An Overview of Color Name Applications in Computer Vision

Joost van de Weijer[1](✉) and Fahad Shahbaz Khan[2]

[1] Computer Vision Center Barcelona, Edifici O, Campus UAB, Bellaterra 08193, Spain
joost@cvc.uab.es

[2] Computer Vision Laboratory, Linköping University, Linköping, Sweden

Abstract. In this article we provide an overview of color name applications in computer vision. Color names are linguistic labels which humans use to communicate color. Computational color naming learns a mapping from pixels values to color names. In recent years color names have been applied to a wide variety of computer vision applications, including image classification, object recognition, texture classification, visual tracking and action recognition. Here we provide an overview of these results which show that in general color names outperform photometric invariants as a color representation.

Keywords: Color features · Color names · Object recognition

1 Introduction

Color is one of the important characteristics of materials in the world around us. As such it is one of the important features for computer vision systems in their task to understand visual data. Its description however is complicated due to many scene accidental events such as unknown illuminant, presence of shadows and specularities, unknown acquisition system and image compression. As a result many researchers ignored color and only extracted information from the luminance channel. However, it has been shown that for many applications, ranging from image retrieval and object recognition to visual tracking and texture recognition, color description is crucial for obtaining state-of-the-art results.

There exist two main methodologies to the color description problem. The first methodology is based on reflection models which describe the interaction of light, material and sensors [5,6,8]. From these reflection models photometric invariant descriptions of the material color can be derived. Given certain assumptions these descriptors can overcome the dependence of the color description on scene accidental events. Examples are color descriptions which are invariant to illuminant color, shadow-shading and specularities[4,19,22]. The main advantage of these methods is that they do not need training data and therefore do not require a laborious and costly labeling phase. The main drawback of these

A. Trémeau et al. (Eds.): CCIW 2015, LNCS 9016, pp. 16–22, 2015.
DOI: 10.1007/978-3-319-15979-9_2

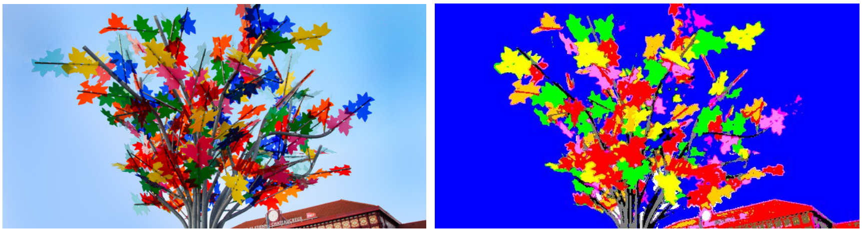

Fig. 1. Example of pixelwise color name annotation. The color names are represented by their corresponding color.

methods is that the assumptions on which they are based (for example white illumination, known acquisition device, etc) limit their application. Typically they require high-quality images without compression artifacts, and are not very effective for the medium quality images which are currently used in the many large scale data sets which have been collected from the internet.

The second methodology to color description is based on color names. Color names are words that refer to a specific color and are used by humans to communicate colors. Examples of color names are 'blue', 'crimson' and 'amber'. Humans use color names routinely and seemingly without effort. They have been primarily studied in the fields of visual psychology, anthropology and linguistics [7]. Basic color terms have been studied in the influential work of Berlin and Kay [2]. They are defined as those color names in a language which are applied to diverse classes of objects, whose meaning is not subsumable under one of the other basic color terms, and which are used consistently and with consensus by most speakers of the language. Basic color names were found to be shared between languages. However the number of basic terms varies from two in some indigenous languages to twelve in for example Russian. Most computer vision works, and also this paper, consider the eleven basic color terms of the English language: black, blue, brown, grey, green, orange, pink, purple, red, white, and yellow.

Computational color naming[1,18,23] aims to learn a mapping from pixel values to color name labels (see Fig. 1). A clear example in computer vision where color names are desired is within the context of image retrieval, where a user might want to query for images with "blue sofas". The system recognizes the color name "blue", and orders the retrieved results on "sofa" based on their resemblance to the human usage of "blue'. Later research showed that color names actually also constitute an excellent color descriptor. They where found to be robust to photometric variations, while having in general higher discriminative power than the photometric invariants.

