Dominant Colors as Image Content Descriptors: A Study with Users

Soraia M. Alarcão Ruben Pavão Manuel J. Fonseca
LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
salarcao@lasige.di.fc.ul.pt, rubenpavao@gmail.com, mjfonseca@ciencias.ulisboa.pt

***Abstract*—Image content are typically described using low level features such as color, texture, shape, or a combination of the previous. A particular use of color is the identification of the dominant colors in images to describe its content, for image retrieval, for instance. In this paper, we present a study with users to verify if the dominant colors can be used as image content descriptors. From the study we identified the dominant and the search colors users associated to a set of images. We supplemented this information with gaze coordinates, collected with an affordable eye tracker, to register the regions at which people looked while identifying colors in the images. The analysis of the data revealed that users used a small set of color names, and that the colors used for searching were similar to those considered dominant, validating the use of dominant colors as image descriptors. As a result of the study, we make available a dataset of 100 images annotated with their dominant colors, the colors that users would use to search for them, and the areas where they looked while identifying both types of colors.**

I. Introduction

Color is one of the most distinctive visual features. Various systems for exploring, searching and presenting images to users take advantage of it through the use of image's dominant colors. Although there are mechanisms to search for or explore images through their dominant colors, these are usually identified from the perspective of the system and making several assumptions (e.g. more importance to the center, salient objects, etc.) and not based on the human perception of colors. That is, typically the dominant color is the one that occupies the largest area of the image. However, from the point of view of people, the dominant colors are not always those that cover more pixels. Additionally, most works consider too many colors as possible dominant colors, making their naming almost impossible, when users may want/need to explore or retrieve images by specifying the colors names.

The aim of this paper is to investigate whether dominant colors can be used as image content descriptors, and whether there is a relation between the regions at which people look and the colors they identify. To that end, we conducted a study with 40 participants in our research lab.

In particular, we designed two setups, one where we asked participants to identify up to three dominant colors in an image, and another where we asked them to mention up to three colors they will use to search for the presented image. Additionally, we collected eye tracking data of the regions of the image at which users looked while identifying colors.

DOI Reference Number: 10.18293/VLSS2018-034

From the data collected, we found that the colors used for searching are similar to those considered dominant, which means that we can develop a retrieval system where images are described by their dominant colors. In terms of the gaze information, we can conclude that there is no strong correlation between the regions at which people looked and the colors they identified. In most cases users looked at one region (e.g. faces) and mentioned a color that is presented in another region (e.g. t-shirt).

Our contributions are: 1) the confirmation of the JNS 11 colors as a valid reduced set of colors; 2) a dataset of 100 images annotated with their dominant colors and search colors identified by people; 3) gaze coordinates of users while identifying dominant and search colors in images.

II. Background and Related Work

In this section, we provide some background about color perception, color naming, dominant colors, and the use of eye-tracking to identify where people look at in images.

A. Color Perception

Color is the perceptual phenomenon related to the spectral characteristics of the electromagnetic radiation in the visible wavelengths (approximately from 380-750 nm). As suggested by human visual perception research [1] color is considered a pre-attentive property known to attract our visual attention above and beyond other object properties such as shape.

Our vision starts on the eye retina with two types of photoreceptors that receive the light stimulus and emit electrical impulses. Rods are responsible to operate at low light levels (scotopic vision), while cones operate at higher light levels (photopic vision). Cones are the ones responsible for the color vision, having a high spatial acuity. These electrical signals are then processed in the cortex, with our previously accumulated visual experience (memory), to form representations (Visual Perception) of color, shape, movement etc.

As so, we can say that color is the result of interpretation in the brain of the perception of light in the human eye and our visual memory.

B. Color Naming

In everyday life, we mainly identify colors by their names, which requires a general color vocabulary that is far from being precise. Given the importance of color naming, a variety of models and studies describing how people associate names and colors were introduced. Berlin and Kay studied the color naming behavior with subjects from multiple languages [2]. They concluded that the basic color terms in a culture can be predicted by the number of color terms the culture has. For English, they identified the following 11 basic terms: black, white, red, green, yellow, blue, brown, pink, orange, purple, and gray. Mojsilovic *et al.* presented a computational model for color categorization and naming of the 11 basic colors plus beige and olive [3].

Weijer *et al.* used real-world images to learn the 11 basic colors [4]. Moroney *et al.* conducted an unconstrained web-based study where they identified the 20 most commonly used color terms: green, blue, purple, red, pink, light, lime, dark blue, brown, yellow, black, orange, sky, bright, violet, olive, navy, sea, teal, and royal [5]. Menegaz *et al.* proposed a discrete model for color naming, where each of the 11 basic color terms was modeled as a fuzzy set [6]. Benavente *et al.* presented a parametric model for automatic color naming, where each of the 11 basic color terms was modeled as a fuzzy set with a parametric membership function [7].

As we can see, various authors adopted the set of 11 colors proposed by Berlin and Kay, probably because it is considered to contain colors that can be named by all cultures. Indeed, in 2000, Chang *et al.* coined it as the "Just Not the Same" colors (JNS), because any two colors from this set are not perceived as the same [8].

