Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 45 (2017) 70–81

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Clustering colors

Action editor: Dietmar Heinke

Igor Douven

Sciences, Normes, Décision (CNRS), Paris-Sorbonne University, France

Received 2 August 2016; received in revised form 10 February 2017; accepted 8 May 2017
Available online 25 May 2017

Abstract

Regier, Kay, and Khetarpal report the results of computer simulations that cluster color stimuli on the basis of their coordinates in CIELAB space, one of two commonly used perceptual color spaces. Regier and coauthors find partitions of those stimuli that are strikingly similar to the way actual color lexicons partition color space. They do not argue for the custom-made clustering method used in their simulations, nor for the assumption of CIELAB space. The present paper aims to answer the question to what extent their computational results depend on these assumptions. It does this by applying a great variety of known clustering methods to Regier et al.'s stimuli, and by assuming not only CIELAB space but also CIELUV space, the other main color space.

Keywords: Categorization; Color; Cluster analysis; Computational modeling; Optimization

1. Introduction

There is an ongoing debate about the metaphysical status of color categories. According to some theorists, color categories are organized around a set of universal focal colors (Berlin & Kay, 1969/1999), while other theorists hold that color categories are culturally relative, grounded in linguistic conventions (Roberson, 2005; Roberson, Davies, & Davidoff, 2000). Evidence from linguistic anthropology is mixed, showing universal tendencies in color categorization across languages, but also deviations from those tendencies (Berlin & Kay, 1969/1999; Cook, Kay, & Regier, 2005; Lindsey & Brown, 2009).

Jameson and D'Andrade (1997) put forward the idea that both the universal tendencies and the deviations might be due to the interaction of a cognitively motivated preference for informative categorizations with a perceptual color space that is irregularly shaped and as a result can be carved up into partitions with different degrees of informativeness. Both elements in this explanation could be argued to be culture-independent—the first possibly being anchored in our innate cognitive makeup, the second in our perceptual apparatus—yet their supposed interaction may leave some room for cultural influences in categorization, given that different categorizations can achieve roughly the same high level of informativeness. Thereby, Jameson and D'Andrade offer an interesting middle ground between the "universalist" and "relativist" positions vis-à-vis color categorization.

Jameson and D'Andrade's proposal has been tested computationally by Regier, Kay, and Khetarpal (2007), who implemented the notion of informative categorization in a clustering method and applied this, in simulations, to the stimuli used for the anthropological studies cited above. These simulations yielded partitions of color space strikingly similar to lexical partitions of that space in many languages from around the world, a finding that supports Jameson and D'Andrade proposed explanation of the presence of universal tendencies in color categorization. The simulations also showed, however, that rather different

E-mail address: igor.douven@paris-sorbonne.fr

http://dx.doi.org/10.1016/j.cogsys.2017.05.004
1389-0417/

looking partitions can be more or less equally informative (in the relevant sense), thus supporting Jameson and D'Andrade's explanation of the registered differences in color categorization.

Jameson and D'Andrade understand the notion of informative categorization in terms of similarity and dissimilarity. Specifically, on their definition, a categorization is more informative the more similar items within the same category are to each other and the more dissimilar items across categories are to each other. Regier et al.'s (2007) clustering method formalizes this principle in a straightforward way. It is to be noted, however, that by now a bewildering variety of clustering methods is available, all of which can be said to explicate the same broad idea of jointly maximizing within-cluster similarity (or within-similarity, for short) and across-cluster dissimilarity (across-dissimilarity) in slightly different, yet seemingly equally legitimate, ways. Regier and colleagues do not compare their custom-built clustering method with any of the clustering methods that are available off the shelf, nor do they explain why they use their particular method rather than any of the others.

This raises a number of questions. First, have Regier and colleagues given us a genuinely new clustering method, or does it boil down to one we already had, in the sense that the two methods will always yield the same clustering outputs? And if the former, to what extent do Regier et al.'s results in the color domain depend on the particularities of their method? Might we be able to improve upon those results by using any of the other methods, in the sense that other methods might yield categorizations of color space that more accurately reflect categories in actual use by humans? We address these questions by applying a great variety of clustering methods to the stimuli of Regier and colleagues' study and comparing the outputs to the relevant anthropological color-naming data.

There is another dimension along which we would like to generalize Regier, Kay, and Khetarpal's results. These researchers apply their clustering method to color stimuli as specified by their coordinates in CIELAB space, which is one of two standard and widely used perceptual color spaces, the other one being CIELUV space. Despite the similarity of these spaces, differences between them are not so minute that they can be guaranteed to have had no noteworthy effect on Regier and colleagues' results. We investigated what effect (if any) the assumption of CIELAB space may have had by carrying out our computations separately for each of the two color spaces.

2. Theoretical background

The data for the earlier-cited anthropological studies were gathered by asking native speakers of various unwritten languages to name the color of each of the 330 Munsell chips shown in Fig. 1 (henceforth often simply referred to as "the chips"), which consist of 320 chromatic chips and 10 achromatic chips, the latter ranging from black to white through various shades of gray. The columns in Fig. 1 represent equally spaced Munsell hues, and the rows represent levels of value; the chromatic chips are all at the maximum saturation available for their hue–value combination.

These chips also served as the items on which Regier et al. (2007) ran their computer simulations. More exactly, they used the chips as specified by their CIELAB coordinates. As mentioned, CIELAB space—or CIE 1976 $L^*a^*b^*$ space, as it is known more officially—is one of the two most commonly used perceptual color spaces; CIE 1976 $L^*u^*v^*$ space (or "CIELUV space") is the other commonly used one. Both spaces are intended to be perceptually uniform in that distances within them are meant to indicate perceived degrees of similarity among colors.[1] Most color researchers consider these spaces to be fairly successful in achieving their aim, though it is known that neither space is perfect (see Fairchild, 2013, Ch. 10).

One gets an impression of both the differences between these spaces and the irregularity of each by considering the locations of the Munsell chips in them, shown in Fig. 2.[2] Looking at this figure, one understands why color space is sometimes described as "a spindle" (e.g., Gärdenfors, 2000, 10 f). However, given either space, that spindle is not completely symmetric. Far from it, in fact: the figure clearly brings out a large bump in the yellow/green[3] region as well as bumps in the purple/blue and red regions, in both CIELAB and CIELUV space. (Here it is important to recall that the chromatic chips are all at maximum saturation so that they lie on the surfaces of the spaces.)

If color space were perfectly symmetric, any rotation of a partition of that space would be as informative (in Jameson and D'Andrade's sense) or well-formed (as Regier and colleagues call it) as any other rotation. But it is precisely because of the noted irregularities that, in principle, one partition can be more informative than all others. In Regier et al. (2007), this claim is made precise with the help of the following definition of well-formedness: Let P be some partition of the chips shown in Fig. 1. Then the function S,

$$S(P) = \sum_{x,y:\mathrm{cat}_P(x)=\mathrm{cat}_P(y)} \mathrm{sim}(x, y),$$

[1] This is what makes these spaces *perceptual* color spaces, in contrast to, for instance, the RGB and CMYK spaces, which serve different purposes. See Malacara (2002, Ch. 6), for a mathematical specification of the CIELUV and CIELAB spaces and for an explanation of how they formally relate to each other as well as to other well-known color spaces.

[2] The CIELAB coordinates of the 330 chips are available at the WCS website, http://www1.icsi.berkeley.edu/wcs/. To obtain the corresponding CIELUV coordinates, we used the ColorConvert function of *Mathematica* 10.

[3] For interpretation of color in Fig. 2, the reader is referred to the web version of this article.