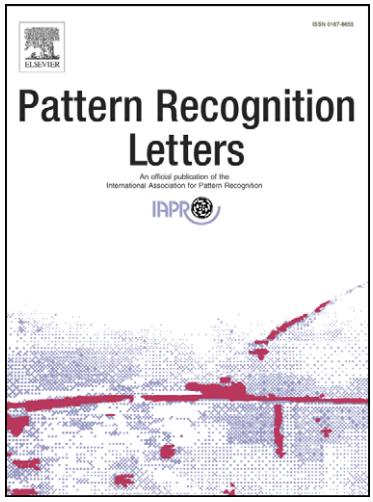


Modeling Human Color Categorization

E.L. van den Broek, Th.E. Schouten, P.M.F. Kisters

PII: S0167-8655(07)00275-9
DOI: 10.1016/j.patrec.2007.09.006
Reference: PATREC 4244

To appear in: *Pattern Recognition Letters*

Received Date: 8 December 2006
Revised Date: 19 July 2007
Accepted Date: 3 September 2007

Please cite this article as: van den Broek, E.L., Schouten, Th.E., Kisters, P.M.F., Modeling Human Color Categorization, *Pattern Recognition Letters* (2007), doi: 10.1016/j.patrec.2007.09.006

Modeling Human Color Categorization

E.L. van den Broek [a] Th.E. Schouten [b] P.M.F. Kisters [c]

[a] *Center for Telematics and Information Technology, University of Twente*
P.O. Box 217, 7500 AE Enschede, The Netherlands

[b] *Institute for Computing and Information Science, Radboud University Nijmegen*
P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

[c] *GX creative online development B.V.*
Wijchenseweg 111, 6538 SW Nijmegen, The Netherlands

Abstract

A unique color space segmentation method is introduced. It is founded on features of human cognition, where 11 color categories are used in processing color. In two experiments, human subjects were asked to categorize color stimuli into these 11 color categories, which resulted in markers for a Color LookUp Table (CLUT). These CLUT markers are projected on two 2D projections of the HSI color space. By applying the newly developed Fast Exact Euclidean Distance (FEED) transform on the projections, a complete and efficient segmentation of color space is achieved. With that, a human-based color space segmentation is generated, which is invariant for intensity changes. Moreover, the efficiency of the procedure facilitates the generation of adaptable, application-centered, color quantization schemes. It is shown to work excellently for color analysis, texture analysis, and for Color-Based Image Retrieval purposes.

Key words: 11 color categories, human color categorization, color space segmentation, Fast Exact Euclidean Distance (FEED) transform

1 Introduction

Digital imaging technology is more and more embedded in a broad domain. As a consequence, digital image collections are booming, which creates the need for efficient data-mining in such collections [10]. An adequate model of

Email addresses: e.l.vandenbroek@utwente.nl (E.L. van den Broek), T.Schouten@cs.ru.nl (Th.E. Schouten), p.kisters@gx.nl (P.M.F. Kisters).

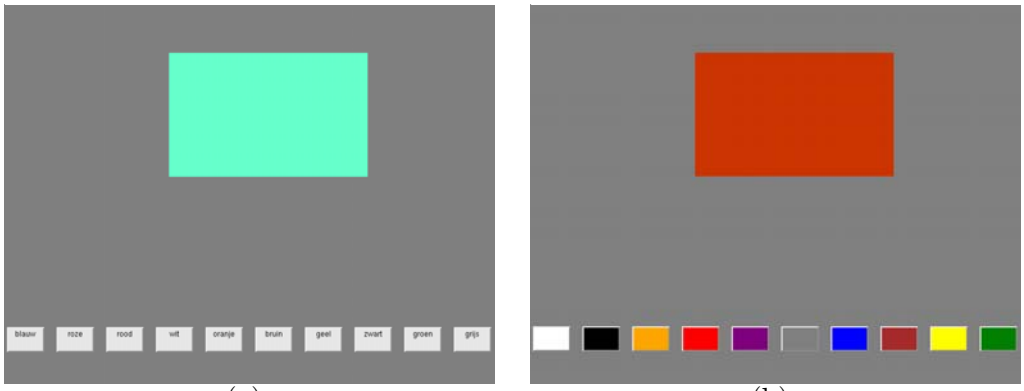
human visual perception would facilitate data-mining [10,15]. Our approach, hereby, is to utilize human cognitive and perceptual characteristics.

For a broad range of data-mining applications (e.g., the cultural domain [3,15]), various image processing techniques have been adopted. However, the current paper focuses on a generic image processing technique: a color quantization scheme based on human perception. This unique color space segmentation is both relevant and suitable for the development and study of Content-Based Image Retrieval (CBIR) in its variety of contexts [11,14,15].

In general, we argue that color should be analyzed from the perspective of human color categories. Both to relate to the way people think, speak, and remember color and to reduce the data from 16 million or more colors to the limited number of 11 color categories: black, white, red, green, yellow, blue, brown, purple, pink, orange, and gray. Research from diverse fields of science emphasize their importance for human color perception [2,6,8]. The use of this knowledge can possibly provide a solution for problems concerning the accessibility and the availability of knowledge, where color analysis is applied in data-mining. In addition, such a human-centered approach can tackle the computational burden of traditional (real-time) color analysis [3,6,7].