C. Dominant Colors

In general, color is a very distinctive feature, and as such several image search systems take advantage of it. In particular, they use the dominant colors of the images as a mechanism to describe and index their content. Usually, these systems rely mostly on color histograms to provide both the description of the colors present in an image and their quantities. Histograms are obtained by counting the number of pixels for each color, after quantizing the image colors into a reduced set of colors.

The VisualSEEk was one of the first systems for searching images using the dominant colors. It used the HSV color space to compute a histogram of 166 colors, from which it identified the dominant ones [9]. Deng *et al.* presented a feature descriptor that uses segmentation and color clustering to identify representative colors in each image's region [10]. Mojsilovic *et al.* proposed a method to compute dominant colors by considering both information captured through the image histogram and extracted from spatial relationships between frequently occurring colors [11].

Atsalakis *et al.* proposed the use of a neural network to automatically identify the significant colors with the minimum number of color classes [12]. Younnes *et al.* [13] and Amante *et al.* [14] proposed methods based on a fuzzy representation of colors to identify the dominant colors. Talib *et al.* proposed a method to reduce the background effect on the computation of dominant colors. Authors assigned weights to each dominant color in accordance with its belonging to the object or the background. The background colors, which are in contact with the image borders and out of salient object area, received a lower weight [15].

Although there are mechanisms for content-based image retrieval using dominant colors, most of them identify the dominant colors from the perspective of the system and not taking into consideration the human perception of colors.

D. Eye-tracking

Eye-tracking consists on cameras continuously tracking the position or orientation of the eyes [16]. Fixation consists on maintaining the visual gaze on a single location, and is useful to determine the focus of attention, i.e., to identify what triggered the attention change. Datasets of images annotated with eye-tracking information are important for the development of saliency models, i.e., to identify which information on an image attracts visual attention from the person looking at it.

In Table I, adapted from [17], we present some of the existing datasets available in the public domain (for detailed information, see [18], [17]). As far as we know, all of them contain eye tracking information but none is related to the tasks of looking at images while identifying the dominant colors or the colors to be used for searching.

Table I
DATASETS OF IMAGES ANNOTATED WITH EYE-TRACKING INFORMATION.

	Fixations		Inter-Fixation	Raw
	Locations	Durations	Durations	Data
DUT-OMRON	yes			
GazeCom Image				yes
MIT CSAIL				yes
MIT LowRes				yes
VAIQ				yes
IRCCyN Image 1	yes	yes		
Memorability	yes	yes		
McGill ImgSal	yes			yes
KTH	yes	yes	yes	
FiFA	yes	yes		yes
LIVE DOVES	yes	yes		yes
MIT CVCL	yes	yes		yes

III. User Study

In this section, we describe the study carried out to collect information about the way users identify colors in images (both for searching and as dominant), the names of colors they mention, and for what regions of the image they look while enumerating the colors.

A. Participants

Forty participants, divided into two groups of 20, completed the study. The first group (G1) was composed of 14 males and 6 females, with an average of 22 years old (SD=2.86). Six users wore glasses and one wore contact lenses. In the second group (G2) there were 12 males and 8 females, with an average of 21 years old (SD=2.96). Seven wore glasses and two contact lenses. All participants were voluntaries and had never used an eye tracker. Participants from group G1 answered question Q1 *"What are the (up to) three colors that you identify as dominant in this image?"*, while participants from group G2 responded to question Q2 *"What (up to three) colors would you use to search for this image?"*.

B. Apparatus and Material

We used a desktop computer with an application to present the images to the users and register the gaze coordinates collected by the eye tracker. We used TheEyeTribe (an affordable eye tracker), placed under a 20" LCD monitor with a resolution of 1600 x 900 pixels. To collect the coordinates, we used the eye tracker API with the maximum sampling rate supported (60 Hz). Participants were placed at a distance between 50 cm to 70 cm of the monitor (and the eye tracker). All users used the same computer and eye tracker, in the same place, with the same setup.

For the study, we used a set of 100 images (all with Creative Commons licensing) collected from Flickr, and organized into 30 categories: animal, architecture, baby, beach, bird, building, car, clouds, dog, flowers, food, girl, graffiti, lake, landscape, nature, night, people, portrait, river, sea, sign, sky, snow, street, sun, sunset, trees, urban, and water. These categories were based on the ones used in the MIRFLICKR dataset. We did not use this dataset because its images have a reduced size (500 x 500 pixels), which would produce poor results for the gaze coordinates.

To gather the images for our dataset, we performed an advanced search on Flickr, using the category name as tag and "Large" as the minimum size. For each category we selected four images (the first, third, fifth and seventh). After this initial step, we ended up with 120 images. From these, we discarded 20 images that were very similar to others in the dataset, thus getting 100 images. All images were resized, keeping the aspect ratio, to have their width or height equal to the width (1600) or height (900) of the screen (e.g. 1350 x 900; 669 x 900). By doing this, we had a direct correspondence between the images and the screen (and eye tracker) coordinates.

