain search results only from pages of a specific country in a specific language.

As color naming is related to human language and visual appearance, it is impossible to estimate single color values for a given color name. We rather associate a color name with a distribution in CIELAB color space that has a peak in that region where the color is most likely located. The more the distribution is narrow the better the color is defined. And vice versa, the more the distribution is wide the more vague is the color name. All color names in our thesaurus are represented by such a distribution. Examples can be seen in Figures 1, 2 and 3.

The color names we use are based on the XKCD Color Survey [1], a large-scale psychophysical color naming experiment that has been carried out online. This survey resulted in 950 distinct color names and associated color values. We translated the 950 color names to nine other languages, which are Chinese, French, German, Italian, Japanese, Korean, Portuguese, Russian, and Spanish, respectively. All translations have been done by native speakers with a good level of English.

In the last sections we discuss the quality of the color estimations in the context of human language. This is important as color names are subject to vaguenesses of a language. We further undertake a more advanced data analysis and show that colors are mainly lightness invariant rather than hue or saturation invariant.

The code, the estimated color distributions and other results are available for non-commercial and research purposes under: `http://ivrg.epfl.ch/ColorThesaurus`.

Related Work

Naming colors has been of importance ever since people started to communicate colors to each other [5]. There are two opponent explanations that describe how a language develops names for colors. The first is that color terms are universal as proposed by the study from Berlin and Kay [2]. The authors state that there are eleven basic color terms that exist in any fully developed language. Not fully developed languages can be classified in different stages, where the most rudimentary languages in the first stage only distinguish between *dark* and *bright*. Language terms are added in the consecutive stages in a somewhat fixed order until the eleven basic color terms are complete. The second opinion is the so called "linguistic relativity", which states that language determines, or at least influences, our perception of the world [3]. In this scenario, people with different native languages group and name colors in different ways. Hence, there is no universal color naming scheme. Our work is not in favor of the one or the other opinion, but rather provides color estimations for different languages that can be used in related research.

Moroney developed a web-based color naming experiment [11, 10] where people are asked to enter the name of a color that is displayed on a uniform patch. The answers for different samples in color space are accumulated in order to match color names with corresponding color values. Mylonas et al. collected similar data in an online experiment and evaluated two algorithms to build a color naming model [12].

The abundance of annotated images on the world wide web has led to other approaches that avoid labour intensive and time-consuming psychophysical experiments. Sekulovski et al. propose a method using mean shift to extract appropriate colors for given song lyrics [13]. They also show estimations for 9 color names in English and Finnish. Weijer et al. present a modified PLSA model that learns color names from images downloaded with Google Image Search or other sources [14]. However, their study is limited to 24 color naming examples.

We previously proposed a statistical framework that relates image characteristics with semantic expressions [8, 9]. This can be used for color naming if the image characteristics are colors and the semantic expressions are color names [8]. The framework also handles other semantic expressions related to color such as *chocolate* or *ferrari*. We base our work on the last study as it is a simple, yet robust, method that easily scales to large databases.

Data mining provides techniques to extract implicit and previously unknown information from raw data [16]. Free access to large collections of annotated images on e.g. Flickr, Google Image Search or in the form of prepared databases [7] provide a rich source of data to mine for information. In our case we want to estimate for a given color name the corresponding significance distribution in color space.

Building a Database

We took the 950 English color names that were derived in the XKCD Color Survey [1] and translated them into nine other languages, namely Chinese, French, German, Italian, Japanese, Korean, Portuguese, Russian, and Spanish, respectively. Each translation has been done by a native speaker with a good level of English.

In some cases the translation of a color name is difficult, because the destination language does not have this precise color name, or because two varieties of a color name in English translate to the same expression in the destination language. Examples are the four color names *burple*[1], *purpleish blue*, *purpley blue*, and *violet blue*, which all translate to the same expression in Chinese.

We download for all color names and all languages 100 images each, using Google Image Search. To guarantee that we acquire only images from the present language we use the **cr** (country restrict) and **lr** (language restrict) fields as defined in Google's Custom Search API [6]. This is important for color names such as *rose* that have the same spelling in English and French. A simple query for *rose* would therefore lead to an undesired mixed search result from both languages. The search query is the "color name" in quotes plus the word color in the respective language. Two example queries are *"cloudy blue"+color* and *"bleu nuageux"+couleur* for English and French, respectively.

