Understanding color trends by means of non-monotone utility functions

Mohammad GHADERI[a,1], Francisco J. RUIZ[b] and Núria AGELL[a]
[a]*ESADE Business School- Ramon Llull University, Sant Cugat, Spain*
[b]*Automatic Control Department, BarcelonaTech. Vilanova i la Geltrú, Spain*

Abstract. In this paper we explore the possibility of capturing color trends and understanding the rationale behind the popularity of a color. To this end, we propose using a preference disaggregation approach from the field of Multi-Criteria Decision Analysis. The main objective is to identify the criteria aggregation model that underlies the global preference of a color. We introduce a new disaggregation method based on the well-known UTASTAR algorithm able to represent preferences by means of non-monotonic utility functions. The method is applied to a large database of ranked colors, from three different years, based on the information published on the webpage of an international creative community. Non-monotone marginal utility functions from each of the coordinates are obtained for each year. These functions contain the color preference information captured, in an understandable way.

Keywords. Multi-Criteria Decision Analysis, Disaggregation preference method, Non-monotonic utility function. Color trends.

Introduction

Color is one of the key features which play an important role on the purchase decisions of consumers. Fashionable colors depend on many uncontrolled factors related to the nature of the product, the target market and other environmental characteristics such as cultural, religious and even climatic variables. Color trends are ephemeral and prevail just in one season, thus, it is crucial for the industry to understand the color fashion trends to offer the product to the market in the most efficient way.

The common practice for the forecasting of color trends in industry are based on the opinion of field experts, which is hard to be substituted by analytical models. In this paper, we explore the option to capture color trends and understand the rationale behind the popularity of each color. To this end, we propose to use a preference disaggregation approach from the field of Multi-Criteria Decision Analysis (MCDA). The aim of this approach is to identify the criteria aggregation model that underlies the global preference in a multi-criteria decision problem by means of the marginal utility function of each of the attributes considered.

UTA (Utilités Additives) is one of the most representative preference disaggregation methods. In most of the fields where UTA and its extensions have been applied, the input attributes are normally expected to be monotone with respect to the preferences. The assumption of monotonicity is widely accepted as reasonable for

[1] [a]*ESADE Business School- Ramon Llull University, Sant Cugat, Spain*

many criteria, such as price, risk level, security, safety, comfort, required time, etc. However, this is not the case for other attributes. For instance, whether a color is preferable or not may depend on the red/green coordinate (if we use CIELab coordinates) but it is not expected that this attribute was monotone. This fact has motivated us to propose an extension for UTA method to address non-monotone preferences. Several attempts have been made in the literature to overcome the aforementioned shortcoming. This paper contributes to the existing literature by introducing a faster and simpler method able to capture the preferential system of the DM in the form of marginal additive non-monotonic utility functions. The method is applied to a database of colors obtained from an international creative community. These functions contain the color preference information captured, in an understandable way.

The paper is organized as follows: first, an introduction to the problem of color preference description is briefly introduced and the basics of color spaces are presented. Section 2 is devoted to the description of the methodology proposed. Experiment description and results corresponding to the application of the proposed method to the color database are presented in section 3. Finally, in the last section conclusions and future work are discussed.

1. Color coordinates and color spaces

Color preferences are the tendency for an individual or a group to prefer some colors over others. People make associations with certain colors due to their past experiences. For an individual, colors associated to good experiences are preferred and colors associated to bad experiences are disliked. For a group, color preferences can be influenced by many global factors include among others: cultural, politics, religion, economy, climate and geography factors.

Designers and manufacturers desire to know what the “in” colors are going to be before their products can be developed. To this end, it will be useful to understand how some colors attributes influence color preferences. Color attributes are normally referred as color coordinates and the space formed by all possible colors is denoted color space.

Several numeric specifications for colors definition can be found in the literature. The most classic and internationally accepted of these are the ones based on tristimulus values or coordinates. The most known of these is RGB, proposed by the *Commission International de l’Eclairage* (CIE) in 1931. RGB uses additive color mixing, because it describes what kind of light (red, green or blue) needs to be emitted to produce a given color. The RGB color model is implemented in different ways, depending on the capabilities of the system used. By far the most common is the 24-bit implementation. This model is thus limited to a range of 256×256×256 ≈ 16.7 million colors. It is a convenient color model for computer graphics, but it can be unintuitive in use. The specification of a desired color can be difficult for untrained people (try selecting brown using an RGB vector).