The 11 color categories are applicable for a broad range of CBIR domains [14], where in specific domains, other sets of colors might be more appropriate. In this paper, we regard the 11 color categories as they are used in daily life. These color categories are constructed and handled by methods that are presented in this paper. However, in the same way, it is possible to incorporate another set of colors, which is user, task, or application specific.

This paper presents a line of research starting with psychophysical experiments in Section 2. This provided us with color markers for a Color LookUp Table (CLUT) in the RGB color space. The boundaries between the color categories in the RGB space are expected to be too complex to be determined, using the limited number of CLUT markers. Therefore, we describe in Section 3 how the RGB space is transformed into two 2D projections of the HSI color space in which the boundaries are less complex. In Section 4, we present the newly developed Fast and Exact Euclidean Distance (FEED) transformation, adapted such that it can handle multi class data; in our case: the 11 color categories. Section 5 describes how the CLUT markers are fed to FEED to find the boundaries and how this is used to segment the complete color space. In Section 6, the segmented color space is validated through a comparison with human color categorization. We finish the paper with a discussion on the work presented in Section 7.



(a) (b)

Fig. 1. Screendump of the user interface of (a) the color memory experiment (gray buttons, labeled with a color name) and (b) the color discrimination experiment (colored buttons without a label). The stimulus did have a size of 9.5×6.0 cm, the buttons measured 1.8×1.2 cm with 0.6 cm between them.

2 Validation of the 11 color categories

Twenty-six subjects with normal or corrected-to-normal vision and no color deficiencies participated voluntary in two tasks. The first task was to write down the first 10 colors that came to mind. This task was embedded in the research because the 11 color categories are still a topic of debate; e.g., see [6,8]. It enabled us to verify its validity for our research. The second task consisted of both a color memory experiment and a color discrimination experiment.

2.1 Method

For the color memory experiment, the subjects were instructed to categorize the stimulus into one of the color categories, represented by buttons with their names (task 2a; see Figure 1a). In the color discrimination experiment, the subjects were asked to press one of the 11 focal-color buttons (i.e., a typical color for a color category) that best resembled the stimulus (task 2b; see Figure 1b). Both experiments consist of four blocks of repetitions of all stimuli (in a randomized order).

The experiments ran in an average office environment on a PC with an Intel Pentium II 450 MHz processor, 128mb RAM, a Matrox Millennium G200 AGP card, and with a Logitech 3-button Mouseman (model: M-S43) as pointing-device. The experiments were conducted in a browser-environment with Internet Explorer 6.0 as browser and Windows 98SE as operating system, using 16-bit colors, where respectively 5, 6, and 5 bits are assigned to the red, green, and blue channel.

Color name	freq. (min.-max.)
red, green, blue, yellow	26 (84.4%-100.0%)
purple	24 (74.5%- 98.8%)
orange	22 (65.7%- 94.3%)
black, white, brown	20 (57.5%- 89.2%)
gray	15 (38.9%- 74.4%)
pink	11 (25.6%- 61.1%)

Color name	freq. (min.-max.)
violet	06 (10.8%- 42.5%)
beige	04 (5.7%- 34.3%)
ocher	03 (3.3%- 30.0%)
turquoise, magenta,	02 (1.1%- 25.5%)
indigo, cyan	
silver, gold, bordeaux-red	01 (0.7%- 20.7%)

Table 1
Frequency and confidence-intervals (p at 99%) of color names mentioned.

The stimuli were the full set of 216 web-safe colors [16]. These are defined as follows: The R, G, and B dimensions (coordinates) range from 0 to 255 and are treated equally. For each dimension, six values are chosen: 0 (0%), 51 (20%), 102 (40%), 153 (60%), 204 (80%), and 255 (100%). Each of these six values is combined with each of the six values of the two other dimensions. This results in $6^3 (= 216)$ triple of coordinates in the RGB-space. These RGB-values result for both Internet Explorer and Netscape under both the Windows and the Mac operating system, in the same (non-dithered) colors, under the prerequisite that the operating system uses at least 8-bit (256) colors.

2.2 Results

the main results, which confirm the existence of the 11 color categories, as is illustrated in Table 1. Noteworthy is that all 26 participants wrote down the colors red, green, blue, and yellow; all belong to the 11 focal colors or color categories [2,6,8]. With 11 occurrences, pink was the least mentioned focal color. Followed by the most frequently mentioned non-focal color: violet, which was mentioned by six of the participants. The complete results of task 1 are presented in Table 1.

In the classification of the web-safe colors by the subjects, three sets of color markers can be distinguished: i) non-fuzzy color markers: web-safe colors categorized by a magnitude of at least 10 subjects to one color category; ii) fuzzy color markers: web-safe colors that are assigned to two different color categories by at least 10 subjects; iii) colors of which the subjects did not agree to what color category they belong. The set of non-fuzzy color markers are used in further processing since the color category they belong to is undisputed. The fuzzy color markers will be used in a later stage to validate the final color segmentation, based on the non-fuzzy color markers. The third and remaining category is excluded as data for further research.