A complete set for one language comprises $100 \times 950 = 95\,000$ images, which has a download time in the order of one day. This process can run in the background as it does not require significant computational power. We assume that the downloaded JPEG images are encoded in sRGB color space.

Statistical Color Value Estimation

We use a statistical framework that relates semantic expressions with image characteristics [8, 9]. It computes a significance score that expresses how much a semantic expression, i.e. color name, is related to a color value. The significance value z is positive (negative) if the color value is dominantly present (absent) in images annotated with the respective color name. The significance values is close to zero if there is no relation. Please see Lindner et al. [8, 9] for details.

Practical Example

We use a histogram in CIELAB color space with $15 \times 15 \times 15$ equidistant bins in the ranges $L \in [0\ 100]$ and $a, b \in [-80\ 80]$, respectively. For each histogram bin we compute the z value with the statistical framework using our image database downloaded with Google Image Search.

Figure 2 shows the resulting z value distribution for *pink* in the English database. The three orthogonal planes show cross sections of the distribution and intersect in the maximum. The z values are encoded with a gray level heat map as indicated by the vertical bar. The color plane at the figure's bottom shows the histogram bin colors for the horizontal plane with constant L value that goes through the maximum z value. We can see that this maximum is at a pink color in CIELAB space and that the z values attenuate with increasing distance from the maximum.

A similar plot is shown in Figure 3, but for the Chinese color name 绿色, which means *green* in English. This plot again shows how the z values decrease in all directions with increasing distance from the maximum.

Interpolating quantized bin centers

We perform two steps to determine for a given color name its estimated color values $L^{\text{est}}, a^{\text{est}}, b^{\text{est}}$. First, we find the maximum bin of the z value distribution. As the bin centers are quantized we do an interpolation step in the neighborhood of the maximum

[1] A combination of *blue* and *purple*

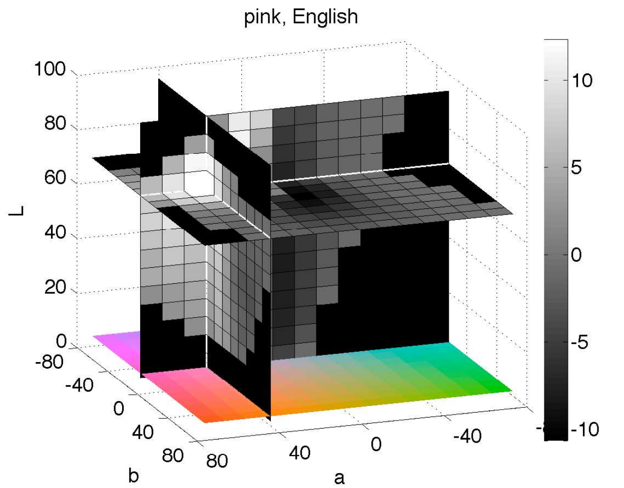


Figure 2: The z value distribution for *pink* in English in a 3-dimensional heat map. The maximum is located at the crossing of the three orthogonal planes. The homogeneous dark areas at the plane borders are out-of-gamut values. At the bottom, we show the histogram bin colors for the constant L plane through the maximum value for a better orientation in CIELAB space.

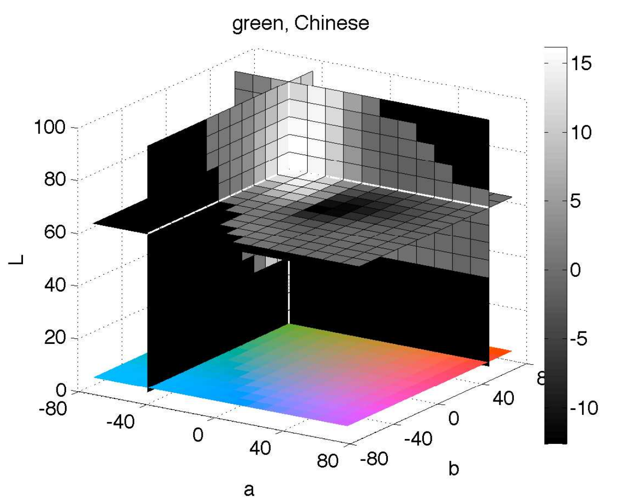


Figure 3: The same plot as in Figure 2, but for the Chinese color name for *green*.