In 1976, the CIE proposed the CIE Lab color scale as an attempt to linearize the perceptibility of color differences [6]. CIE Lab (CIELab) is the most complete color model used conventionally to describe colors visible to the human eye. Its three parameters represent the luminance (L) of the color, its position between red and green

(a) and its position between yellow and blue (b). It is generally argued that CIELab is more intuitive than RGB and its coordinates L,a,b are more readily and easily recognized. For this reason, in this paper we'll use the CIELab color coordinates representation. Figure 1 represents geometrically the two most color representation systems RGB and CIELab.

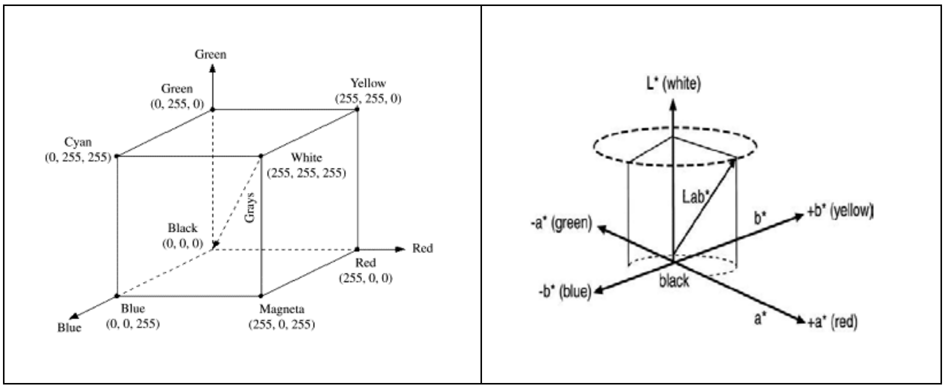

Figure 1. RGB (left) and CIE L*a*b* (right) coordinates.

2. The preference disaggregation methodology proposed

Most of machine learning tools can automatically discover relations between attributes and preferences, but, usually, they are considered as a black-box in the sense that this relation is difficult to understand in a rational way. In this study we are interested in representing this relation in an understandable way. For this reason, we will use a preference disaggregation method from MCDA capable of identifying the criteria aggregation model that underlies the preference result from the analysis of the global preferences.

UTA (Utilités Additives) is one of the most representative preference disaggregation methods. It was first introduced by Jacquet-Lagrèze and Siskos as a Linear Programming (LP) model to capture the preferential system of the Decision Maker (DM) through nonlinear (piecewise linear) monotonic additive utility functions. The aim of the UTA method is to reproduce the ranking made by the DM over the set of alternatives by minimizing the level of ranking errors. Ranking errors are generally defined as the distance between the utility levels of two consecutive alternatives that are ranked incorrectly. However, the definition of the error slightly differs in the specific variant of UTA. The method leads to a simple Linear Programming (LP) model where the optimal solution can be easily obtained.

Several extensions of UTA method have been introduced in the MCDA literature since then, incorporating variations on the original algorithm and considering different forms of global preference and optimality criteria. In this paper we propose an extension for UTA method to address non-monotone preferences.

In the following subsections, we present the most representative UTA method for ranking (UTASTAR) and the non-monotone variant proposed to address the problem of color preferences.

2.1 UTASTAR method

Suppose that there are m criteria $g_1, g_2,\ldots,g_m$ to assess N preordered set of alternatives $a_1,a_2,\ldots,a_N$ (in which a_1 is the most and a_N is the least preferred alternative in the ranking list) and x_i^n is the performance of the alternative a_n over the criteria g_i. Given a preordering set of the alternatives by the DM, the aim of the UTASTAR algorithm is to extract and represent the underlying logic behind the given ranking through estimating a set of m monotonic and additive utility functions, as consistent as possible with the preferential system of the DM. The formulation of the UTASTAR method involves defining α_i breakpoints and henceforth $\alpha_i - 1$ subintervals $[g_i^0, g_i^1]$, $[g_i^1, g_i^2]$, …,$[g_i^{\alpha i-2}, g_i^{\alpha i-1}]$ on the i^{th} criterion, in which g_i^0 and $g_i^{\alpha i-1}$ are the minimum and maximum values over the i^{th} scale, respectively. The marginal value at a breakpoint g_i^1 on criterion i is expressed as in equation (1).