In general, color matching using a Color LookUp Table (CLUT), based on the color markers derived from the experimental results, could enhance the

color matching process significantly and may yield more intuitive values for users [2,6]. In addition, such a coarse color space quantization of 11 color categories reduces the computational complexity of color analysis drastically, compared to existing matching algorithms of image retrieval engines that use color quantization schemes (c.f. PicHunter [1]: HSV $4 \times 4 \times 4$ and QBIC [3]: RGB $16 \times 16 \times 16$). The coarse 11 color categories quantization makes it also relatively invariant with respect to intensity changes [6]. For more detailed discussions concerning color quantization, we refer to Schettini, Coicca, and Zuffi [11].

3 The segmentation framework

The experiments presented in Section 2 provided us with categorized color markers for a Color LookUp Table (CLUT). These color markers are considered scarce data for segmenting a color space into 11 categories. Therefore, a framework is needed to that provides means to maximize the efficiency. In this section, we explain the framework that is used for the segmentation process.

3.1 HSI: segmentation color space for scarce data

The color markers are RGB coordinates; however, the RGB color space is not perceptually intuitive. Hence, the position and shape of the color categories within the RGB color space are complex. Therefore, for the full color space categorization, the HSI color space is used, which is (i) perceptually intuitive, (ii) performs as good as or better than perceptual uniform color spaces such as CIE LUV [9], and (iii) the shape and position of the color categories are less complex functions of location and orientation than with the RGB color space. See Figure 2 for a visualization of the RGB and HSI color spaces as well as their relation.

Prosperously, the 216 web-safe colors are clearly distinct for human perception. As a consequence, in a perceptual intuitive color space some distance is present between them. Moreover, the perceptual intuitive character of the HSI color space results in an orientation of adjacent colors such that the web-safe colors are spatially arranged by color category.

Let us now briefly discuss the HSI color space, an intuitive color space that proved to work good under varying circumstances [9,11]. The axes of the HSI space represent hue (i.e., basic color index), saturation (i.e., colorfulness or chromatic purity), and intensity (i.e., amount of white present in the color). The shape of HSI color space can be displayed as a cylinder: intensity is the

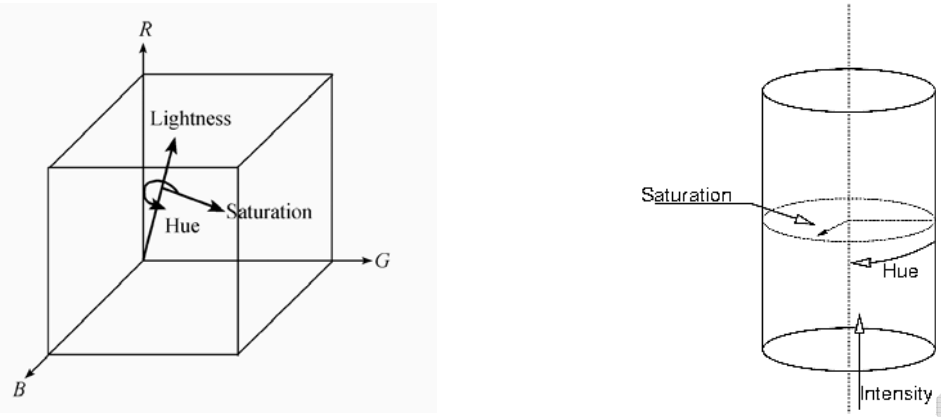


Fig. 2. Left: The relation between the RGB and the HSI color space, from the perspective of the RGB color space. Right: The cylinder shaped representation of the HSI (Hue, Saturation, and Intensity) color space, as used in this research.

central rod, hue is the angle around that rod, and saturation is the distance perpendicular to that rod; see also Figures 2 and 3. The color categories' orientation is as follows: Around the intensity axis, the achromatic categories are located, as is shown in Figure 3. The achromatic region has a conical shape and is described with small saturation values, the complete range of intensity, and the complete range of hue values. Around this achromatic region, the chromatic categories are located. Chromatic categories have high saturation values and occupy a part of both the total hue and the total intensity range.

3.2 From 3D HSI color space to two 2D representations

Since the HSI color space is a 3D space, the boundaries between color categories consist of 3D functions. However, the amount of HSI CLUT markers is too limited to determine the exact boundaries through a 3D segmentation, which would evolve in very weak estimations of the shape of color categories in color space. However, the perceptually intuitive axes of the HSI color space do allow a reduction in the complexity of boundary functions without loosing essential features of the boundaries. The intuitive values that the axes represent facilitate the separation of chromatic and achromatic categories, using two 2D projections. Thereby, we use three assumptions:

(1) The boundaries between achromatic categories and chromatic categories do not excessively change over the hue range; see also [6].
(2) The boundaries between chromatic categories do not excessively change over the saturation axis and can be approximated by a linear function toward the central rod of the color space; i.e., the intensity axis [6]. The intuitive features of the HSI space provide strong arguments for the latter assumption: Consider a chromatic point of the outer boundaries of the

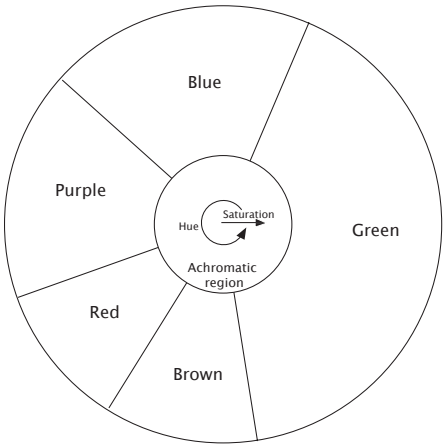


Fig. 3. Separation of the various regions in the hue-saturation plane with intensity value 300. Shown is both the separation between chromatic and achromatic regions and the separation between different hues (or colors), within the chromatic region.