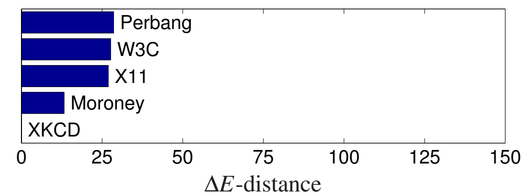
bin. We compute a weighted mean over the 27 bin neighborhood $\mathcal{N}$ in 3-dimensional CIELAB space, where the weights are given by the z values: $L^{\text{est}} = \sum_{i\in\mathcal{N}} z_i L_i \Big/ \sum_{i\in\mathcal{N}} z_i$, where L_i is the L value that corresponds to bin i. The a^{est} and b^{est} values are computed accordingly.

Precision Analysis

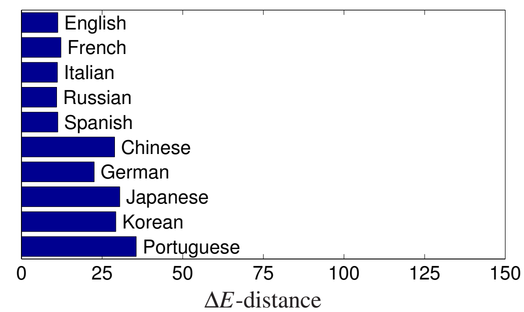
A standard validation technique to analyze the precision of estimated color values is to compare them against ground truth data. However, this is difficult in color naming due to the lack of reliable ground truth data. In fact, it is almost impossible to create reliable ground truth data, because color naming involves natural language, which is too vague for a strictly quantitative validation.

Let us consider the color name *maroon* whose sRGB values are given in several color databases: 64, 35, 39 (Perbang, an online color database [4]), 128, 0, 0 (W3C's CSS Color Module Level 3 [15]), 176, 48, 96 (X11 Color Names [17]), 140, 28, 61 (Moroney's web-based experiment [11]), and 101, 00, 33 (XKCD Color Survey [1]). Our estimate for English is 100, 32, 40. It is difficult to decide with certainty which of the color values represents the true *maroon*.

Figure 4(a) shows the ΔE distances between the color values for *maroon* from the XKCD database and the other databases. The distances between arbitrary pairs of databases are even larger; the maximum is for Perbang's and W3C's values: ΔE=49. For a better visual comparison, the horizontal axes in Figures 4 and 5 have the same scale.



(a) distance between XKCD and different databases



(b) distance between XKCD and ten different languages

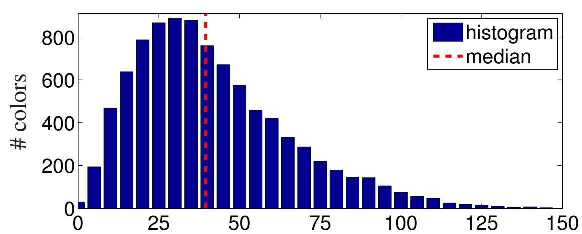
Figure 4: Top: ΔE distances between the color value for *maroon* from the XKCD database and the values from other databases. Bottom: ΔE distance between the XKCD value for *maroon* and our estimations for all languages. The horizontal axes have the same scale as the ones in Figure 5 for a better visual comparison.

We argue that a discussion about the true color value of *maroon*, and any other color name, is strongly influenced by opinions/tastes and can not be taken as a fact. Consequently, a performance evaluation such as measuring the widely used ΔE distance in CIELAB space between our estimates and a ground "truth" has to be carefully interpreted (see Fig. 4).

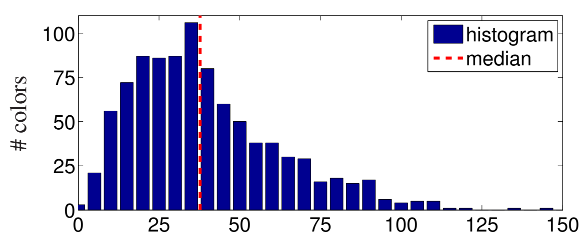
It is also non-trivial to compare results from translations of a single color name into different languages. Our French translation for *maroon* is *bordeaux* and we estimate it as 83, 20, 30. If we translate the French expression back to English we could also say *bordeaux red* or *dark red*, which makes the French estimation justifiable. The German translation is *kastanienbraun*, which literally means *chestnut brown*. Hence, our estimation has a brown tint 70, 29, 27. The Italian translation is *rosso bordeaux*, which means *reddish bordeaux* and our estimation is accordingly more reddish 101, 33, 41. For Portuguese we have *castanho* (*chestnut*) and obtain 73, 54, 41. The Chinese translation is 栗色 (*chestnut* + *color*) 63, 33, 25, the Korean is 적갈색 (*reddish* *brown* + *color*) 39, 0, 0, and the Russian is бордовый (wine red) 85, 19, 31. The Japanese color name is the same as the Chinese, because the translator could not find a corresponding expression and thus used the Chinese vocabulary; a common practice in Japan. Nevertheless, we estimate a different value as we use Google's language and country restrictions: 96, 62, 48.