$$u_i(g_i^l) = \sum_{j=1}^{l}(u_i(g_i^j) - u_i(g_i^{j-1})) = \sum_{j=1}^{l} w_{ij} \tag{1}$$

and the marginal value for an alternative a_n whose performance on the i^{th} scale is $x_i^n \in [g_i^l, g_i^{l+1}]$ is obtained by linear interpolation between $u_i(g_i^l)$ and $u_i(g_i^{l+1})$, as follows:

$$u_i(x_i^n) = \sum_{j=1}^{l} w_{ij} + \frac{x_i^n - g_i^l}{g_i^{l+1} - g_i^l} w_{i,l+1} \tag{2}$$

The global utility of an alternative a_n is obtained by the sum of all of the marginal utilities, as in equation (3).

$$U(a_n) = \sum_{i=1}^{m} u_i(x_i^n) \tag{3}$$

The linear programming problem by the UTASTAR is provided in (4).

$$\begin{aligned}
&min\ z = \textstyle\sum_{n=1}^{N}(\sigma^+(a_n) + \sigma^-(a_n)) \\
&\text{subject to} \\
&U'(a_n) - U'(a_{n+1}) \geq \delta \text{ iff } a_n \succ a_{n+1}, \forall n = 1, 2, ..., N-1 \\
&U'(a_n) - U'(a_{n+1}) = 0 \text{ iff } a_n \sim a_{n+1}, \forall n = 1, 2, ..., N-1 \\
&\textstyle\sum_{i=1}^{m}\sum_{j=1}^{\alpha_i - 1} w_{ij} = 1 \\
&U'(a_n) = U(a_n) - \sigma^+(a_n) + \sigma^-(a_n) \\
&w_{ij}, \sigma^+(a_n), \sigma^-(a_n) \geq 0, \forall i, j, n
\end{aligned} \tag{4}$$

in which $\sigma^+(a_n)$ and $\sigma^-(a_n)$ are the underestimation and overestimation error terms. The term $\sigma^+(a_n)$ (respectively $\sigma^-(a_n)$) is the lowest amount that must be deducted from (added to) the estimated global utility of a_n to satisfy the DM preferential order over a_n

and a_{n+1}. The term δ is a parameter with a small value, and the first two constraints represent the preorder relations provided by DM. The third constant ensures that the relative weights of the criteria sum up to 1, and the objective function minimizes the deviation of the utility function proposed by the model and the one assumed as the tacit knowledge of the DM. By solving this model, the marginal utility function over each criterion scale will be constructed based on the expression in (1).

2.2 Non-monotonic UTA-based Algorithm

The input attributes in UTASTAR method are normally expected to be monotone with respect to the preferences. However this is not a reasonable requirement for colorimetric components. Obviously no one can expect a monotonic relationship between color preference degree and its degree of greenness, or its luminance. Therefore, modification of UTASTAR algorithm in the sense that it will be able to represent preferential system of the DM by non-monotonic utility functions is of a great importance in this setting.

Although several attempts have been conducted in the literature to overcome the mentioned shortcoming [2],[3],[4],[5], all are computationally intensive or require extra information from the DM. The method we applied here, inspired by the UTA methodology, is fast and tractable, unlike the existing ones. The general idea is to relax sign constraint in the decision variables representing difference of utility level between two consecutive breakpoints. Therefore, marginal utility function can change the monotonicity at any breakpoint. This might leads to two problems: the first one is the overfitting problem in the case that the monotonicity changes many times. This is prevented by defining a small, but reasonable, number of breakpoints. The second problem is about normalization. By the normalization, we mean that the minimum and maximum global utility must be equal to zero and one, respectively. The challenge is that we cannot predict where the maximum utility will be achieved in order to impose a constraint on the sum of them over the set of criteria. Furthermore, we do not know the attribute level corresponding to the minimum marginal utility on each criterion to set them equal to zero. To solve this problem, an iterative approach is followed. Whenever the maximum global utility is less than one, it's value is forced to increase in the next iteration, by adding a new constraint considering the performance level corresponds to highest marginal utility in the current stage. The added constraint is applied just in the next iteration, and will be removed from the LP model in the next iterations, because it is not necessarily satisfied in the final solution. Whenever the maximum global utility is greater than one, a restrictive constraint is imposed to ensure that the global utility of the attribute levels corresponding to the highest marginal utility in the current stage will not have a value more than one in all the next iterations. Furthermore, to satisfy another condition of normalization which is minimum global utility being zero, a penalization term is added to the objective function to penalize any violation of this assumption.