HSI space (with maximum saturation). When the saturation value is lowered, the perceived color becomes 'decolorized' or pale. Nevertheless, in general the color is perceived as belonging to the same color category.

(3) The two boundaries between achromatic categories can each be expressed with a single intensity value.

Given the latter assumptions, segmentation can be done in three steps. First, the separation of achromatic categories and chromatic categories through a 2D plane leaving out the hue axis resulting in a saturation-intensity plane. Second, the segmentation of the chromatic colors by leaving out the saturation axis: the hue-intensity plane. Third, the segmentation of the individual achromatic categories is performed in a saturation-intensity plane.

4 Feeding the color markers to a distance transform

The previous section presented a framework that utilized two HSI 2D planes. In these 2D planes, the categorized color data will occur as grouped clouds of data points. However, the information that is available about humans' color categorization through the experiments (see Section 2) does only assign a limited number of points in these planes. Therefore, we applied distance mapping, where each point gets a distance measure to the set of categorized points by humans.

The speed of distance transforms is determined by the precision needed. Images are a priori an approximation of reality due to various forms of noise. It might be that introducing additional approximations in the image processing chain to make it faster, has no effect on the final quality of the application. The

best way to test that is to compare with no additional approximations in the chain. Therefore, we preferred an as accurate as possible distance transform, preferably exact.

Distance transforms can be applied on all kinds of data. In this paper, we discuss the 11 color categories, which are generally applicable. However, the categories that are needed depend on the specific application; e.g., a catalog of paintings or a stained cell tissue database. There might be the need to adapt the color categories quickly to specifications based on a certain domain or task. Moreover, the perception of individual users differs and systems are needed that use user profiles, which would be in our case: a user specific (i.e., personalized) color space segmentation. The latter is of great importance since users are in interaction with the systems, which use image analysis techniques, and judge their results. Therefore, we wanted a fast color space segmentation regarding computer and human resources.

The distance transform to be applied both needs to be fast enough and preferably exact. For this purpose, we developed the Fast Exact Euclidean Distance (FEED) transform, which is introduced in the next section.

4.1 Fast Exact Euclidean Distance (FEED)

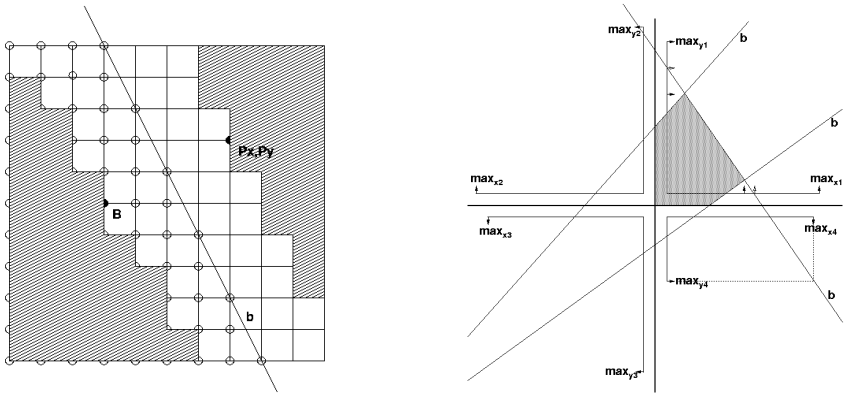
In contrast with the state-of-the-art approaches such as Shih and Wu's two scan algorithm (EDT-2) [13], we have implemented the Euclidean Distance (ED) transform starting directly from its definition. Or rather its inverse: each object pixel q, in the set of object pixels (O), *feeds* its ED to all non-object pixels p. The naive algorithm then becomes:

initialize $D(p) \;=\; \text{if } (p \in O) \text{ then } 0, \text{ else } \infty$
  foreach $q \in O$
    foreach p
      update : $D(p) = \min(D(p), \mathrm{ED}(q, p))$

This algorithm is extremely time consuming, but is speeded up by:

- restricting the number of object pixels q that have to be considered, to the border pixels of the objects
- pre-computation of $\mathrm{ED}(q, p)$
- restricting the number of background pixels p that have to be updated for each border pixel using bisection lines (see Figure 4)

The method used for searching for other object pixels, bookkeeping of the bisection lines, and determining which background pixels to update is carefully designed (see also Figure 4). This, to ensure that it takes much less time than



(a) (b)

Fig. 4. (a) Principle of limiting the number of background pixels to update. Only the pixels on and to the left of the bisection line b have to be updated. B is the border pixel under consideration, p is an object pixel. (b) Bookkeeping of the sizes (the "max") of each quadrant. Updating process: On each scan line the bisection lines b determine the range of pixels to update.

the time gained by not updating the other background pixels.