The ΔE distances between the XKCD value and our estimations for all languages are plotted in Figure 4(b). We can split the languages into two groups. First, languages in which *maroon* has been translated to some type of *red* (top 5 in Fig. 4(b)). In these cases the ΔE distances are lower than for any database (see Fig. 4(a)). In the other group of languages the translation is related to *chestnut* and *brown*. In these cases the estimations are more brownish and the ΔE distances are higher.

Figure 5(a) shows the ΔE distances for all color names in all languages between the estimated values and the XKCD value for English. Considering the large distances for a single color name between different databases (e.g. up to ΔE=49 for *maroon*), the estimations are in a reasonable range. Figure 5(b) shows the ΔE distances for only the English color names. As the color names come from the same language there are no additional deviations due to the translation. Consequently, the ΔE distances are smaller than in the global set.



(a) ΔE-distance: all languages



(b) ΔE-distance: English

Figure 5: ΔE distances between the English XKCD color values and our estimations for all languages (top) and only the English terms (bottom). The distances are in a reasonable range considering that color values are subject to vaguenesses of human languages and deviations from translations. This is demonstrated in the text at the example of the color *maroon* and in Figure 4. The dashed red lines indicate the medians of the distributions.

Figure 7 shows color patches for 50 color names in ten languages. The patches are sorted by increasing hue angle of the English color estimation. We see that these example estimations are correct within expected variations due to language and translation imprecisions.

We show two failure cases in Figure 6 in order to discuss the limitations of the statistical approach. Korean is a single outlier among all estimations for *raspberry*. The Korean expression for this color is 나무딸기 where the first two characters mean *wood*

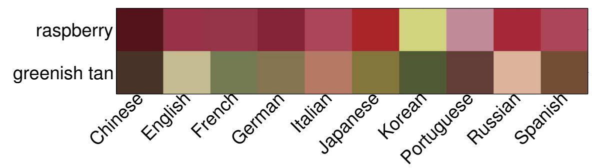


Figure 6: Two failure cases. *Raspberry* fails, because the Korean images with raspberries contain a significant amount of leaves. Hence the estimations is a green color. *Greenish tan* is ambiguous and leads to greenish colors in English, French, and Korean and to skin colors in the other languages.

and the last two *strawberries*. The image results in Korean show raspberries in the woods with a significant amount of green leaves so that green is the most dominant color. The color name *greenish tan* produces ambiguous results. For some of the languages, the framework estimates rather green colors and for others rather tanned skin colors. An interesting case is the German translation *grünlich hellbraun* (*greenish light brown*), which is due to the fact that the German expression for *tanned* literally means "*browned*". However, *greenish light brown* is an expression that is so rarely used that even Google Image search can not provide search results for this query. In this case the term *light brown* dominates the modifier *greenish*.

We can see that imprecisions of natural languages limit the precision of the statistical framework in cases where there is semantic ambiguity or where a semantic concept is difficult to express in the given language. However, this is not a drawback of the automatic estimation, because some of these imprecision can effect conventional psychophysical color naming tasks as well.

Advanced Analysis

The abundance of data allows a more advanced analysis of the estimated significance distributions. In this section we demonstrate two properties: first, higher z values implicate a higher accuracy of the estimated color and second, color names have more variance along the lightness axis than along the two chromatic plane axes in CIELAB space.