In recent years, the two approaches to color description, namely, the physics-based and the color name methods, have been compared on a wide variety of computer vision applications. Constantly, color names where found to outperform the physics-based approaches by a significant margin. Color names have been extensively tested in image classification tasks [11,13], object recogni-

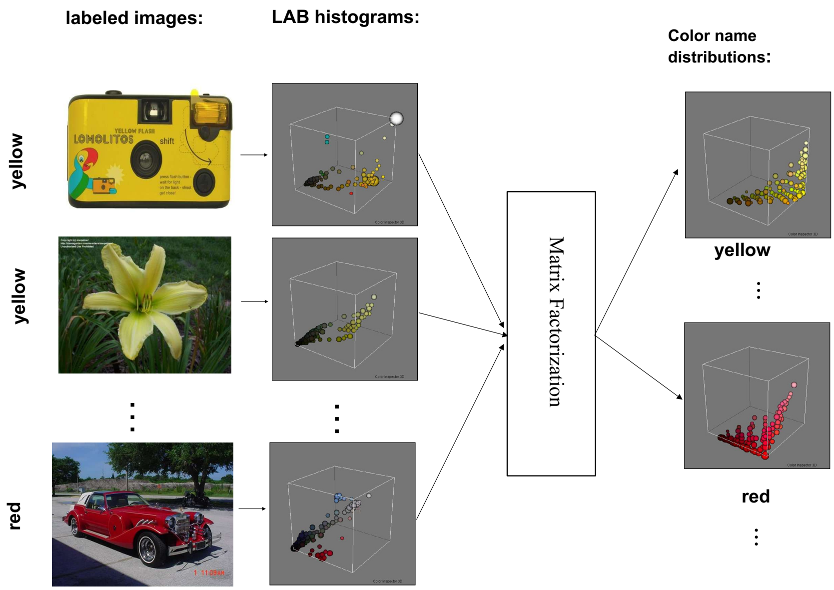

Fig. 2. Overview of approach to learn color names presented in [23]

tion [12], and action recognition [9]. Similar results where reported for texture classification[10], visual tracking [3] and person reidentification [24]. For image retrieval results for color names are reported in [16][25] and recently it was used for improved illumination estimation[20]. The main reason for this success is the high discriminative power which color names possess, while being robust to photometric variations in images.

In this paper, we first outline the different approaches which exist to computational color naming. In section 3 we present an overview of the results we obtained when comparing color names to other color representations. After which we discuss conclusion and future outlook for color name research in section 4.

2 Color Names

There are two main approaches to computational color naming, the main difference is the nature of their training data, either being calibrated or uncalibrated. In a calibrated setup, multiple subjects are asked to label hundreds of color chips within a well-defined experimental setup [15,17]. The colors are to be chosen from a preselected set of color names (predominantly the set of 11 basic color terms). Examples of color name mappings based on calibrated data are the work of Mojsilovic [18] and Benavente et al. [1]. The latter proposed a parametric method to model color names. They introduce a fuzzy model which is fitted to data obtained from psychophysical experimental data. This data is well calibrated, meaning that the color samples are presented in a controlled laboratory environment under stable viewing conditions with known illuminant. These studies are very relevant within the fields of linguistics and color science. However, for applications in computer vision which are often based on uncalibrated data, these mappings learned under perfect circumstances were often found to underperform.

In [23], we proposed a different approach to learning color name mappings, which is based on uncalibrated data obtained from 'Google image'. For this data the illuminant, acquisition system and amount of compression are unknown. An overview of this method is provided in Fig. 2. The images from 'Google images' are represented as LAB histograms after which they are joined in one large data matrix. An adapted probabilistic latent semantic analysis (PLSA) is applied to factorize the matrix into mixture coefficients and topic distributions. Here the topic distributions are the probabilities of colors to occur for a certain color name. An advantage of this method, over the calibrated mappings described above, is that it is more robust to scene accidental events, such as slight illuminant changes, shadows, image compression, etc. Therefore this approach is more popular for computer vision applications, and in the next section we provide an overview of its usage.

It is interesting to interpret color names in terms of photometric invariance theory which was developed by physics-based research to color. Color names typically are elongated along the intensity axes (or achromatic axis) of the color space, and more compact in the hue direction. As a result color names typically describe a group of colors which have similar hue but vary significantly in saturation and luminance. The hue is known to be a photometric invariant with respect to shadow, shading and specularities. Since the color names typically group colors of equal hue it can be said to be photometrically robust. Color names however significantly differ from photometric invariants along the achromatic axes. Here photometric invariance do not differ between light and dark colors and are often instable for dark and achromatic colors. On the other hand there are three color names, black, grey and white, which allow to distinguish between these sections of the color space. As a result, color names do not have the drop in discriminative power along the achromatic axis which is observed for photometric invariants [14].

3 Color Names Applications in Computer Vision

Here we provide an overview of the experiments over the last couple of years, where we have used color names and compared them to photometric invariants.