C. Research Questions

Taking into consideration the goals of our study, we identified six research questions that we wanted to answer:

- RQ_1 Can we reduce the name of all mentioned colors to a small subset (palette) of colors?
- RQ_2 Do users use the colors they consider dominant in an image to search for it?
- RQ_3 Where do people look at more often in an image while mentioning its colors?
- RQ_4 Do users look at the regions where the mentioned colors are?
- RQ_5 Does the category of the image affect the gaze pattern of the users?
- RQ_6 Does the type of color (e.g. warm, pure, etc.) influence the set of mentioned colors?

https://theeyetribe.com/
http://press.liacs.nl/mirflickr/

D. Procedures

The sessions took place in a room properly prepared for the study, with adequate lighting and isolation from external interferences. We started the study by showing to the users three plates (4, 7 and 17) from the Ishihara 24 plates test [19], to check for color blindness. Participants who did not pass the test were discarded.

For those who passed the test, we started by collecting demographic information about them, namely age, gender and whether they were wearing glasses or contact lenses, and calibrated the eye tracker. Then, we presented 100 images to each user, one at a time, during seven seconds. For each image users verbally enumerated the names of the colors, while our application registered the coordinates of the image at which they looked using the eye tracker.

Half of the users (G1) enumerated up to three colors that they consider to be the dominant ones, while the other half (G2) enumerated up to three colors that they would use if they wanted to search for the image. The names of the colors were not defined a priori, so users could say any name they wanted. We registered those names as users enumerated them.

IV. Results

This section presents the main results from our study and answers our research questions. Finally, we describe the resulting dataset containing the images, their dominant and search colors, and the gaze coordinates collected.

A. Color Names

After collecting the color names and the gaze coordinates for each user and image, our first step was to group the names of the colors mentioned by users, to see if we could reduce them to a small palette. We performed this separately for each group (G1 - dominant colors, G2 - search colors).

From the analysis of the names, we found that they could be grouped into a reduced number of colors. In fact, the names mentioned more often by the users were the 11 JNS colors, defined by Berlin and Kay. Table II presents the colors enumerated by the participants and how we grouped them into the 11 colors palette. As we can see, for each color of the palette, the color most mentioned was equal to that of the palette. In fact, 90.7% (dominant colors) and 94.0% (search colors) of the names mentioned by the users belonged to the 11 colors palette. These results are in line with our previous study [14], where we found an agreement of 94.6%.

From this, we can conclude that the 11 JNS color palette is appropriated for the identification of dominant colors and the specification of colors for searching. Furthermore, it contains colors whose names people can easily enumerate, enabling them to specify colors using various modalities, such as speech, writing or sketches, making the creation of queries for content-based retrieval or color exploration systems more natural, easier, and simpler to perform.

We could have used the palette introduced by Ware in the scope of an application for nominal information coding [20, p. 126], which is composed of the 11 JNS colors plus the cyan,

Table II
COLORS ENUMERATED BY THE PARTICIPANTS AND HOW WE GROUPED THEM INTO THE 11 COLORS PALETTE.

Color Palette	Dominant Colors (G1) Total	#	Names	Search Colors (G2) Total	#	Names
White	942	935	White	993	992	White
		3	Off-White		1	White Light
		3	White Light			
		1	Transparent White			
Black	520	519	Black	555	554	Black
		1	Ebon		1	Black Gray
Gray	336	320	Gray	290	284	Gray
		7	Light Gray		3	Light Gray
		5	Dark Gray		1	Grayish Brown
		1	Cement		1	Gray Cream
		1	Gray-medium		1	Silver
		1	Gray Tree			
		1	Silver			
Red	487	451	Red	507	481	Red
		14	Brick		14	Brick
		12	Bordeaux		6	Bordeaux
		4	Wine		3	Red Pink
		2	Red Pink		1	Dark Red
		1	Red Brown		1	Wine
		1	Red-sly		1	Garnet
		1	Red wine			
		1	Vermilion			
Brown	724	497	Brown	687	524	Brown
		59	Beige		51	Beige
		59	Skin color		49	Skin Color
		46	Cream		23	Cream
		25	Light Brown		20	Light Brown
		21	Dark Brown		11	Dark Brown
		2	Cream Brown		2	Beige Yellow
		2	Sepia		1	Yellowish Brown
		1	Light Beige		1	Reddish Brown
		1	Dark Beige		1	Brown Earth
		1	Camel		1	Skin Brown
		1	Yellowish Brown		1	Cream Brownish
		1	Brown Beige		1	Maroon
		1	Camel Brown		1	Honey
		1	Gray-brown			
		1	Greenish Brown			
		1	Dirty Brown			
		1	Brown Earth			
		1	Brownish Brown			
		1	Brown Tree			
		1	Creamy			
Orange	170	160	Orange	172	166	Orange
		1	Reddish orange		5	Redhead
		1	Orange Brick		1	Reddish orange
		1	Peach			
		7	Redhead			
Yellow	561	525	Yellow	582	566	Yellow
		13	Golden		6	Golden
		5	Light Yellow		5	Sand Yellow
		5	Yellow Roasted		2	Diarrhea Yellow
		4	Ocher		1	Yellow Yellow
		3	Dark Yellow		1	Sand Yellow
		3	Blond		1	Earth Yellow
		2	Yellowish			
		1	Yellow Vomit			
Green	898	836	Green	959	937	Green
		29	Dark Green		10	Dark Green
		8	Light Green		4	Light Green
		6	Lettuce Green		3	Greenish Yellow
		5	Forest Green		1	Vegetation Green
		4	Acid Green		1	Greenish Brown
		2	Greenish Yellow		1	Greenish Blue
		2	Green Petroleum		1	Grass Green
		2	Greenish		1	Aqua Green
		1	Olive Green			
		1	Greenish Blue			
		1	Lime Green			
		1	Pale Green			
Blue	803	707	Blue	853	794	Blue
		31	Dark Blue		20	Dark Blue
		27	Turquoise		16	Light Blue
		6	Aquamarine		8	Turquoise
		5	Light Blue		4	Sea Blue
		5	Indigo Blue		3	Cyan
		5	Sea Blue		1	Greyish Blue
		5	Navy Blue		1	Dark Blue Gray
		2	Cobalt Blue		1	Navy Blue
		2	Blue Baby		1	Night Blue
		2	Blue Cyan		1	Sky Blue
		2	Blue Green			
		2	Sky Blue			
		1	Greyish Blue			
		1	Blue Gray			
Purple	94	69	Purple	105	64	Purple
		14	Lilac		24	Lilac
		11	Violet		15	Violet
					2	Light Purple
Pink	116	106	Pink	136	128	Pink
		4	Magenta		5	Magenta
		1	Pink Skin Color		2	Hot Pink
		1	Pink Bordeaux		1	Pink Skin Color
		1	Light pink			
		1	Pink Fluorescent			
		1	Salmon			
		1	Fuchsia			