Higher significance implicates higher accuracy

Let $\mathcal{L} = \{$Chinese, English, French, German, Italian, Japanese, Korean, Portuguese, Russian, Spanish$\}$ be the set of all languages, $\hat{z}_{l,w}$ the maximum z value of the significance distribution of color name w and language $l \in \mathcal{L}$, and $\mathbf{c}^{\text{est}}_{l,w} = (L^{\text{est}}, a^{\text{est}}, b^{\text{est}})^T$ the estimated color triplet in CIELAB space. We then compute for each color name w the average maximum z value over all languages l, denoted $\bar{z}_w$, and the average ΔE distance between any two estimations of different languages $l_1 \neq l_2$, denoted $\overline{\Delta E}_w$:

$$\bar{z}_w = \frac{1}{|\mathcal{L}|} \sum_{l \in \mathcal{L}} \hat{z}_{l,w} \quad (1)$$

$$\overline{\Delta E}_w = \frac{1}{|\mathcal{L}|(|\mathcal{L}|-1)} \sum_{l_1 \in \mathcal{L}} \sum_{l_2 \in \mathcal{L} \setminus \{l_1\}} ||\mathbf{c}^{\text{est}}_{l_1,w} - \mathbf{c}^{\text{est}}_{l_2,w}||_2 \quad (2)$$

where $|\cdot|$ signifies the cardinality operator and $||\cdot||_2$ the Euclidean distance, respectively.

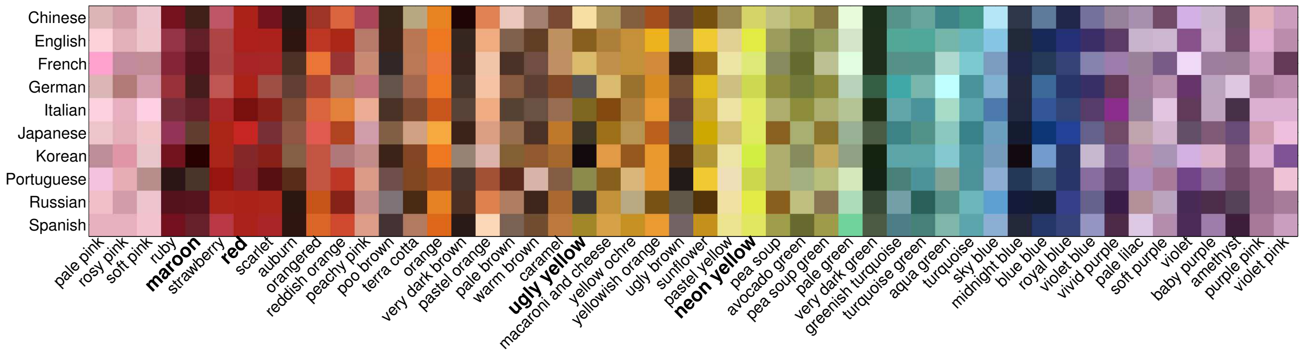


Figure 7: Overview of 50 color names in ten languages. The samples are sorted by increasing hue angle of the English term from left to right. Varying color patches along a column can be due to translation imprecisions as previously discussed at the example of *maroon*. Color names that are referred to in the article are highlighted in **bold font**.

$\overline{\Delta E}_w$ can be visualized as the average deviation of a color name for different languages. For example the deviations for *neon yellow* are smaller than for *ugly yellow* as can be seen in Figure 7, which is reflected in the corresponding values: $\overline{\Delta E}_{neon\ yellow} = 11.5$ and $\overline{\Delta E}_{ugly\ yellow} = 42.1$, respectively. $\overline{\Delta E}_w$ can be high due to estimation errors or translation difficulties as previously discussed for *maroon*.

Figure 8 shows the mean, 25% and 75% quantiles of the $\overline{\Delta E}_w$ values as a function of the corresponding $\bar{z}_w$ value. It is visible that the deviation decreases for increasing average significance. The average significance values for the above example are $\bar{z}_{neon\ yellow} = 8.7$ and $\bar{z}_{ugly\ yellow} = 5.0$, which is in accordance with the overall trend.

We conclude that estimations become better for higher significance values. This is the case when the translated color names are well defined and the related images all have a single dominant color.

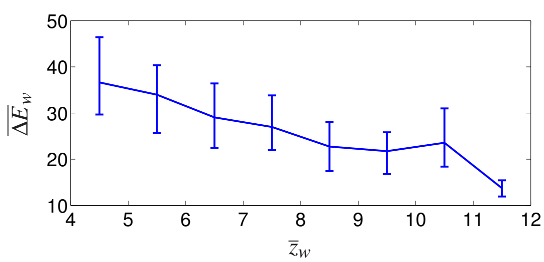


Figure 8: $\overline{\Delta E}_w$ (mean, 25% and 75% quantiles) as a function of $\bar{z}_w$. The deviation of color name estimations for different languages decreases on average with increasing significance.