As discussed above, the main advantage of color names is that they maintain the discriminative power while possessing a certain degree of photometric robustness. In addition, they yield a very compact color representation of only eleven dimensions. This also compares favorable with respect to photometric invariants. For example the hue and opponent angle invariants of [21] are 36 dimensions. The popular color SIFTS [19] perform the SIFT operation on the

Table 1. Comparison of color-name results versus baseline (either luminance or standard RGB) and versus the best reported photometric invariant. Results are provided for several applications and different data sets. The final column indicates the performance measure which is used. The results show that color names consistently outperform the photometric invariants.

applications	data set	baseline	phot. inv.	color names	measure
image class.	flower-17 [11]	69.0	87.0	**88.0**	class. rate
	birds-200 [13]	12.9	14.0	**17.0**	class. rate
object detection	cartoons [12]	27.6	35.3	**41.7**	mAP
	pascal 2007 [12]	32.2	30.6	**34.8**	mAP
action recognition	willow[9]	66.6	67.2	**68**	mAP
	Pascal 2010 [9]	55	54.8	**56.3**	mAP
	standford-40 [9]	39	38.7	**39.8**	mAP
texture	KTH-TIPS-2a [10]	55.5	54.3	**56.8**	class. rate
	KTH-TIPS-2b [10]	42.1	**45.4**	44.2	class. rate
	FMD [10]	20.3	22.2	**25.6**	class. rate
	Texture-10 [10]	52.3	52.7	**56.0**	class. rate
visual tracking	OTB [3]	54.5	57.6	**74.0**	distance prec.

separate color channels of colorspaces. As a result the dimensionality increment is either of 128 or 256 dimensions.

When incorporating color into a computer vision application, one has to decide on the color feature to use, and on how to combine the shape and color information. Standard approaches include early and late fusion methods, but especially for image classification it was found that more complex fusion approaches can significantly improve the overall results [13] [11]. The main idea behind these more complex fusion methods is that they aim to 'bind' the color and shape information locally.

Our results of using color names for various application domains have been summarized in Table 1. The results are compared to a baseline performance, which is either the results based on luminance only or on RGB (if luminance results are not reported). In addition, the best results reported with photometric invariants for these data sets are given. One can observe that for all applications the best results are obtained with color names. It is also interesting to note that photometric invariants do not always outperform the baseline. The highest performance increases are reported for object detection and visual tracking.

Next to our efforts to evaluate color name performance several other research groups have resulted similar conclusions. A recent paper, which proposes to use 16 color names (fuchsia, blue, aqua, lime, yellow, red, purple, navy, teal, green, olive, maroon, black, gray, silver and white) for person re-identification obtains excellent results without using uncalibrated images to train. Most probably this is due to the fact that colors are grouped based on their distance to the color centers which represent these color names. As a result they do not obtain the very compact color distributions for the achromatic colors which are typical for color mapping learned from calibrated data. In addition Zheng et al. [25] report excellent image retrieval results based on color names.

4 Discussion and Future Research Directions

We have summarized recent evaluation results on color descriptors. Results on various computer vision applications, including image classification, object recognition, texture classification, visual tracking and action recognition, show that color names outperform color descriptors based on photometric invariance.

Several future research directions can be considered to further improve color representations. In a recent paper [14], we show that there is a third approach to color description. This method directly optimizes the discriminative power of the color representation given a classification problem. For the same dimensionality as the color names (11 dimensions) this method reported slightly inferior results than color names. However, for higher dimensions, this method obtained better results on various data sets. This method could be further improved by learning from larger data sets.

Also the approach of Zheng et al. [25], which extends the set of color names to 16 color names, could be further investigated. Analysis of the optimal number of color names, and the correct probabilistic representations of these overlapping color name sets should be considered. Finally, applying recent advances in convolutional neural networks (deep learning) to the problem of discriminative color representations learning is also expected to improve results.

Acknowledgments. This work is funded by the Project TIN2013-41751 of Spanish Ministry of Science and the Catalan project 2014 SGR 221.