Results obtained by applying this method, which is able to represent preferential system of the DM by set of additive non-monotonic utility functions, on the color dataset are presented in the next section.

3. Experiment description

The purpose of this study is to be able to represent the color preference in an understandable way. This representation can be useful when a color should be chosen for a new product. To this end, a database of color preferences has been collected and the non-monotone method presented is applied. The following subsections presents de database used and the results of the experiment.

3.1. Color Database

Data collected in this research, expressing the collective color preference, were obtained from the website *www.colourlovers.com*. It is managed by an international creative community that focuses on color inspiration and color trends for both personal and professional creative projects. Each year since 2010, community members vote their favorite color created by themselves during the previous year. In this way, votes can be taken as a preference measure of each color.

Colors, represented by their RGB colorimetric components, along with its number of votes were collected for three different years, 2010, 2011, and 2012. Colorimetric components are transformed into CIELab color space. There are no simple formulas for conversion between RGB and Lab, because RGB is device-dependent. Then, RGB is first transformed to a specific absolute color space using CIE standard illuminant D50 and then it is transformed into CIELab. In order to increase the validity of our findings, we exclude colors with less than 100 votes from the dataset. Finally, for each year, colors are ranked according to their votes. In conclusion, we obtain a dataset with 114 colors each one of them represented by its year, two pairs of coordinates and with an output representing its ranking order in the specific year considered.

3.2. Experimental results

The algorithm was run on a 64 bit OS and 2.53 GHz Intel Core2Duo using MATLAB R2012b. The extracted marginal utility functions obtained in CIELab color spaces for each of the three years are provided in the figure 2. Each row of this table of figures represent the non-monotone marginal utility function for attributes L, a and b. Each graph represents the influence of one of the attributes on the global color preference of the year.

The obtained graphs demonstrate how the difference of attractiveness among set of colors, perceived by a set users, can be represented by the color coordinates. Concretely, if color A is perceived more popular than color B based on the collective preferences of users, the figure explains the reason of that. Furthermore, it provides the contribution of each color coordinate into explaining the underlying preference structure.

From figure 1, it can be seen that the shapes of the marginal utility function of the first attribute is approximately the same in all the three different datasets. Generally speaking, as the value of the luminance increases, the marginal utility value first decreases, then increases, and decreases again at the end. Therefore we can conclude that the marginal utility function over the attribute L (luminance) has an S-shape function, while the general shape for the other two marginal utility functions are not very clear and considerably differs in the different years.

Table 1 shows the weights of each colorimetric dimension in the CIELab color space for each year