but from our analysis people mentioned cyan a very reduced number of times (only twice for dominants and three times for search). Thus, and despite this 12 colors palette being used by Google and Bing in their image search engines, we found that the 11 JNS colors palette is more natural to users.

B. Dominant Colors vs Search Colors

One of our research questions (RQ_2) seeks to know whether the colors that people use to search for an image are related to the dominant colors of that image. To that end, we started by identifying the most voted colors for each image and for each situation (dominant and search).

We consider a color to be a dominant or search color for an image if it has more than 10% of the votes for that image. We defined this threshold based on our previous tests, where we found that a color with less than 10% has a very low importance on an image [14].

With this approach, we could assign more than the three colors that we asked users to mention, i.e., we decided not to limit the number of colors to three because: 1) some colors can have the same percentage of votes, and we should not ignore one of them just because there are more than three colors; and 2) people perceive colors differently, e.g., some shades of red can be perceived as orange or as brown. Thus, if a significant amount of people identify that in a specific image the existing reds are “brown” or “orange”, this should be reflected on the colors that describe the image. As an example, consider an image that has the following distribution of votes: 35% black, 24% red, 14% white, 14% yellow, 6% blue, 4% orange, and 3% gray. The resulting set of colors will be black, red, white, and yellow, since they have more than 10% of the votes.

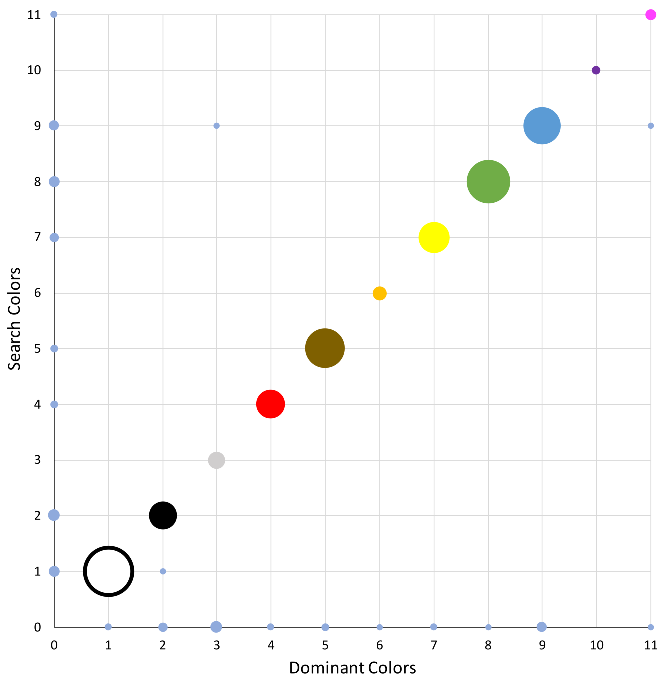


Figure 1. Distribution of the votes for all the images across the eleven colors: white (1), black (2), gray (3), red (4), brown (5), orange (6), yellow (7), green(8), blue (9), purple (10), and pink (11). Zero represents a color that was used as dominant/search but was not used for search/dominant.