Tints of a color stretch mainly along the L axis

So far we only considered the maximum bin of the significance distribution and its direct neighbors to estimate a color's CIELAB values. However, the distribution itself contains more information that can be exploited for a deeper insight.

The significance distribution has a blob around the maximum bin and its values decrease with increasing distance from the center, as can be seen in Figures 1, 2 and 3. We compute the 2nd derivative at the estimated color $\mathbf{c}^{\text{est}}$ to determine how quickly the significance values decrease:

$$\left.\frac{\delta^2 z(\mathbf{c})}{\delta L^2}\right|_{\mathbf{c}=\mathbf{c}^{\text{est}}} \approx \left.\frac{z(L^{\text{est}}-\Delta L) - 2z(L^{\text{est}}) + z(L^{\text{est}}+\Delta L)}{\Delta L^2}\right|_{a=a^{\text{est}},b=b^{\text{est}}} \tag{3}$$

where ΔL is the distance between two neighboring bins along the L axis. The equation is analogous for the a and b directions.

The second derivative is always negative in this case, because the z distribution has a maximum at $\mathbf{c}^{\text{est}}$. Therefore, the plot in Figure 9 shows its absolute value, i.e. curvature, for convenience. It is visible that the curvature along the L axis is smaller than along the a and b axes.

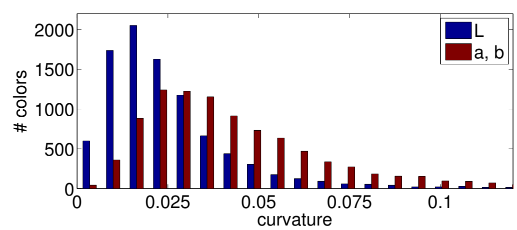


Figure 9: Histogram of the absolute value of the 2nd derivative, i.e. curvature, at the maximum turning point of the z distribution. It is visible that the curvature is smaller in the direction of the L axis than for the a and b axes. This means that color names are more independent of small lightness changes than changes in the chromatic plane.

A similar result is obtained when one fits a Gaussian curve to the z values around the estimated color $\mathbf{c}^{\text{est}}$ in CIELAB-space. We use a symmetric $5\times5\times5$ neighborhood around the center bin and use Matlab's `fminsearch` command to fit a Gaussian function:

$$g(\mathbf{c}) = A\cdot\exp\left[-\frac{1}{2}\left(\frac{(L-L^{\text{est}})^2}{\sigma_L^2} + \frac{(a-a^{\text{est}})^2}{\sigma_{a,b}^2} + \frac{(b-b^{\text{est}})^2}{\sigma_{a,b}^2}\right)\right] \tag{4}$$

where $\mathbf{c} = (L,a,b)^T$ is a position in CIELAB, σ_L the standard deviation in L direction and $\sigma_{a,b}$ the standard deviation in the a and b directions, respectively. The histogram in Figure 10 shows

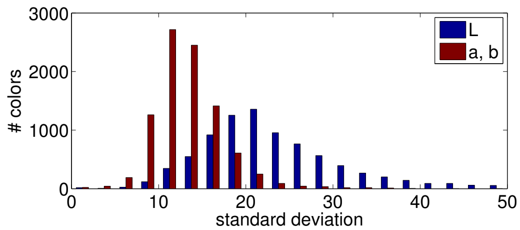


Figure 10: Histogram of the standard deviations of the Gaussian curve around the color centers. A color name's spread in CIELAB is approximately twice as large in the L direction than in the two chromatic directions.

that the spread in the L direction is approximately twice as large as in the a and b directions.

This is an intuitive result when looking at basic color names such as *red* or *green*, because they are hue names and allow for more variation along the lightness axis. Our large scale analysis shows that this is not restricted to basic color names but a general trend for all the 9000 color names studied.

Conclusions

We present a color thesaurus with color values for over 9000 color names in ten different languages, namely Chinese, English, French, German, Italian, Japanese, Korean, Portuguese, Russian, and Spanish, respectively. Unlike previous large-scale experiments [10, 12], we do not carry out labour intensive psychophysical experiments, but use a statistical framework [9] that can be used to extract color values for a given semantic expression [8].