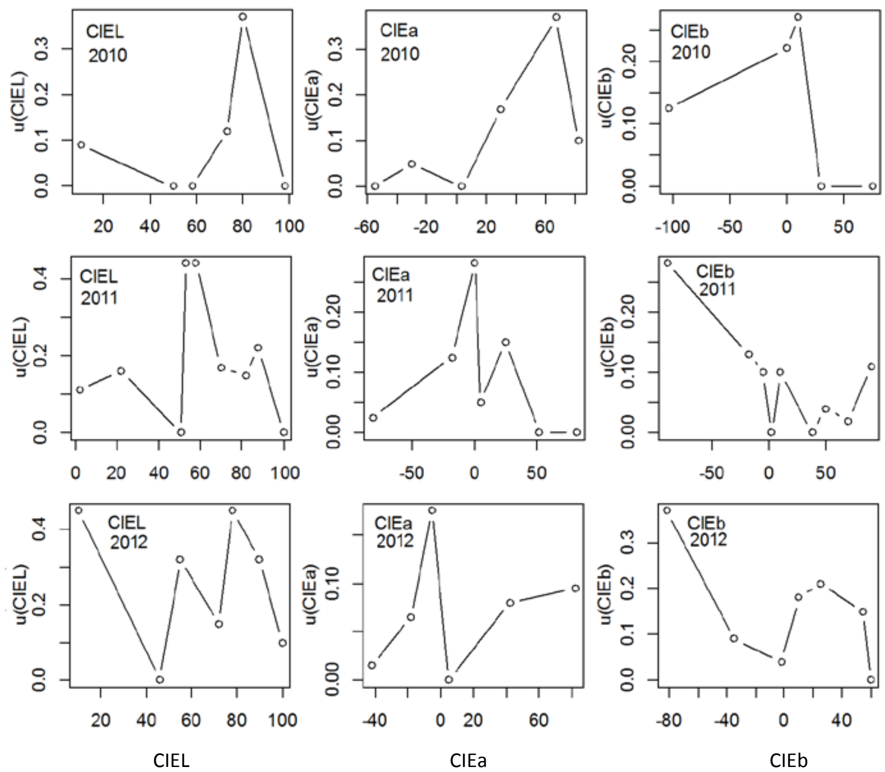

Figure 2. Marginal utility functions obtained in the CIELab color space

Table 1 Weights of each colorimetric dimension in the CIELab color space for each year

	CIEL	CIEa	CIEb
2010	37.00%	36.00%	27.00%
2011	44.10%	27.60%	28.30%
2012	44.50%	17.90%	37.60%

As can be seen, L dimension is always the most important dimension in the CIELab space. It has the highest weight in all the three rows, supporting the idea that attractiveness of a color mostly depends on its luminance. In another words, luminance (L) plays the most important role in determining the extent to which the color is going to be perceived favorable, therefore should be considered more carefully than other two attributes.

To evaluate performance of the learning algorithm, the accuracy of the results is calculated by comparing the ranking achieved by estimated utility and the ranking achieved by the voters' opinion. The result is as follows.

Table 2 Accuracy of the results

	CIELab
2010	68.18%
2011	50.80%
2012	74.20%

Given the marginal utility functions, it is also possible to detect the color that provides maximum utility to the DMs. The following table provides numerical values of those colors, along with their visualization.

Table 3 Most favorable colors based on the extracted utility functions in the CIELab space

	L*	a*	b*	
2010	80.2	67.6	9.7	
2011	58.2	-0.6	-93.7	
2012	79	-5.6	-83.1	

It is clear that the three colors should not be necessarily the same, as the set of the voters differs in the three years. But the interesting thing is that all the three colors have pretty close value of the dimension luminance, while values for the other two dimensions differs a lot in the three rows.

4. Conclusions and future research

This paper presents a methodology, based on a non-monotonic UTA algorithm, to capture color trends and understand the rationale behind the popularity of colors. An experiment has been performed to analyze if the presented algorithm can provide some insights on strict preference underlying color popularity. Results show that the luminance is the dimension which mostly affects color preference, it always dominates the other two dimensions red/blue, green/yellow. The results also shed light on the general shape of the marginal utility function of luminance, the most important dimension, which is shown to be S-shape.

In further research we will try to analyze the dynamics of these trends by taking into account the sequence of marginal utility functions to forecast the influence of each of the color attributes in future preferences.

Acknowledgments

This work is supported by the Spanish project SENSORIAL (TIN2010-20966-C02-02) Spanish Ministry of Education and Science. Mohammad Ghaderi has partially been supported by SUR (Secretaria d'Universitats i Recerca) of the DEC (Departament d'Economia i Coneixement) of the Government of Catalonia (2014FI_V00637).

References

[1] Jacquet-Lagrez, E., Siskos, Y., Assessing a set of additive utility functions for multicriteria decision making: The UTA method, *European Journal of Operational Research 10, 151-164* (1982).

[2] Despotis, D. K., Zopounidis, C., Building additive utilities in the presence of non-monotonic preferences, *Advances in Multicriteria Analysis, 5, 101-114* (1995).

[3] Kliegr T., UTA-NM: Explaining Stated Preferences with Additive Non-Monotonic Utility Functions, *Preference Learning (PL-09) ECML/PKDD-09 workshop* (2009).

[4] Eckhardt, A., Klieger, T., Preprocessing Algorithm for Handling Non-Monotone Attributes in the UTA method, *Preference Learning: problems and applications in AI (PL-12) workshop* (2012).

[5] Doumpos, M., Learning non-monotonic additive value functions for multicriteria decision making, *OR* Spectrum, 34, 89-106 (2012).

[6] F. Billveyer, Principles of Color Technology, 2n edition Jonh Wiley & Sons, New York (1981).