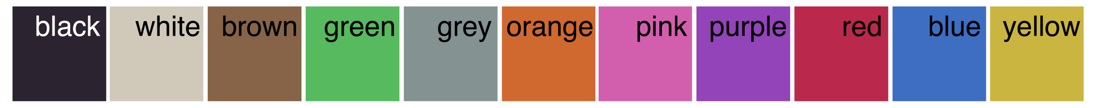
Color Reduction Based on Human Categorical Perception

Robert Laganière, Di Pang, and Ahmad Al-Kabbany
laganier@eecs.uottawa.ca, dpang097@uottawa.ca, alkabbany@ieee.org
VIVA Lab, School of Electrical Engineering and Computer Science, Canada
Arab Academy for Science and Technology, Department of Electronics and Communications Engineering, Egypt

1 Problem Statement and Motivation

- Color is a low-level feature that is used in several applications — detection, recognition, retrieval, and tracking of objects.
- Several perceptual studies [1] have shown that categorical perception plays an important role in the process of color discrimination and color memorization.
- Inspired by the human categorization in color theory [2, 3, 4], this paper proposes a lookup table for compact representation of color spaces, which reduces the space of color to 11 categories.
- A reduced color space will be particularly suitable to content-based image retrieval and classification tasks.

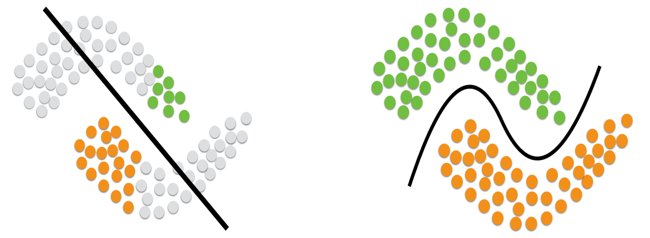


2 Previous Work

- Given the 11 color names, chip-based methods [5, 6, 7] rely on a controlled experimental setup in which the color category of particular color chips is decided by human subjects. Based on the distance from those "anchor chips", the category of the rest of triplets in a color coordinate system can be determined.
- Those chip-based methods perform poorly, though, as they sample the RGB color coordinate system only sparsely.
- To overcome such drawbacks, the techniques of [3, 8] adopt a modified version of probabilistic latent semantic analysis (PLSA) to learn color names from real-world images queried from Google, taking into account the color label of queried images.
- In [9], the authors started with an off-the-shelf color naming. Relying on learned human prioritization of colors on natural images, they assigned a dominant and an associative color name to every image region. Hence, they boosted the pixel-level color naming to the image/region level.

3 Learning by Transduction

- Learning by transduction exploits both labelled and unlabelled data points to infer a discriminative model.
- This is particularly useful if the labelled points are scarce and if the data points are well-separated in the feature space.



(a) (b)

- For discretization, graph Laplacian methods are adopted. They are based on a discrete approximation of the s-weighted Laplacian operator [10].
- In these methods, a graph with nodes representing the data points is constructed, where the edge weights are induced by a kernel and represent the similarities between the data points.
- The discrete approximation for the original optimization problem is given by

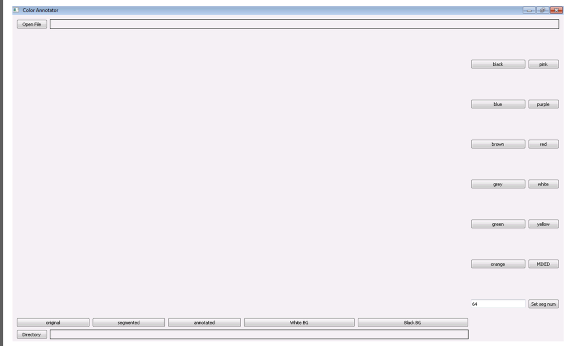
$$\min_{F \in \mathbb{R}^n} \; (F-Y)^T C (F-Y) + F^T L\, F, \qquad (1)$$

- The n-dimensional vector F can then be obtained by solving the linear system given by

