Deep Color Semantics for E-commerce Content-based Image Retrieval

Pakizar Shamoi, Atsushi Inoue, Hiroharu Kawanaka
Department of Information Systems Management, Kazakh-British Technical University, Almaty, Kazakhstan
Department of Computer Science, Eastern Washington University, Washington, USA
Graduate School of Engineering, Mie University, Tsu, Japan

Abstract

This paper aims to develop a methodology to retrieve images based on fuzzy dominant colors expressed through linguistic descriptions. This process involves two steps: assigning fuzzy colorimetric profile to the image and processing the user query. People regard color as an aesthetic issue, especially when it comes to choosing the colors for their clothing, apartment design and other objects around. It is often quite difficult to label these colors exactly using finite set of categories. Fuzzy color model that we are proposing represents the collection of fuzzy sets providing the conceptual quantization of crisp HSI space having soft boundaries. Most online shops tend to use conventional tag-based image retrieval systems. Making use of color visual content is still not disseminated in e-commerce. Subjectivity and sensitivity of humans in color perception and bridging the semantic gap between low-level visual features and high-level concepts are major issues that we plan to tackle in this research.

1 Introduction

Nowadays, the increasing availability of huge amounts of multimedia information and the corresponding rapid growth of image databases and its users in various domains requires effective and efficient image retrieval systems for managing the visual data [Goel, 2014]. Various fields may benefit from smart content-based image retrieval (CBIR), like e-commerce, GIS, art galleries collections, medical image processing, etc. A number of CBIR systems have already been proposed, such as QBIC, MARS, Virage, VisualSEEk, PhotoBook, among others. However, no similar works have been done in e-commerce area.

This paper aims to develop a methodology to retrieve images based on fuzzy dominant colors expressed through linguistic descriptions. We employ fuzzy set theory as it is widely recognized as a toll for effective imprecision modelling in image processing [Norita, 1994; Hilderbrand and Fathi, 1999]. Since color naming is inherently imprecise, we can use fuzzy semantics of color names in the HSI (Hue, Saturation, Intensity) color space. This process involves two steps: assigning fuzzy colorimetric profile to the image and processing the user query. Proposed fuzzy sets for the hue attribute take into account the non-uniformity of color distributions. L, S attributes are also represented through linguistic qualifiers.

Section 1 is this introduction. Section 2 explains the subjective and sensitive nature of human color perception. Then, Section 3 provides an overview of current image retrieval techniques in e-commerce field. Proposed fuzzy color space and the corresponding approach for image retrieval based on this space are presented in Section 4. Based on this method an intelligent e-commerce-related system has been developed. Some examples along with experimental results are discussed in Section 5. Finally, concluding remarks are drawn in the last section.

2 Human Color Perception

Humans perceive colors in a very subjective manner. Matching human sensitivity and subjectivity in color perception is a great challenge to the research community.

What is human color perception ? Since individuals have differences in visual sensitivity, they have different color perceptions. Recently (on 27th of February) a single dress image polarized the whole Internet into two aggressive groups of people arguing whether a picture depicts a blue dress with black lace fringe or white with gold lace fringe (see Figure 1). This is a great example to explain human color perception. As we know, light enters the human eye through the lens (i.e. various wavelengths corresponding to various colors). Then, the light hits the retina in the back of our eye, exactly at a place where pigments "wake up" neural connections to the special brain part that processes those signals into an image. Our visual system tends to throw away information about the illuminant and extract information about the real reflectance. So, it automatically tries to subtract the chromatic bias of the daylight axis when a human looks at the object. The fact that the daylight varies from pinkish to blue-white (dawn and noon) makes different people see colors presented at some object differently. The trick with this image is that it hits some

Figure 1. The original disputing image is in the middle. At left it is white-balanced and seems white-gold. At right, white-balanced to blue-black. In the reality, it is blue-back.

kind of perceptual boundary. In most cases, the system works fine and differences are not so critical [Ford, 2014]. However, they still exist. That is the primary motivation to count subjectivity in color perceptions.