After assigning the most voted colors (dominant and search) to all images, we aligned the similar dominant and search colors for each image. We ended up with a set of 400 pairs, some composed of two colors that are similar on both sides (e.g. green-green) and others where we have only one color on one of the sides (e.g. green-none, or none-green). The latter means that there was no similar color on the other side.

Figure 1 presents the distribution of these pairs across the eleven colors. In the diagonal, we can see the colors that were used simultaneously as dominant and for search, while the size of the bubble represents the amount of times that this pair occurred. We have a correspondence of 80.5% between the dominant colors and the search colors, 0.75% where the two colors are different, 10.25% where we have a color for search but not for dominant, and 8.50% on the opposite case.

To assess the agreement between the dominant and search colors, we used similarity metrics to quantify how similar two sets of colors are (dominant colors are denoted by D, while search colors are denoted by S). The measures used were the *Jaccard index* [21] (see Eq. 1), the *Sørensen-Dice index* [22], [23] (see Eq. 2), and the *Overlap coefficient* [24] (see Eq. 3). For all these metrics, the closer its value is to one (or 100%), the more similar the two sets are.

$$jaccard(D,S) = \frac{\left|D \cap S\right|}{\left|D\right| + \left|S\right| - \left|D \cap S\right|} \tag{1}$$

$$sorensenDice(D,S) = \frac{2\,\left|D \cap S\right|}{\left|D\right| + \left|S\right|} \tag{2}$$

$$overlap(D,S) = \frac{\left|D \cap S\right|}{min\Big(\left|D\right|, \left|S\right|\Big)} \tag{3}$$

Let us consider the following example: we have an image with dominant colors white, red, and green, while the search ones are white, red, green, and blue. White, red, and green colors are common to dominant and search, but blue is not.

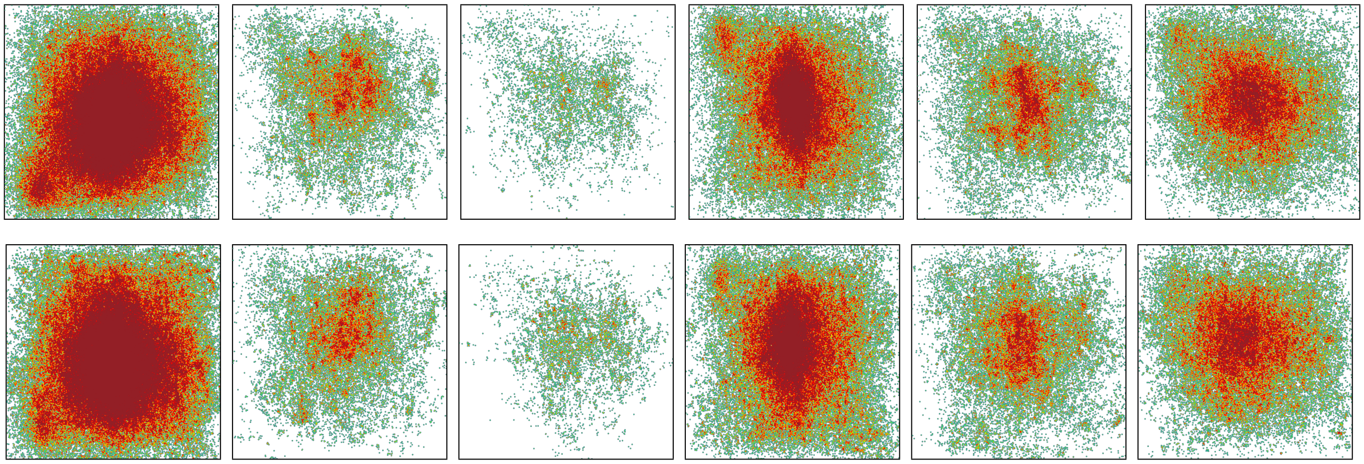
If we we are concerned with exact matches, we should use the $jaccard$ or $sorensenDice$ to assess the agreement. In such case, we would have an agreement of 75% for $jaccard$ and 86% for $sorensenDice$, i.e., in both cases we would be penalizing the result due to the existence of an extra color (blue). Otherwise, we should used $overlap$ that will only consider the exact matches, even if there are more colors assigned to dominant than search, or vice-versa. In this case, we would have an agreement of 100%.

Table III presents a summary of our dataset. We present the number of images per category, the average number of dominant colors and search colors assigned to each category, and the average agreement percentage for each similarity metric. As we can see, around half of the categories (53.44%) have the same average for dominant and search colors, while 33.33% have an average of search colors bigger than the dominant.

For the dominant colors, the following categories have at

Table III
OVERVIEW OF OUR DATASET, SHOWING THE NUMBER OF IMAGES PER CATEGORY, THE AVERAGE NUMBER OF COLORS PER CATEGORY AND THE AVERAGE VALUES FOR EACH METRICS.