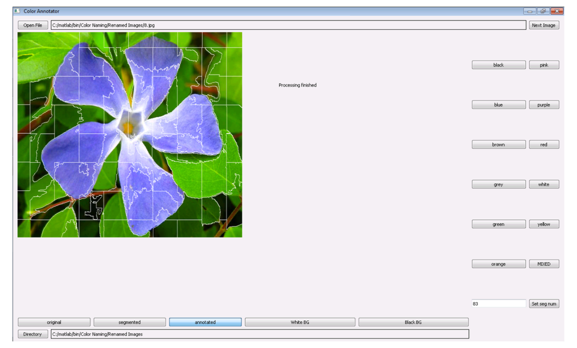
$$(L+C)\, F = C\, Y. \qquad (2)$$

4 Proposed Method - Main Idea and Label Gathering Experiment

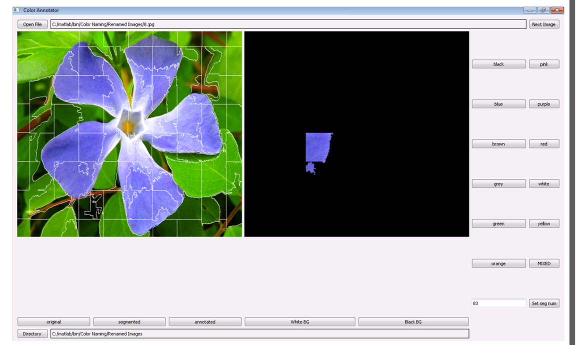
- We propose a semi-supervised learning-based scheme for color reduction using graph Laplacian-based transduction [11].
- Color names (labels) for sparse triplets in the RGB cube are first provided by a group of 31 volunteers (tested for vision problems) to enforce a perceptually relevant labelling. Experiments were carried out in an average office environment with monitors equipped with 16 million-color graphics cards.



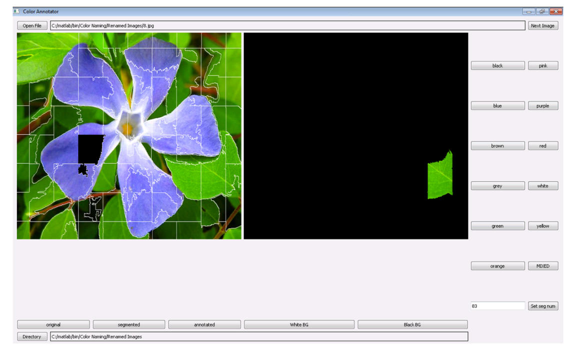
The Annotator's GUI



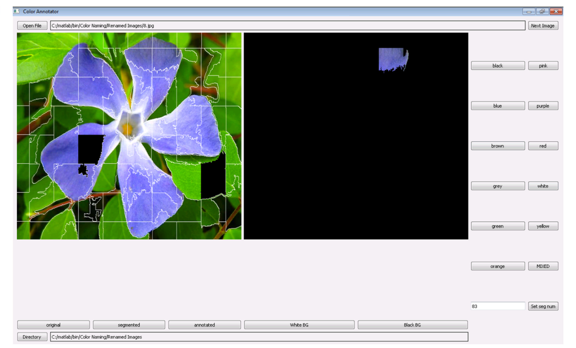
Loading and Segmenting Image



Choosing a Super-pixel (SP)



Labeling 1st SP and Choosing 2nd SP

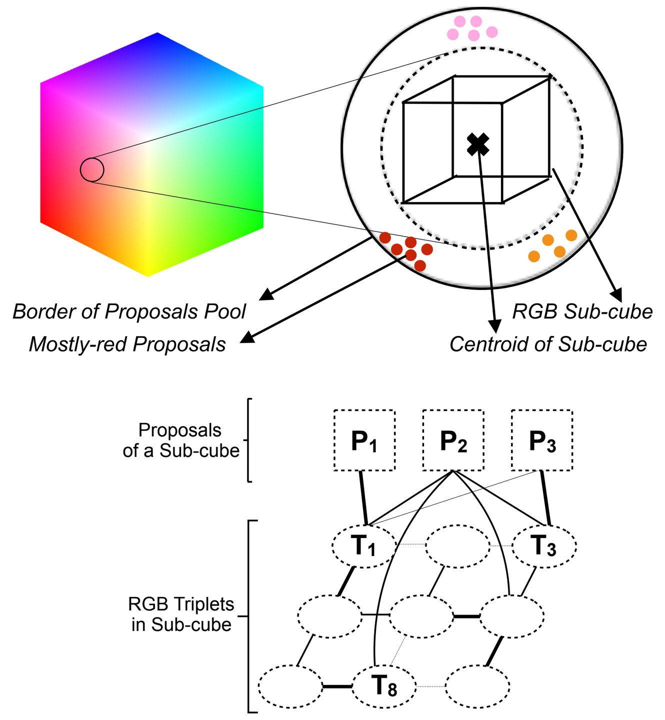


Labeling 2nd SP and Choosing 3rd SP

- Then, the acquired sparse labels are propagated to the whole RGB cube by means of the transduction of a graph whose nodes/vertices are all the triplets in the RGB cube.
- The output is a lookup table that maps every RGB triplet to a fuzzy color naming vector that indicates how likely a triplet is affiliated with each of the 11 basic color labels.
- Each volunteer was assigned three images to be labelled within a maximum of 20 minutes. Every image was shown as a set of super-pixels and the volunteers were instructed to give every super-pixel one of the 11 Basic Color Labels .
- Contrary to chip-based methods, we require the labelled triplets to sample the RGB cube efficiently. We chose images to be labeled by solving a knapsack problem.