This resulted in an exact but computationally less expensive algorithm for the ED transform: the Fast Exact Euclidean Distance (FEED) transformation. It is recently introduced by Schouten and Van den Broek [12]. For both algorithmic and implementation details we refer to this paper.

4.2 Benchmarking FEED

In the latest experiments, we have compared FEED with Shih and Wu's EDT-2 [13] and with the city-block (or Chamfer 1,1) distance, as baseline. In Table 2, the timing results are provided for the city-block measure, for EDT-2, and for FEED. As was expected, with a rough estimation of the ED, the city-block distance outperformed the other two algorithms by far. More surprising is that FEED is more than twice as fast as EDT-2 (see Table 2). The time complexity of the city-block and EDT-2 methods is $\mathcal{O}(n)$ with n the number of pixels in the image. FEED behaves over a large range of tested images as having a time complexity of $\mathcal{O}(n)$; however, this could not be proved.

The aim of FEED is to utilize exact ED transforms. Hence, next to the timing results, the percentage of errors made in obtaining the ED is of interest to us. The city-block transform resulted for all images in an error-level of less than 5%; see Table 2. Shih and Wu claimed that their EDT-2 provided exact EDs. However, in 1% of the cases errors occur in their algorithm, as reported in Table 2. So, FEED is the only algorithm that provided the truly exact ED for all instances.

Images	Timing (errors) per distance transform		
	city-block	EDT-2	FEED
standard	8.75 s (2.39%)	38.91 s (0.16%)	17.14 s
rotated	8.77 s (4.66%)	38.86 s (0.21%)	18.02 s
larger obj.	8.64 s (4.14%)	37.94 s (0.51%)	19.94 s

Table 2
Average timing results (in seconds) for three sets of images on the city-block transform, Shih and Wu's two scan method (EDT-2) [13], and for FEED [12]. Between brackets the errors (in %) of the city-block (or Chamfer 1,1) and EDT-2 transform. Note that no errors of FEED are mentioned since FEED provides truly exact EDs.

4.3 FEED for multi class data

Now, let us consider the case that multiple labeled classes of data points (e.g., color markers assigned to color categories) are present and, subsequently, FEED is applied for data space segmentation. In such a case, the class of the input pixel that provides the minimum distance is placed in a second output matrix. To achieve this, the update step of FEED is changed to:

$$\begin{aligned} \text{update}: \ & \text{if } (\text{ED}(q,p) < D(p)) \\ & \text{then } (\ D(p) = \text{ED}(q,p); \quad C(p) = C(q)\ ) \end{aligned}$$

where C is a class matrix, in our case the colors assigned to one of the 11 color categories.

Similar changes were applied to the city-block and EDT-2 methods; thus, three distance transforms and classification methods are obtained. These methods are then applied on a set of hue-intensity images as used in the remaining of this paper. The results are shown in Table 3. The addition of the classification increases the time for all three methods. As was expected FEED is faster than EDT-2 and slower than the city-block approximation of the ED. But, only FEED provides no misclassified pixels.

The minimum distance value then indicates the amount of probability (or weight) that the pixel belongs to the class. This can be visualized by different color ranges, for each class. By extracting the pixels that express minimum probability, a Voronoi diagram can be generated. In our application this diagram defines the borders between the color categories (see Figures 5c and 6c).

5 The segmentation process

Sections 3 and 4 described the means to do color space segmentation with scarce data. In this section, we describe the actual segmentation process of

Method	Timing and classification errors		
	DT	Classification	% wrong
city-block	23.3 ms	28.5 ms	1.9 %
EDT-2	64.4 ms	73.2 ms	0.12 %
FEED	32.7 ms	41.5 ms	0 %

Table 3
Average timing results in milliseconds on an ADM Athlon XP 2100 machine for a set of hue-intensity images of size 815×941 for the city-block transform, Shih and Wu's two scan method (EDT-2) [13], and for FEED [12]. The last colomn gives (in %) the number of wrongly assigned pixels. FEED provides truly exact EDs.

the color space with as input the color markers from the experiment described in Section 2. Result of this process is a fully segmented color space and, subsequently, a complete Color LookUp Table (CLUT) for the 11 color categories.

The first phase in preprocessing is the conversion of the RGB color markers (see Section 2) to HSI color markers. The conversions as given by Gevers and Smeulders [5] were adopted:

$$H(R,G,B) = arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B))}\right) \tag{1}$$