Category	#	AvgDC	AvgSC	jaccard	sorensenDice	overlap
Animal	4	3.75	4.00	72.5	82.8	85.5
Architecture	3	3.67	3.67	85.0	92.7	100
Baby	3	4.00	4.00	70.0	81.7	89.0
Beach	4	3.20	3.60	71.3	86.0	100
Bird	4	3.50	3.75	85.5	91.5	100
Building	3	3.00	3.33	75.0	84.3	89.0
Car	4	3.25	3.75	87.5	93.0	100
Clouds	2	3.00	3.00	75.0	83.5	83.5
Dog	3	4.00	4.00	86.7	92.7	100
Flowers	4	3.50	3.50	75.0	84.8	100
Food	3	3.67	3.67	85.0	91.7	100
Girl	4	3.25	3.75	88.8	93.8	100
Graffiti	4	4.00	3.50	76.3	87.5	100
Lake	3	4.00	4.00	89.0	93.3	93.3
Landscape	4	2.75	3.00	79.3	86.8	91.8
Nature	3	3.67	3.67	75.0	84.3	100
Night	3	4.00	3.67	78.3	87.0	91.7
People	4	3.50	3.50	75.0	84.8	91.8
Portrait	3	4.67	4.00	74.0	85.0	100
River	3	4.00	4.00	100	100	100
Sea	3	3.67	3.67	83.3	89.0	89.0
Sign	3	4.00	3.67	91.7	95.3	93.8
Sky	4	3.50	3.50	90.0	93.8	93.8
Snow	3	3.00	3.00	90.0	93.8	93.8
Street	3	4.33	4.33	63.3	77.0	83.3
Sun	3	3.33	3.33	83.3	90.7	100
Sunset	3	3.33	3.67	91.7	95.3	100
Trees	3	3.33	3.33	83.3	90.7	100
Urban	4	3.25	4.00	83.8	90.3	100
Water	3	4.00	3.67	93.3	96.3	100
Total	100	3.59	3.66	82.1	89.3	96.2



(a) All (b) Animal (c) Food (d) Nature (e) People (f) Urban

Figure 2. First row depicts the heatmaps for the dominant colors, and the second row for the search colors. At a given position, a darker shadow of red represents a stronger number of eye gazes, yellow and green represent a medium number, and blue a lower number. (a) all the categories; (b) animal, bird, and dog categories; (c) food category; (d) beach, clouds, flowers, landscape, lake, nature, river, sea, sky, snow, sun, sunset, trees, and water; categories; (e) baby, girl, people, and portrait categories; (f) architecture, buildings, car, graffiti, night, sign, street, and urban categories. (best seen in color)

least an average of four dominant colors: portrait, street, baby, dog, graffiti, lake, night, river, sign, and water; while building, clouds, snow, and landscape categories have three or less dominant colors. Regarding the search colors, the following categories have at least an average of four colors: street, animal, baby, dog, lake, portrait, river, and urban, while clouds, snow, and landscape categories have three or less colors.

If we analyze our results considering the most restrictive measures, we have a $jaccard$ agreement varying from 72% to 100%, and a $sorensenDice$ agreement varying from 82.23% to 100%. The most permissive of the three measures, the $overlap$ varies from 89% to 100%. If we now consider the overall dataset, we have an average agreement of 82.12%±17.04% using $jaccard$, 89.28%±10.80% using $sorensenDice$, and 96.22%±9.99% using $overlap$.

From these values, we can conclude that there is virtually no difference for users when asked about dominant colors in an image and colors to be used for searching for that images. In conclusion, a possible algorithm that identifies dominant colors in images according to human perception, will also serve to highlight the colors that would be used by a user to search for the same image.

C. Focus Regions

Before we analyzed the gaze information, we validated for each participant if there were any corrupted data to be removed (e.g., coordinates outside the image). Across all the images and participants, we had a total of 287 538 gaze coordinates for the dominant colors and 268 716 for the search colors. We discarded around 7% of corrupted data from the former and around 11% from the latter.

To analyze and identify the gaze patterns, we created heatmaps for each image, groups of categories, and the overall dataset, considering the dominant and search colors separated. Since we have images with different orientations and sizes, we normalized the gaze coordinates for each image according to their max width and height. This way, we ensure that our conclusions are correct regardless of the orientation and size of the images. Figure 2 presents the normalized heatmaps of our dataset for both dominant and search colors. To simplify the analysis, we created groups of categories by joining related ones (e.g. animal, bird, and dog). We can see that people look at the central area of images, regardless of being questioned about dominant or search colors (Figure 2a). This is also true for the different groups of categories (Figures 2b - 2f).

Figure 3 present examples of images from our dataset with the corresponding heatmaps overlapped, and the dominant and search colors associated to each one.

Figures 3a and 3f depict a building illuminated at night. People looked more at the center of the image, where we can find the main part of the building, the lamp light and the red lights of traffic. The white, black and yellow colors reflect this gaze behavior, but black (for dominant and search) and blue (for search) are not predominant in the areas where people looked. Figures 3b and 3g depict a street with parked cars. Although, people identified white (surroundings and buildings), gray (car on front and street), red (car), and green (trees) as the dominant and search colors, in both cases, they mainly looked at the red car.