Color is one of the features that the humans remember the most, and we can say that it is a reflection of humans' likes and dislikes to a certain extent [Shamoi et al., 2014]. Usually, color perception involves attaching a label to a color so as to categorize it. What is more interesting, people regard color as an aesthetic issue, especially when it comes to choosing the colors for their clothing, apartment design and other objects around. It is often quite difficult to label this colors exactly using finite set of categories, like red, black, bright, etc.

All in all, color plays an extremely important role in an overall impression that some object creates on humans. We can use this notion and pay special attention to color in retrieving certain items on the web, e.g. clothing commodities.

3 Image Retrieval in E-commerce

As we are witnessing now, an online shopping has become a very easy process, with a number of different intelligent technologies simplifying the work user needs to perform to find a perfect match. One of such technologies is smart search engines, that are able to not only find some piece of apparel by its name or brand, but also by its color, like "deep red" or, for instance, by its aesthetic category, like "elegant" or "romantic". Nevertheless, machine itself does not understand aesthetics or colors in the same way as humans do - it only retrieves the results with corresponding tags specified by humans [Sherman and Price, 2003]. For example, if a webpage contains, for example, the word "bright", it will be returned as a result for the search query "bright shirts". The obvious drawback of this approach is that, e.g. some bright shirts, that are not indexed, will never be included into the search results.

Another great technology that has made our online shopping easier is recommendation systems. For example, when user tries to find some commerce related item using Google search engine, system uses user's preferences, profile information and interaction history to provide recommendations for the object and can re-rank search results accordingly. In addition, some online stores, like Amazon, deploy item-to-item correlation recommender systems based on purchase data [Schafer, 1999]. What it means is that if user added, for example, a dress into his shopping cart, and many other people, who bought this dress, purchased some purse, the system will make a proposition to user to co-purchase this purse along with the dress. However, such systems are based entirely on user input data. Many other combinations of matching apparel pieces may exist (e.g. based on a color harmony), but machine will never know it without human intervention.

Most online shops tend to use conventional text-based image retrieval (TBIR) systems, in which items are retrieved from the database based on the given tags, keywords or text annotations. In contrast, CBIR makes use of visual content, like color, texture, shape, to fetch images from databases. Although it is a popular technique nowadays (e.g. Google search by image) it's still not disseminated in e-commerce. Even the most popular shopping portals, like Amazon, eBay, Taobao, etc. do not provide color indexing, so color names and other text-based descriptions are treated as common keywords/tags. The reason for that is the problem of semantic gap, i.e. very often user query requirements and capabilities of the retrieval system mismatch [Goel, 2014]. Although there are some attempts to solve these problems, like adoption of an ontology for unifying the semantically close terms, they are not sufficient to find the semantic connections between concepts. For instance, it is difficult for current systems to find out that deep red is more similar to crimson rather than to Turkish red.

Users of such IR systems express their query using the text containing the keywords. Therefore, reliable operation of such systems makes it necessary to understand the correspondence, or mapping, between the content and text. To be more specific, between the linguistic terms and colors. There are 2 problems here. Firstly, sometimes users are not able to express their intentions precisely using the text. Secondly, retrieval system lacks the understanding of the query.

We believe that image retrieval systems need to use deeper semantics defined on certain color space. Subjectivity of humans and managing the correspondence between low-level visual features and high-level semantic content are major issues that we plan to tackle in this research.

4 Fuzzy Color Modelling

As we have explained in previous sections, image indexing and retrieval by color schemes are very important in finding certain items, e.g. online clothing search. However, simple keyword matching and semantic mediations suffer from the semantic gap [Nachtegael et al., 2007]. So, we need much more deeper semantics of colors to bridge this gap.

4.1 Fuzzy Color and Fuzzy Color Space

Fuzzy color space that we are proposing represents the collection of fuzzy sets providing the conceptual quantization with soft boundaries of crisp HSI color space (FHSI). HSI space is convenient for our purposes, but one problem with this color space is that it is not a uniform color space. So, humans don't perceive small variations of hue, especially when color is green or blue. Fuzzy set is powerful in modeling this non-uniformity. Indeed, we can easily solve this problem by using trapezoidal membership functions for such kind of hues having a wide interval. Another option was to use some uniform color space, like the one from CIE* family, but they have some limitations that are not acceptable for our method (e.g. intensive computations for the conversion from and to RGB space, difficulties in getting a tone modifier [Younes et al., 2006]).