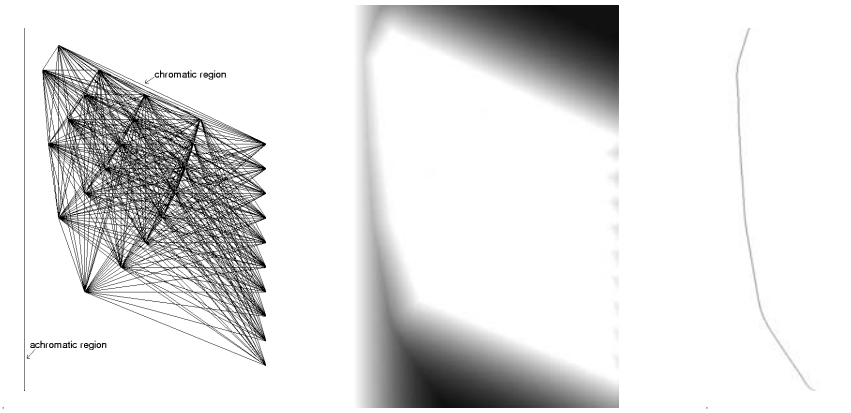
$$S(R,G,B) = 1 - \frac{3 \cdot min(R,G,B)}{R+G+B} \tag{2}$$

$$I(R,G,B) = R+G+B \tag{3}$$

Please note that the original conversion was adapted by the use of a factor 3 instead of 1 in Equation 2. This changes the range of the saturation (S) from $[\frac{2}{3}, 1]$ to $[0, 1]$.

The next phase in preprocessing is the generation of the 2D planes. The 11 color categories of the HSI CLUT were divided into two groups: the achromatic categories (i.e., black, gray, and white) and the chromatic categories (i.e., blue, yellow, green, purple, pink, red, brown, and orange). In this way, each group could be processed in a separate 2D plane: the saturation-intensity and the hue-intensity plane, as is described in Section 3.2 and illustrated in Figures 5 and 6.

As a last phase of preprocessing, the HSI color markers, were plotted in the saturation-intensity and the hue-intensity planes. Next, for each color category, a fully connected graph is generated, using a line generator (see Figures 5a and 6a). For each category, we can assume that all points within the boundaries of the connected graphs belong to the color category to which all individual data points were assigned. The graphs were filled resulting in convex hulls: an initial estimation of the color categories within the HSI color space.



(a) (b) (c)

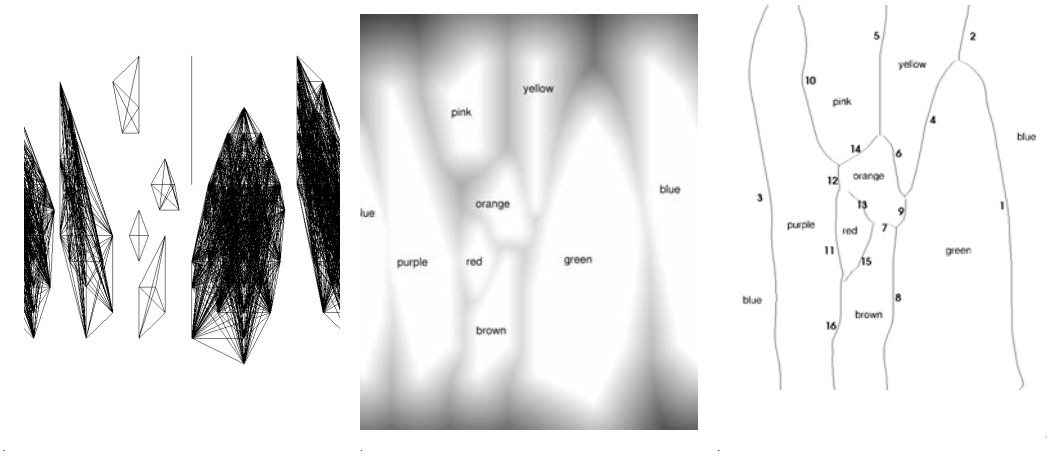
Fig. 5. The processing scheme of the separation of the chromatic from the achromatic color categories, in the saturation-intensity plane, using human color categorization data (see Section 2): (a) The fully connected graphs of the categorized CLUT markers. (b) The weighted distance map, created using Fast Exact Euclidean Distance (FEED) transformations [12]. (c) The resulting chromatic-achromatic border. Note that saturation is presented on the horizontal axis and intensity on the vertical axis.

First, the saturation-intensity plane allows segmentation of the color space between achromatic categories and chromatic categories. In this projection, the achromatic categories are distinguished from the chromatic categories as a line and a cloud of data points (see Figure 5a). Note that, when leaving out the hue axis, the main color information is left out and thus all individual chromatic categories are resembled in a single cloud of data points.

Second, the chromatic category data is projected in the hue-intensity plane. The result is a plane with non-overlapping clouds of categorized points, as illustrated in Figure 6a.

Third, the segmentation of the individual achromatic categories is performed. Since these categories do not represent any basic color information, the hue axis does not contain useful information for these categories. Thus, the segmentation of these individual achromatic color categories is done in a saturation-intensity plane (see Figure 5). A drawback for the differentiation between the achromatic colors is the lack of achromatic color markers. Therefore, we take two intensity values that describe the boundaries between individual achromatic categories in three sections of equal length.