5 Proposed Method - Label Propagation

- First, we compute a Normalized Word Count for every RGB triplet in the labelled images
- That count is a vector of 11 elements, each of which is the number of times an RGB was assigned a particular label, divided by total number of its occurrence in labeled images
- A triplet in labelled images might have existed several times, and might have been assigned different labels by different/same participants since the context of an image was found to influence color naming.
- Label propagation step is started by dividing the RGB cube into H sub-cubes, and computing a pool of proposals — the triplets not farther than a certain Euclidean distance away (in RGB space).
- Figure 2 depicts an example of an RGB sub-cube, its corresponding pool of proposals, and its dominant color names (red, pink, and orange).
- We loop over the sub-cubes, and over every dominant color name. For every loop, we construct a graph with nodes representing the RGB triplets in the sub-cube and the proposals affiliated with the color name of the current loop.



6 Proposed Method - Label Propagation [Cont.]

- Then, we solve a graph transduction problem by minimizing an objective function and solving the corresponding linear system given by eqn.(1) and eqn.(2).
- We get a vector of scores whose length equals to the number of triplets in the sub-cube, and whose elements indicate the degree of acceptance of the proposals by triplets in the sub-cube
- The entries of the Laplacian matrix of this graph are calculated using the kernel function given by

$$k(X_i, X_j) = \frac{\widetilde{k}(X_i, X_j)}{[\widetilde{d}(X_i)\ \widetilde{d}(X_j)]^{\lambda}}, \qquad (3)$$

7 Proposed Method - Label Propagation [Cont.]

- After obtaining the vector F, we define the color naming vectors of the triplets in the sub-cube by

$$\mathcal{C}_{ih} := \frac{1}{N} \sum_{n=1}^{N} F_n^s \times \mathcal{W}_n, \qquad (4)$$

A Bayesian Formulation for Color Naming

- For assigning a color name to a triplet, we have a likelihood term which is the color naming fuzzy vectors of the lookup table, and a prior term induced by a cooccurrence matrix, M. The color naming problem can be formulated in a Bayesian fashion as:

$$p(l|\mathcal{T}_i, L) \propto p(\mathcal{T}_i|l)p(l|L) \qquad (5)$$

8 Results

- We compared our results with 3 SOTA techniques [3, 8, 9] on 2 standard datasets (eBay and BSD500)
- Experiments were conducted on 2 color coordinate systems, RGB and CIEL*a*b, and the performance was found to be consistent across both systems.

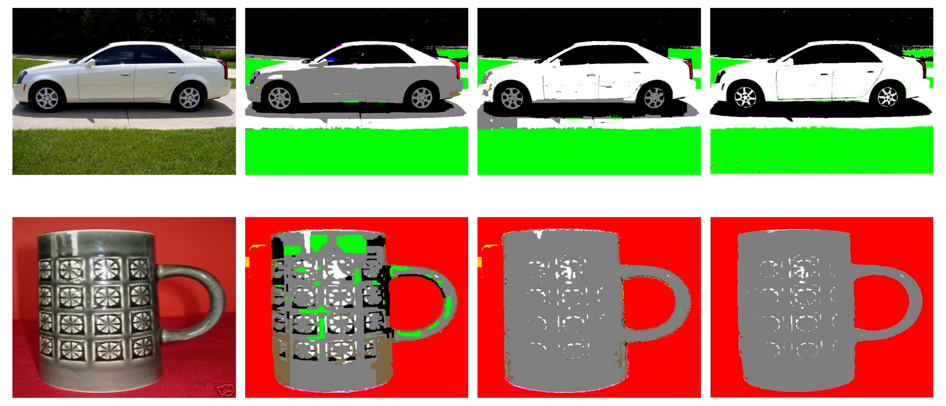


Fig. 6: Results of the proposed color naming method on the eBay dataset. From left to right: the original image, our graph cuts-smoothed CTI-Rk3 and CTR-Rk3 color maps, and the map of PLSA-bg* [3].

