Names and Shades of Color for Intrinsic Image Estimation: Supplementary Material

Marc Serra, Olivier Penacchio, Robert Benavente, and Maria Vanrell
Computer Vision Center / Computer Science Dept.
Universitat Autònoma de Barcelona
Campus UAB, Building O, 08193 Bellaterra (Barcelona), Spain
{ mserra, penacchio, robert, maria }@cvc.uab.cat

1. Introduction

In this supplementary material we provide some details about the discretization of the color-name descriptor explained in Section 2.1 of the paper "Names and Shades of Color for Intrinsic Image Estimation". We also present some additional results that could not be included in the paper due to a space constraints: we show an example of the performance of our algorithm on a natural image and provide, for each image of the MIT dataset, the error values for the shading and the reflectance estimates, as well as the global error.

2. Color-name descriptor discretization

Given an image, the output of the color-name descriptor consists of 11-dimensional arrays containing in each position the probability of a given RGB value to belong to each of the 11 universal color classes defined by Berlin and Kay in their anthropological work [2]. These vectors satisfy the two following properties:

- Their values are positive real numbers.
- Their values sum to 1.

In order to make the problem computationally efficient, we need to reduce the number of possible labels while preserving the second property. Thus, we discretize the set of possible values, making the non-zero values to lay within the set {0.25, 0.5, 0.75, 1} and imposing their sum to be 1 .

This leads to 4 kinds of labels, each of which can be seen as a permutation with repetition of n elements, where the first element is repeated a times, the second b rimes, the third c times, etc. Therefore, we can calculate the number of possible labels of each kind using the following formula:

$$P_n^{a,b,c\dots} = \frac{n!}{a!b!c!\dots}, n = a+b+c+\dots \tag{1}$$

These are the 4 kinds of labels we can find:

- **Labels containing a single non-zero value (1):** Permutations of eleven elements, where the first element, 0, is repeated 10 times, and the second element, 1, just appears once.

$$P_{11}^{10,1} = \frac{11!}{10!1!} = 11$$

- **Labels containing two equal non-zero values (0.5, 0.5):** Permutations of eleven elements, where the first element, 0, is repeated 9 times, and the second element, 0.5, twice.

$$P_{11}^{9,2} = \frac{11!}{9!2!} = 55$$

- **Labels containing two different non-zero values (0.25, 0.75):** Permutations of eleven elements, where the first element, 0, is repeated 9 times, and the second (0.25) and third (0.75) elements, once.

$$P_{11}^{9,1,1} = \frac{11!}{9!1!1!} = 110$$

- **Labels containing three non-zero values (0.5, 0.25, 0.25):** Permutations of eleven elements, where the first element, 0, is repeated 8 times, the second element, 0.5, once, and the third element, 0.25, twice.

$$P_{11}^{8,2,1} = \frac{11!}{8!2!1!} = 495$$

Although theoretically it gives us 671 possible labels, most of them are never found in practice. This makes sense because in the color-naming model we use, [1], no color border is shared by more than 3 or 4 colors. Therefore, many labels, such as the one defining a color as having probability 1/2 of being red and 1/2 of being green, never come to happen. Thus, only considering labels with up to three positive coordinates is enough to accurately describe the whole RGB space. In the end, only 250 different labels are actually used.

3. Results

In this section we provide some additional results. In Figure 1, we show an example of how our method works with natural images, using an image which has been previously used by other authors [3, 4].

We also provide the individual error values for each image on the MIT dataset with the four measures we use in section 5 of the paper. For each image, we provide its error on the shading and the reflectance images, as well as the global averaged error. Boldface values in the tables correspond to the mean error values presented in Tables 1 and 2 of the paper.

References

[1] R. Benavente, M. Vanrell, and R. Baldrich. Parametric fuzzy sets for automatic color naming. *Journal of the Optical Society of America A*, 25(10):2582–2593, 2008.

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

[3] A. Bousseau, S. Paris, and F. Durand. User assisted intrinsic images. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2009)*, 28(5):130:1–130:10, 2009.

[4] L. Shen and C. Yeo. Intrinsic images decomposition using a local and global sparse representation of reflectance. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 697–704, 2011.

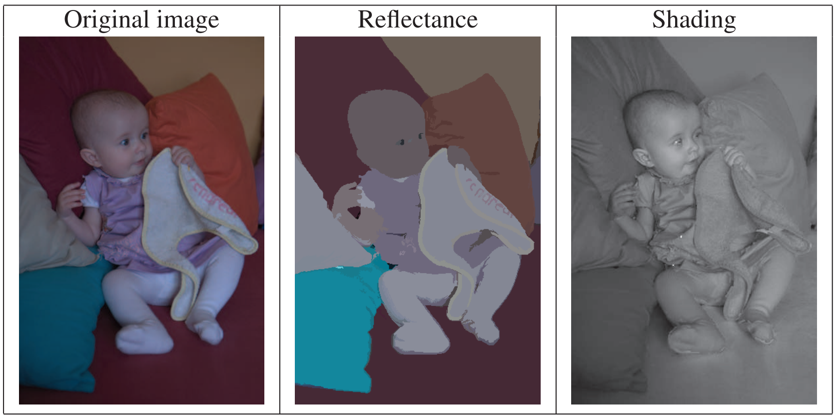

Figure 1: Shading and reflectance images recovered by our approach from a real image.