Figures 3c and 3h show a dog resting on grass. In this case, people mainly looked at the dog face and dog-collar. The predominant colors were green (grass), blue (dog-collar), and finally brown (dog body and face). It is interesting to notice that regardless of the small size of the dog-collar (when compared with the size of the dog), the blue color had more votes than the brown. Figures 3d and 3i depict a purple flower. Here, people mainly looked at the stigma of the flower (white/yellow area in the middle of the flower), the top part of the flower, and some leaves. The identified search colors were purple and green (flower), while for dominant colors, the black and blue colors were also identified.

(a) White, black, yellow (b) White, gray, red, green (c) Green, blue, brown (d) Black, green, purple, blue (e) White, green, brown

(f) White, black, yellow, blue (g) White, gray, red, green (h) Green, blue, brown (i) Green, purple (j) White, green, brown

(k) Gray, yellow, pink, brown (l) Orange, yellow, pink, brown (m) White, green, blue, brown (n) White, blue (o) Black, orange, yellow

(p) White, gray, yellow, brown (q) White, black, yellow, blue (r) White, green, blue, brown (s) White, blue (t) Black, orange, yellow

Figure 3. Examples of heatmaps and images for the dominant colors (first and third row) and search colors (second and forth row). (best seen in color)

Figures 3e and 3j show a young girl laying on the grass. We can see that people mainly looked at the girl face, but indicated white (dress), green (grass), and brown (hair and maybe skin) as the predominant colors for both search and dominant. Figures 3k and 3p depict a nightscape with buildings across the river. Similarly to Figures 3a and 3f, people mainly looked at the center of the image where the buildings and lights are concentrated. For this image, the dominant colors were gray (from the sky and maybe buildings), yellow (from the buildings lights), pink (maybe the central building resembles light pink, and the top structure at its left, dark pink), and brown (surroundings and shadows). It is interesting that in search colors, people also looked at the top of the building with a white light (right top part of image) and the building front illuminated with a white light (right middle part of image). As a result, white was one of the predominant colors identified.

In Figures 3l and 3q, we have the face of a man surrounded by packages of chocolates. In both cases, people mainly looked at his face. However, in both cases, people identified the colors of the chocolate packages (e.g., orange, yellow, white, black). Figures 3m and 3r depict a river with some vegetation. People looked more at the top of the image, where the vegetation and the narrowest river area are. In both cases, people identified white (from the water foam), green (vegetation), blue (from the narrowest part of the river), and brown (from the banks and wider area of the river) as the predominant colors.

In Figures 3n and 3s, we have the sky with clouds. People mainly looked at the center of the image, where the biggest portion of the clouds are. Not surprisingly, the predominant colors identified were white and blue. Finally, in Figures 3o and 3t, we have a sunset on the river. People looked to the sun and the area around it. However, the most predominant color was black, where people barely looked at.

D. *Discussion*

Based on the results from our study, we will answer now the research questions that we raised in Section III.

According to Table II, we can say that the answer to our RQ_1 is yes, that is, we can reduce all the color names mentioned by users to a small subset of colors, such as the 11 colors palette suggested by Berlin and Kay. From the comparison and the assessment that we made on Section IV-B, we verified that there is a strong similarity between the dominant colors and search colors mentioned for each image. Thus, we can say that the colors that users would use to search for an image are the dominant colors of the image (RQ_2).

Although, the gaze pattern differs a bit among the groups of categories (RQ_5), as illustrated in Figure 2, the most looked region is the center of images (RQ_3). Moreover, we noticed that people do not look at some regions of the image, but enumerate their colors, and look at other parts of the image (e.g. faces, bright spots, lights) and do not mention their colors (RQ_4). We noticed that people identify as predominant colors, colors from small areas of the image probably because they have striking colors (e.g., red car, blue dog-collar) (RQ_6).

In summary, we can say that users mentioned colors from the whole image and not only from the area where they looked at. In particular, we noticed that users focus on faces, but identify as predominant colors those of the surrounding objects (e.g. hair, clothes). This focus on faces was also observed by Cerf *et al.* in their study [25]. Finally, and although people use the same "scanning" method for the identification of the dominant and search colors, they slightly tend to disperse more their gaze while identifying colors for searching purposes.

E. Resulting Dataset

Our dataset, named UL-GDSC (Gaze on Dominant and Search Colors), is composed of 100 images collected from Flickr and resized to match the largest size of the screen (width of 1600 or height of 900 pixels). Images are organized in 30 categories, as shown in Table III, and are annotated with their dominant colors, the colors that people would use to search for them, and the coordinates where people gaze at while identifying the colors. We made UL-GDSC dataset publicly available to the community.

Each image has two sets of colors (dominant and search colors) based on the colors that received more than 10% of the votes. On average, images have three to four colors associated. The gaze coordinates in the dataset are the average of three consecutive raw coordinates provided by the eye tracker. Thus, we were able to have more stabilized gaze coordinates, with the cost of having less values per second, since we indirectly reduced the sampling rate (from 60 Hz to 20 Hz).