The two 2D projections of the HSI color space with the filled convex hulls were fed to FEED (see Section 4), two distance maps were generated (see Figures 5b and 6b). From these, distance maps Voronoi diagrams were generated, as are shown in Figures 5c and 6c. This resulted in a segmentation between the achromatic and chromatic categories followed by a segmentation between individual chromatic categories. From the Voronoi diagrams, HSI values of the



(a) (b) (c)

Fig. 6. The processing scheme of the separation of the chromatic color categories in the hue-intensity plane, using human color categorization data (see Section 2) (note that the hue-axis is circular): (a) The fully connected graphs of the categorized CLUT markers. (b) The labeled weighted distance map created using Fast Exact Euclidean Distance (FEED) transformations [12]. (c) The resulting borders between the chromatic color categories. Note that hue is presented on the horizontal axis and intensity on the vertical axis.

	Purple	Pink	Orange	Red	Brown	Yellow	Green	Blue	Black	Gray	White
Purple	C	N	-	N	-	-	-	N	-	-	-
Pink	N	C	N	N	-	-	-	-	-	-	-
Orange	-	N	C	-	N	N	-	-	-	-	-
Red	N	-	N	C	N	-	-	-	-	-	-
Brown	-	-	N	N	C	-	-	-	-	-	-
Yellow	-	-	-	-	-	C	N	-	-	-	-
Green	-	-	-	-	-	N	C	N	-	-	-
Blue	N	-	-	-	-	-	-	C	-	-	-
Black	-	-	-	-	-	-	-	-	C	N	-
Gray	-	-	-	-	-	-	-	-	N	C	N
White	-	-	-	-	-	-	-	-	-	N	C

Table 4
Color categories (indicated by C) and their neighbor color categories (indicated by N) in the segmented HSI color space, which resembles human color categorization as unraveled through the experiments (see Section 2).

borders between color categories were deduced. From this representation, a fast categorization mechanism is easily established; by storing the categorized colors and convert them back to RGB values. This results in a $256 \times 256 \times 256$ CLUT with categorized RGB values.

6 Validation of the color space categorization

The Color LookUp Table (CLUT) as described in the previous section, assigns all possible colors to one of the 11 color categories. To verify the internal validity of the color space categorization, a validation scheme was executed, which comprised two tests: (i) categorization of non-fuzzy color markers and (ii) categorization of fuzzy color markers. The segmented color space is valid if and only if it assigns both types of color markers to the same color categories as the subjects did.

As defined in Section 2, the non-fuzzy color markers are those colors categorized consequently to one color category by the participants of the experiments. The fuzzy color markers are those colors categorized to two or more categories, each by at least 10 subjects.

The assignment of the color markers can be influenced by a broad range of factors; e.g., environmental factors, system settings, and personal preferences [6,8]. In particular, the latter is the case for the fuzzy color markers. Despite the complexity of the categorization of these color markers, the subset of categories these markers are assigned to are perceptually closely related [2,6]. Since the segmented color space models human color categorization, these color categories should be each others neighbors in the segmented HSI color space.

The color space segmentation was founded on the non-fuzzy color markers. Therefore, the correct classification of these markers functioned merely as an additional check of the implementation. Not surprisingly, all non-fuzzy color markers were classified correctly, indicated with Cs in Table 4. Subsequently, all fuzzy color markers were classified using the color space segmentation. Each of the fuzzy color markers was assigned to one of its possible color categories, as is shown in Table 4 with Ns and Cs. Hence, the color space segmentation mimics human color categorization and correctly classifies both more prototypical colors (i.e., the non-fuzzy color markers) and the colors over which no consensus is among people (i.e., the fuzzy color markers).

7 Discussion

We have explained our approach toward color analysis, which exploits human perception instead of mere image processing techniques. The importance of the 11 color categories (or focal colors) is discussed and sustained by a question and answer and by two experiments. The resulting experimental data (i.e., color markers) is used as input for a coarse color space segmentation process.

The HSI color space is segmented, using two 2D projections of the HSI color space on which the recently developed Fast Exact Euclidean Distance (FEED) transform for multi class data is applied. The segmented HSI color space is transformed to a Color LookUp Table (CLUT) for the 11 color categories. We will now discuss its flexibility, a result of its modular implementation, followed by other issues of concern.

The advantage of the color space segmentation method as proposed is threefold: (i) It is based on human color categorization (ii) segmentation can be done on scarce data, and (iii) it is easily adapted to other, application, task, and/or user dependent color categories. Using this segmented color space (or better the CLUT) as a quantization scheme for image retrieval, two more advantages [7] can be mentioned: (i) it yields perceptually intuitive results for humans and (ii) it has a low computational complexity.

Color categories, as used in the current research, have also proved to be very useful in research toward color constancy [6]. Among other things, Hansen, Walter, and Gegenfurtner [6] propose to "infer color constancy from the boundaries between color categories". So, an executable model of human color categorization could aid research toward another intriguing features of human color perception: color constancy.

The flexibility of the model is illustrated by the possibility to do a smooth personalization of the generic color space segmentation through a quick calibration. This is easily done by utilizing personal judgments of fuzzy color markers that can be used to redefine the boundaries between the color categories. Through the standard processing steps, this finally results in a tailored CLUT. More generally, three other modifications of the processing scheme can be applied easily: i) another set of colors can be incorporated, which is user (e.g., in the case of color blindness), task, or application specific, ii) Instead of the HSI color space, each arbitrary color space can be used [2,6,11], which can be achieved simply by incorporating another conversion scheme in the process; and iii) The FEED transform can be replaced by an arbitrary alternative transform; e.g., [4,13].