References

1. Benavente, R., Vanrell, M., Baldrich, R.: Parametric fuzzy sets for automatic color naming. Journal of the Optical Society of America A **25**(10), 2582–2593 (2008)
2. Berlin, B., Kay, P.: Basic color terms: their universality and evolution. University of California, Berkeley (1969)
3. Danelljan, M., Shahbaz Khan, F., Felsberg, M., Van de Weijer, J.: Adaptive color attributes for real-time visual tracking. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2014. IEEE (2014)
4. Finlayson, G.D., Schiele, B., Crowley, J.L.: Comprehensive colour image normalization. In: Burkhardt, H.-J., Neumann, B. (eds.) ECCV 1998. LNCS, vol. 1406, p. 475. Springer, Heidelberg (1998)
5. Geusebroek, J., van den Boomgaard, R., Smeulders, A., Geerts, H.: Color invariance **23**(12), 1338–1350 (2001)
6. Gevers, T., Smeulders, A.: Color based object recognition. Pattern Recognition **32**, 453–464 (1999)
7. Hardin, C., Maffi, L. (eds.): Color Categories in Thought and Language. University Press, Cambridge (1997)
8. Healey, G.: Segmenting images using normalized color. IEEE Trans. Syst., Man, Cybern. **22**, 64–73 (1992)

9. Khan, F.S., Anwer, R.M., van de Weijer, J., Bagdanov, A.D., Lopez, A.M., Felsberg, M.: Coloring action recognition in still images. International Journal of Computer Vision **105**(3), 205–221 (2013)
10. Khan, F.S., Anwer, R.M., van de Weijer, J., Felsberg, M., Laaksonen, J.: Compact color-texture description for texture classification. Pattern Recognition Letters **51**, 16–22 (2015)
11. Khan, F.S., van de Weijer, J., Vanrell, M.: Modulating shape features by color attention for object recognition. International Journal of Computer Vision (IJCV) **98**(1), 49–64 (2012). http://www.cat.uab.cat/Public/Publications/2012/SVV2012
12. Khan, F., Anwer, R., van de Weijer, J., Bagdanov, A., Vanrell, M., Lopez, A.: Color attributes for object detection. In: IEEE Conference on Computer Vision and Patter Recognition (2012)
13. Khan, F., Van de Weijer, J., Bagdanov, A., Vanrell, M.: Portmanteau vocabularies for multi-cue image representation. In: Twenty-Fifth Annual Conference on Neural Information Processing Systems (NIPS 2011) (2011)
14. Khan, R., Van de Weijer, J., Khan, F.S., Muselet, D., Ducottet, C., Barat, C.: Discriminative color descriptors. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2866–2873. IEEE (2013)
15. Lammens, J.: A computational model of color perception and color naming. Ph.D. thesis, Univ. of Buffalo (1994)
16. Liu, Y., Zhang, D., Lu, G., Ma, W.Y.: Region-based image retrieval with high-level semantic color names. In: Proceedings of the 11th International Multimedia Modelling Conference, MMM 2005, pp. 180–187. IEEE (2005)
17. Menegaz, G., Troter, A.L., Sequeira, J., Boi, J.M.: A discrete model for color naming. EURASIP Journal on Advances in Signal Processing 2007 (2007)
18. Mojsilovic, A.: A computational model for color naming and describing color composition of images. IEEE Transactions on Image Processing **14**(5), 690–699 (2005)
19. van de Sande, K.E.A., Gevers, T., Snoek, C.G.M.: Evaluating color descriptors for object and scene recognition. PAMI **32**(9), 1582–1596 (2010)
20. Vazquez-Corral, J., Vanrell, M., Baldrich, R., Tous, F.: Color constancy by category correlation. IEEE Transactions on Image Processing **21**(4), 1997–2007 (2012)
21. van de Weijer, J., Schmid, C.: Coloring local feature extraction. In: Proc. of the European Conference on Computer Vision **2**, pp. 334–348. Graz, Austria (2006)
22. van de Weijer, J., Schmid, C.: Applying color names to image description. In: IEEE International Conference on Image Processing (ICIP). San Antonio, USA (2007). http://www.cat.uab.cat/Public/Publications/2007/VaS2007a
23. van de Weijer, J., Schmid, C., Verbeek, J., Larlus, D.: Learning color names for real-world applications. IEEE Transactions on Image Processing **18**(7), 1512–1524 (2009). http://www.cat.uab.cat/Public/Publications/2009/VSV2009
24. Yang, Y., Yang, J., Yan, J., Liao, S., Yi, D., Li, S.Z.: Salient color names for person re-identification. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part I. LNCS, vol. 8689, pp. 536–551. Springer, Heidelberg (2014)
25. Zheng, L., Wang, S., Liu, Z., Tian, Q.: Packing and padding: Coupled multi-index for accurate image retrieval. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2014. IEEE (2014)

http://www.springer.com/978-3-319-15978-2

Computational Color Imaging
5th International Workshop, CCIW 2015, Saint Etienne, France, March 24-26, 2015, Proceedings
Trémeau, A.; Schettini, R.; Tominaga, S. (Eds.)
2015, XVII, 244 p. 133 illus., Softcover
ISBN: 978-3-319-15978-2