In computer systems, colors are usually represented as a triplet of numbers corresponding to coordinates in a certain color space. In turn, fuzzy color is a fuzzy subset of points of some crisp color space [Soto-Hidalgo, 2013], which is HSI space in our case. Let D_H, D_S, D_I be domains of the H, S, I attributes respectively.

Definition 4.1 *FHSI (fuzzy HSI) color C is a linguistic label whose semantic is represented in HSI color space by a normalized fuzzy subset of* $D_H \times D_S \times D_I$.

From the above definition it is obvious that for each fuzzy color C there exist at least one representative crisp color whose membership to *C* is 1. Now let's extend the concept of fuzzy color to a concept of a fuzzy color space.

Definition 4.2 *FHSI (fuzzy HSI) color space is set of fuzzy colors that define a partition of* $D_H \times D_S \times D_I$.

Table 1 below shows the information about each fuzzy variable in our color space, like term set, domain and universal set.

Fuzzy Variable	Fuzzy Sets	Domain	Universal Set
Hue	{"Red", "Orange", "Yellow", "Green", "Cyan", "Blue", "Violet", "Magenta"}	X = [0, 360]	U = {0 ,1, 2, ..., 359, 360}
Saturation	{"Low", "Medium", "High"}	X = [0, 100]	U = {0, 1, …, 99, 100}.
Intensity	{"Dark", "Deep", "Medium", "Pale", "Light"}	X= [0, 255]	U = {0, 1, … 254, 255}

Table 1. Description of fuzzy attributes of the proposed FHSI color space.

We limit the hue variable to nine fundamental colors. This choice is very intuitive since it approximately corresponds to the rainbow colors. The membership functions for the Hue variable were constructed based on results we obtained from the survey based on human cognition (Direct rating method). The resultant membership functions can be seen in Fig. 2. We did it in our previous works, please refer to [Shamoi et al., 2014] for more detailed explanation.

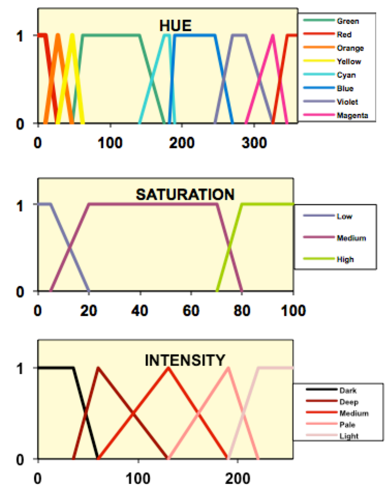


Figure 2. Fuzzy sets for each attribute of FHSI color space.

This simple method allows us to directly model colors such *dark blue* or *bright red*.

We developed fuzzy color space and the main motivation for that is that indistinguishability is a fuzzy concept for humans, since, to a certain degree, colors are indistinguishable for us. That is why crisp boundaries are counterintuitive for us.

4.2 Similarity Between Fuzzy Colors

Literature suggests a number of ways to find the similarity degree between two colors. Among them are: doing comparison based on RGB values (simple, but not precise and effective, since RGB space is not perceptually uniform), based on hue values only, LAB measure etc.

Actually, H, S and I attributes do not have the same importance when human judges how similar two colors are. For instance, colors of the same nuance (i.e. having the same S and I attributes) have small distance, although they are not similar for human visual system.

In [Shamoi et al., 2014] we proposed the mechanism to evaluate the perceptual difference (and similarity, respectively) between fuzzified color descriptions. We also provided objective measures for expressing the image similarity in a way that matches human evaluation. Our formula takes into account the notion that different hues have various value ranges, according to linguistic conventions of the society (e.g. green color).