- We show results for CTI (no prior information) and CTR (taking prior information into consideration) approaches for color naming.
- We also show results for rank 1 (Rk1) through rank 3 (Rk3) classification.
- For visualization purposes, Rk3 results are shown as graph cuts-smoothed color maps.

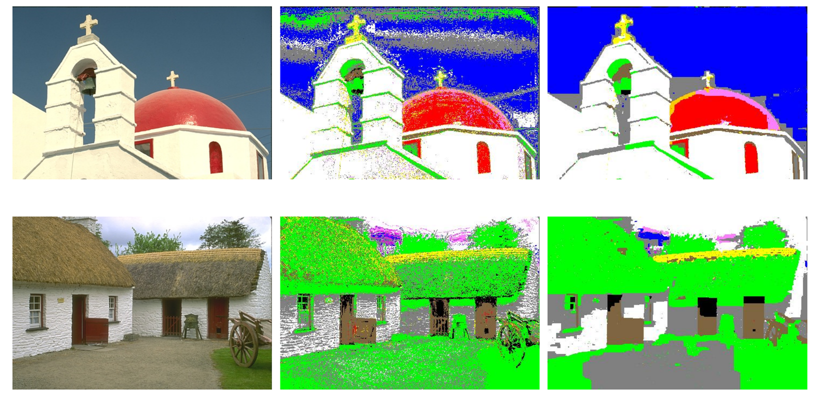


Fig. 7: Results of the proposed method on the BSD500 segmentation dataset. From left to right: original image, CTI-Rk1 and graph-cuts-smoothed CTI-Rk3 maps.

Table 1: Comparison of pixel annotation scores of the methods proposed in [8] and [9] with our method. The third and the fourth sections of the table show the results of annotating pixels with and without considering the image label prior. For the two methods, we show the score for rank 1 (Rk1), rank 2 (Rk2), and rank 3 (Rk3) labelings.

Method	Cars	Shoes	Dresses	Pottery	Overall
PLSA-ind [8]	56	77	80	70	70.6
PLSA-reg [8]	74	94	85	82	83.4
DA [9]	63	88	85	79	78.8
CTI-Rk1	51.1	64	69.2	57.2	60.4
CTI-Rk2	70.7	81.7	86.5	77.5	79.1
CTI-Rk3	78.6	88.2	91.3	85.4	85.9
CTR-Rk1	88.1	94	94.9	92.3	**92.3**
CTR-Rk2	88.6	94.2	95	92.7	92.6
CTR-Rk3	89.1	94.4	95.1	93.4	93

9 References

1. Keiji Uchikawa and Hiroyuki Shinoda, "Influence of basic color categories on color memory discrimination," Color Research and Application , vol. 21, no. 6, pp. 430–439, 1996.
2. E.L. van den Broek, Th.E. Schouten, and P.M.F. Kisters, "Modeling human color categorization," Pattern recognition letters , vol. 29, no. 8, pp. 1136–1144, June 2008.
3. J. van de Weijer, C. Schmid, and J. Verbeek, "Learning color names from real-world images," in IEEE TIP , June 2007, pp. 1–8.
4. Brent Berlin and Paul Kay, Basic Color Terms: their Universality and Evolution , University of California Press, Berkeley and Los Angeles, 1969.
5. M. Vanrell R. Benavente and R. Bladrich, "A data set for fuzzy colour naming," COLOR Res. Appl. , vol. 31, no. 1, pp. 48–56, 2006.
6. K. Kelly and D. Judd, Color: Universal Language and Dictionary of Names , Number 440. MD: National Bureau of Standards, Gaithersburg, 1976.
7. G. Menegaz, A. Le Troter, J. Sequeira, and J. M. Boi, "A discrete model for color naming," EURASIP Journal on Advances in Signal Processing , vol. 2007, no. 1, pp. 029125, 2006.
8. J. van de Weijer, C. Schmid, J. Verbeek, and D. Larlus, "Learning color names for real-world applications," IEEE TIP , vol. 18, no. 7, pp. 1512–1523, July 2009.
9. Y. Liu, Y. Liang, Z. Yuan, and N. Zheng, "Learning to describe color composition of visual objects," in ICPR, 2012.
10. O. Duchenne, J.-Y. Audibert, R. Keriven, J. Ponce, and F. Segonne, "Segmentation by transduction," in CVPR, 2008.
11. Mikhail Belkin and Partha Niyogi, "Semi-supervised learning on Riemannian manifolds," Machine Learning, vol. 56, no. 1-3, pp. 209 – 239, 2004.