A salient aspect of the UL-GDSC is that it contains not only the colors that people identified as dominant and for searching, but also the eye movements people performed while doing it.

V. Conclusion

In this paper, we presented the results of a study with users to identify the predominant colors in images and at which regions they look while mentioning those colors. From the data collected, we were able to confirm that the JNS palette contains a set of colors that is representative of the color names that users mentioned.

Additionally, we measured the similarity between the dominant colors associated to an image and the colors used to search for it, and found that they are very similar. So, we can use the dominant colors of the images as a content descriptor, since users would use them for searching.

The analysis of the gaze data revealed that overall there is no strong relation between the colors of the regions where people look at and the predominant colors identified in the image. Furthermore, people look mainly at the center of the image, regardless of its category.

http://www.di.fc.ul.pt/~mjf/research/ul-gdsc/

Acknowledgment

This work was supported by national funds through Fundação para a Ciência e Tecnologia, under LASIGE Strategic Project - UID/CEC/00408/2013.

References

[1] A. M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive Psychology*, vol. 12, no. 1, p. 97136, Jan. 1980.

[2] B. Berlin and P. Kay, *Basic color terms : their universality and evolution*. University of California Press, 1969.

[3] A. Mojsilovic, "A computational model for color naming and describing color composition of images," *Transactions on Image Processing*, pp. 690–699, 2005.

[4] J. van de Weijer, C. Schmid, and J. Verbeek, "Learning color names from real-world images," in *Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.

[5] N. Moroney, "Unconstrained web-based color naming experiment," in *Color imaging VIII: Processing, hardcopy, and applications*. International Society for Optics and Photonics, 2003, pp. 36–47.

[6] G. Menegaz, A. Le Troter, J. Sequeira, and J.-M. Boi, "A discrete model for color naming," *Journal on Advances in Signal Processing*, 2007.

[7] R. Benavente, M. Vanrell, and R. Baldrich, "Parametric fuzzy sets for automatic color naming," *Journal of the Optical Society of America A*, pp. 2582–2593, 2008.

[8] E. Y. Chang, B. Li, and C. Li, "Toward perception-based image retrieval," in *Workshop on Content-based Access of Image and Video Libraries*, 2000, pp. 101–105.

[9] J. R. Smith and S.-F. Chang, "Visualseek: a fully automated content-based image query system," in *ACM international conference on Multimedia*, 1997, pp. 87–98.

[10] Y. Deng, B. Manjunath, C. Kenney, M. S. Moore, and H. Shin, "An efficient color representation for image retrieval," *Transactions on Image Processing*, pp. 140–147, 2001.

[11] A. Mojsilovic, H. Hu, and E. Soljanin, "Extraction of perceptually important colors and similarity measurement for image matching, retrieval and analysis," *Transactions on Image Processing*, pp. 1238–1248, 2002.

[12] A. Atsalakis and N. Papamarkos, "Color reduction and estimation of the number of dominant colors by using a self-growing and self-organized neural gas," *Engineering Applications of Artificial Intelligence*, 2006.

[13] A. A. Younes, I. Truck, and H. Akdag, "Image retrieval using fuzzy representation of colors," *Soft Computing*, pp. 287–298, 2007.

[14] J. C. Amante and M. J. Fonseca, "Fuzzy color space segmentation to identify the same dominant colors as users," in *International Conference on Distributed Multimedia Systems*, 2012, pp. 48–53.

[15] A. Talib, M. Mahmuddin, H. Husni, and G. Loay E, "A weighted dominant color descriptor for content-based image retrieva," *Journal of Visual Communication and Image Representation*, pp. 345–360, 2013.

[16] J. Lazar, J. H. Feng, and H. Hochheiser, *Research Methods in Human-Computer Interaction*. John Wiley & Sons, 2010.

[17] S. Winkler, F. M. Savoy, and R. Subramanian, "X-eye: A reference format for eye tracking data to facilitate analyses across databases," in *Human Vision and Electronic Imaging*, 2014.

[18] S. Winkler and R. Subramanian, "Overview of eye tracking datasets," in *International Workshop on Quality of Multimedia Experience*, 2013.

[19] S. Ishihara, *Tests for color-blindness*. Handaya, Tokyo, Hongo Harukicho, 1917.

[20] C. Ware, *Information Visualization: Perception for Design*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2004.

[21] M. J. Zaki and W. M. Jr, *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge University Press, 2014.

[22] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, pp. 297–302, 1945.

[23] T. A. Sørensen, "A method of establishing groups of equal amplitude in plant sociology based on similarity of species content, and its application to analyses of the vegetation on {Danish} commons," *Biologiske Skrifter*, pp. 1–34, 1948.

[24] Y. A. Pesenko, "Principles and Methods of Quantitative Analysis in Faunistical Researches," *Moscow (Nauka) [in Russian]*, 1982.

[25] M. Cerf, J. Harel, W. Einhaeuser, and C. Koch, "Predicting human gaze using low-level saliency combined with face detection," in *Advances in Neural Information Processing Systems*, 2008, pp. 241–248.