	MSE			Correlation		
Image	Shading	Reflectance	Global	Shading	Reflectance	Global
box	0,1749	0,2534	0,2142	0,8188	0,5567	0,6878
cup1	0,0044	0,0321	0,0183	0,9934	0,8675	0,9305
cup2	0,0683	0,1001	0,0842	0,8644	0,6934	0,7789
deer	0,0983	0,1099	0,1041	0,7313	0,8271	0,7792
dinosaur	0,0823	0,0411	0,0617	0,8145	0,9512	0,8829
frog1	0,1781	0,2051	0,1916	0,7094	0,7005	0,7050
frog2	0,1326	0,1480	0,1403	0,7217	0,2263	0,4740
panther	0,0055	0,0130	0,0092	0,9860	0,8267	0,9064
paper1	0,0028	0,0034	0,0031	0,9922	0,8868	0,9395
paper2	0,0064	0,1122	0,0593	0,9848	0,2639	0,6243
raccoon	0,0047	0,0043	0,0045	0,9931	0,9443	0,9687
squirrel	0,1674	0,1876	0,1775	0,6346	0,8061	0,7204
sun	0,0250	0,0148	0,0199	0,9153	0,9594	0,9374
teabag1	0,0864	0,0745	0,0804	0,4646	0,8574	0,6610
teabag2	0,0440	0,0491	0,0465	0,8222	0,9448	0,8835
turtle	0,1966	0,0412	0,1189	0,7238	0,6802	0,7020
apple	0,1587	0,1600	0,1593	0,8469	0,6785	0,7627
pear	0,0865	0,1158	0,1011	0,8213	0,0315	0,4264
phone	0,0575	0,0195	0,0385	0,8566	0,9742	0,9154
potato	0,0409	0,0379	0,0394	0,8781	-0,0204	0,4288
Mean 16 obj.	0,0799	0,0869	**0,0834**	0,8231	0,7495	**0,7863**
Mean 20 obj.	0,0811	0,0862	**0,0836**	0,8287	0,6828	**0,7557**

Table 1: MSE and correlation results on the reduced (16 objects) and full (20 objects) MIT datasets.

	LMSE			aLMSE		
Image	Shading	Reflectance	Global	Shading	Reflectance	Global
box	0,0314	0,0714	0,0514	0,3764	0,1693	0,2728
cup1	0,0019	0,0159	0,0089	0,1179	0,4285	0,2732
cup2	0,0492	0,0326	0,0409	0,4420	0,5026	0,4723
deer	0,0559	0,0467	0,0513	0,5306	0,3658	0,4482
dinosaur	0,0402	0,0201	0,0302	0,3055	0,1165	0,2110
frog1	0,0439	0,0815	0,0627	0,2854	0,5330	0,4092
frog2	0,0479	0,0390	0,0434	0,2364	0,5910	0,4137
panther	0,0029	0,0035	0,0032	0,0509	0,0829	0,0669
paper1	0,0021	0,0022	0,0021	0,0721	0,1238	0,0980
paper2	0,0038	0,0203	0,0121	0,1584	0,1609	0,1597
raccoon	0,0035	0,0035	0,0035	0,0383	0,1170	0,0776
squirrel	0,0804	0,0848	0,0826	0,6089	0,4593	0,5341
sun	0,0088	0,0035	0,0062	0,1539	0,0575	0,1057
teabag1	0,0650	0,0523	0,0587	0,6490	0,2138	0,4314
teabag2	0,0329	0,0355	0,0342	0,3013	0,1772	0,2393
turtle	0,0773	0,0278	0,0525	0,2084	0,8307	0,5196
apple	0,0202	0,0199	0,0200	0,3582	0,9178	0,6380
pear	0,0144	0,0171	0,0157	0,5782	0,7957	0,6869
phone	0,0096	0,0080	0,0088	0,1560	0,0714	0,1137
potato	0,0239	0,0202	0,0221	0,4882	0,9979	0,7430
Mean 16 obj.	0,0342	0,0338	**0,0340**	0,2835	0,3081	**0,2958**
Mean 20 obj.	0,0308	0,0303	**0,0305**	0,3058	0,3856	**0,3457**

Table 2: LMSE and aLMSE results on the reduced (16 objects) and full (20 objects) MIT datasets.