In general, the indirect segmentation of the 3D color space (by using 2D representations) can be considered as an disadvantage of the current processing scheme. Nevertheless it had two benefits for the segmentation process: (i) a limited number of color markers is sufficient and (ii) its computational costs are much lower. In our opinion this strategy, given the scarce categorization data that was available, lead to the best possible result. In some other setting, a 3D implementation of FEED could possibly limit the number of errors in the end result. In general, the choice whether or not to adopt a direct 3D distance transform is founded on the trade-off between the available color markers, precision, and speed.

The human color categorization model as described in the current paper has been benchmarked with various color quantization schemes and distance measures [14]. Moreover, since 2005, the model of human color categorization serves as the foundation of the online color-based image retrieval system `http://www.M4ART.org` [15]. M4ART contains over 30,000 photos, mainly of pieces of art, provided by the Rijksmuseum (National museum of the Netherlands) and ArtStart.nl, which coordinates the Dutch art rental centra.

In general, it is our belief that the combination of human perception and statistical image processing will improve the performance of data-mining systems that rely on color analysis. Moreover, such a combination enables us to eventually bridge the semantic gap present in automated color analysis. From this perspective, an unique color space segmentation is proposed, which is generic, computationally inexpensive, and easy to tune or to personalize.

Acknowledgments

This research was supported by the Netherlands Organization for Scientific Research (NWO) under project number 634.000.001. We thank Leon van den Broek, Frans Gremmen, Thijs Kok, Makiko Sadakata, Louis Vuurpijl (Radboud University Nijmegen), Merijn van Erp (Planon B.V.), Maarten Hendriks (MEDOX.nl), and Eva van Rikxoort (Image Sciences Institute) for their support.

References

[1] I. J. Cox, M. L. Miller, S. M. Omohundro, and P. N. Yianilos. PicHunter: bayesian relevance feedback for image retrieval. In W. G. Kropatsch, Y. Aloimonos, and R. Bajcsy, editors, *Proceedings of International Conference on Pattern Recognition*, pages 361–369. Vienna, Austria, August 1996.

[2] G. Derefeldt, T. Swartling, U. Berggrund, and P. Bodrogi. Cognitive color. *Color Research & Application*, 29(1):7–19, 2004.

[3] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by Image and Video Content: The QBIC system. *IEEE Computer*, 28(9):23–32, 1995.

[4] C. Fouard, R. Strand, and G. Borgefors. Weighted distance transforms generalized to modules and their computation on point lattices. *Pattern Recognition*, 40(9):2453–2474, 2007.

[5] Th. Gevers and A. W. M. Smeulders. Color based object recognition. *Pattern Recognition*, 32(3):453–464, 1999.

[6] T. Hansen, S. Walker, and K. R. Gegenfurtner. Effects of spatial and temporal context on color categories and color constancy. *Journal of Vision*, 7(4):2, 1–15, 2007.

[7] F.-D. Jou, K.-C. Fan, and Y.-L. Chang. Efficient matching of large-size histograms. *Pattern Recognition Letters*, 25(3):277–286, 2004.

[8] P. Kay and T. Regier. Language, thought and color: recent developments. *Trends in Cognitive Sciences*, 10(2):51–54, 2006.

[9] T. Lin and H.J. Zhang. Automatic video scene extraction by shot grouping. In A. Sanfeliu and J. J. Villanueva, editors, *Proceedings of the 15th IEEE International Conference on Pattern Recognition*, volume 4, pages 39–42, Barcelona, Spain, 2000.

[10] V. A. Petrushin and L. Khan. *Multimedia Data Mining and Knowledge Discovery*. Springer-Verlag: Berlin - Heidelberg, 2007. ISBN-13: 978-1-84628-436-6.

[11] R. Schettini, G. Ciocca, and S. Zuffi. *A survey of methods for colour image indexing and retrieval in image databases*. J. Wiley, 2001.

[12] Th. E. Schouten and E. L. van den Broek. Fast Exact Euclidean Distance (FEED) Transformation. In J. Kittler, M. Petrou, and M. Nixon, editors, *Proceedings of the 17th IEEE International Conference on Pattern Recognition (ICPR 2004)*, volume 3, pages 594–597, Cambridge, United Kingdom, 2004.

[13] F. Y. Shih and Y.-T. Wu. Fast Euclidean distance transformation in two scans using a 3×3 neighborhood. *Computer Vision and Image Understanding*, 93(2):195–205, 2004.

[14] E. L. van den Broek, P. M. F. Kisters, and L. G. Vuurpijl. Content-based image retrieval benchmarking: Utilizing color categories and color distributions. *Journal of Imaging Science and Technology*, 49(3):293–301, 2005.

[15] E. L. van den Broek, T. Kok, Th. E. Schouten, and E. Hoenkamp. Multimedia for Art ReTrieval (M4ART). *Proceedings of SPIE (Multimedia Content Analysis, Management, and Retrieval)*, 6073:60730Z, 2006.

[16] W3 Schools. HTML Colors, URL: `http://www.w3schools.com/html/html_colors.asp` [Last accessed on July 19, 2007].