Instead of just estimating a single color value for a given color name, we compute a distribution that indicates how significant each point in CIELAB is for that expression. This is important, because a color name is not precisely defined, but covers a specific volume in the color space.

The vagueness of color names is demonstrated at the example of *maroon*. Different databases report varying values that have ΔE distances of up to 49 between each other. Another source of imprecision is that a translation does not exactly match with the original expression. For instance, the translations of *maroon* have subtle nuances that are reflected in the color estimations.

A more advanced data analysis shows that higher z values implicate a higher accuracy of the estimated color values, which justifies the usage of statistical tests for color naming. Further, we show that colors are rather oriented along the L axis than the chromatic axes, indicating that color names tend to be lightness invariant.

Website: `http://ivrg.epfl.ch/ColorThesaurus`

Acknowledgements

We thank the following people for their remarkable help to translate the color names from English to their native languages (in alphabetical order of the language name): Zhengren Li (Chinese), Hannah Muckenhirn (French), Céline Heldner (German), Alexis Kessel (Italian), Ken Larpin (Japanese), Hyun-Kyu Kim (Korean), Mélanie Francisco (Portuguese), Alexandra Khlebnikova (Russian), Myriam Zapico (Spanish).

References

[1] XKCD color survey, http://blog.xkcd.com/2010/05/03/color-survey-results/, last checked Sept. 2012.

[2] B. Berlin and P. Kay. *Basic Color Terms: Their Universality and Evolution*. University of California Press, 1969.

[3] R. Brown and E. Lenneberg. A study in language and cognition. *Journal of Abnormal and Social Psychology*, 49(3):454–462, 1954.

[4] Color Database. http://www.perbang.dk, last checked Sept. 2012.

[5] J. Davidoff, I. Davies, and D. Roberson. Colour categories in a stone-age tribe. *Nature*, 398:203–204, 1999.

[6] Google Custom Search API. https://developers.google.com/custom-search/docs/xml_results, last checked Sept. 2012.

[7] M. J. Huiskes and M. S. Lew. The MIR flickr retrieval evaluation. In *ACM MIR*, 2008.

[8] A. Lindner, N. Bonnier, and S. Süsstrunk. What is the color of chocolate? - extracting color values of semantic expressions. In *IS&T CGIV*, 2012.

[9] A. Lindner, A. Shaji, N. Bonnier, and S. Süsstrunk. Joint statistical analysis of images and keywords with applications in semantic image enhancement. *ACM MM*, 2012.

[10] N. Moroney. http://www.hpl.hp.com/personal/nathan_moroney/, last checked Sept. 2012.

[11] N. Moroney. Unconstrained web-based color naming experiment. In *IS&T EI*, volume 5008, Color Imaging VIII: Processing, Hardcopy, and Applications, pages 36–46, 2003.

[12] D. Mylonas, L. MacDonald, and S. Wuerger. Towards an Online Color Naming Model. In *IS&T CIC*, pages 140–144, 2010.

[13] D. Sekulovski, G. Geleijnse, B. Kater, J. Korst, S. Pauws, and R. Clout. Enriching text with images and colored light. In *IS&T EI*, volume 6820, Multimedia Content Access: Algorithms and Systems II, 2008.

[14] J. van de Weijer, C. Schmid, J. Verbeek, and D. Larlus. Learning color names for real-world applications. *TIP*, 18(7):1512 – 1523, 2009.

[15] W3C Recommendation, World Wide Web Consortium. *CSS Color Module Level 3*, 7 June 2011.

[16] I. H. Witten, E. Frank, and M. A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*, volume 3. Morgan Kaufmann, 2011.

[17] X11 color names. http://cvsweb.xfree86.org/cvsweb/*checkout*/xc/programs/rgb/rgb.txt?rev=1.1, last checked Sept. 2012.

Author Biography

Albrecht Lindner is a PhD student in the School of Computer and Communication Sciences at EPFL (Switzerland) under the supervision of Prof. Sabine Süsstrunk. He works on large-scale statistical methods for semantic image understanding and enhancement. His research is sponsored by Océ Print Logic Technologies. In 2008, Albrecht obtained his MS degrees in Electrical Engineering from the University of Stuttgart (Germany) and Télécom Paristech (France).