So, for two *H* values, H_1 and H_2 , corresponding to F_1 and F_2 fuzzy sets, we find the minimum among absolute differences in H_1 and H_2 membership to F_1 and F_2. In other words, we find the hue which is closer to both values. Then we subtract this value from 1 to get the resultant coefficient:

$$a = \begin{cases} 1 - \min\left(\left|\mu_{F_1}(H_1) - \mu_{F_1}(H_2)\right|, \left|\mu_{F_2}(H_1) - \mu_{F_2}(H_2)\right|\right), & \mu_{F_1}(H_2) \neq 0, \mu_{F_2}(H_1) \neq 0 \\ 1 - \left|\mu_{F_1}(H_1) - \mu_{F_1}(H_2)\right|, & \mu_{F_1}(H_2) \neq 0 \text{ and } \mu_{F_2}(H_1) = 0 \\ 1 - \left|\mu_{F_2}(H_1) - \mu_{F_2}(H_2)\right|, & \mu_{F_1}(H_2) = 0 \text{ and } \mu_{F_2}(H_1) \neq 0 \\ 1, & \text{otherwise} \end{cases}$$

We use this coefficient only in case $F_1 = F_2$ or F_1 and F_2 are neighboring fuzzy sets, and both H_1 and H_2 should have non-zero membership to either F_1 or F_2. So, the resultant perceptual difference d_p is:

$$d_p(H_1, H_2) = d(H_1, H_2) * a$$

In addition, the S attribute works as a weighting factor for the intensity and hue. Specifically, when we compare two colors, we need to remember that:

- If (*S is high*) *H* is more important
- If (*S is low*) *I* is more important

Finally, the perceptual difference between two chromatic colors $C_1(H_1, S_1, I_1)$ and $C_2(H_2, S_2, I_2)$ in FHSI color space can be found using the formula below (the normalized values are denoted with upper-cased D):

$$D_p(C_1, C_2) = \\ = \frac{D_p(H_1,H_2) * \left(1 + \mu_{high}(S_1) + \mu_{high}(S_2)\right) + D_p(S_1,S_2) + D_p(I_1,I_2) * (1 + \mu_{low}(S_1) + \mu_{low}(S_2)}{3}$$

Although various color similarity measures were already proposed in many research works, little research was done on how to identify colors which are in a harmony**.**

4.3 Harmony Between Fuzzy Colors

Visually, harmony is something that is pleasing to the eye. When people speak about color harmony, they are evaluating the joint effect of two or more colors. Experience and experiments with subjective color combinations show that individuals differ in their judgments of harmony and discord. The color combinations called “harmonious” in common speech usually are composed of closely similar chromos (e.g. tones, tints and shades), or else of different colors of the same nuance [Itten, 1973]. They are combinations of colors that meet without sharp contrast.

There exist a number of conventional rules of defining harmonious colors. It is generally accepted that the human eye is satisfied or in equilibrium, only when the complementary relation is established. Two or more colors are mutually harmonious if their mixture yields a neutral gray. Any other color combinations, the mixture of which does not yield gray, are expressive or discordant. The most popular conventional techniques for combining colors based on the color wheel are represented in Table 2 below. For more information you can refer to [Itten, 1973].

Harmony is one of the most interesting and intriguing principles in color theory, which is still raw. The problem with harmony is that it is a very complex notion, with many factors having an impact on it, including affective, cognitive and contextual ones. These factors determine how certain individual perceives certain color in a certain situation or context. Therefore, color harmony is very difficult to predict. In the sample application we describe below we use color harmony principles in order to find the set of apparels that fit to a certain apparel that the user inputted to a system.

Table 3 presents color harmony groups proposed by our research team. These groups were obtained based on the deep analysis of basic principles of color theory and fashion images. In our system, for each possible color we need to be able to identify group(s) of colors with which this particular color is in a harmony. This is the case where we need to be able to identify the very similar color with the inputted color.

We checked the competence of the proposed and traditional harmony groups, along with the other methods by conducting an online survey, which is based on a Polling method. The survey was intended to gather people’s opinions on a harmony of various color combinations. The results of this questionnaire concerning the conventional color schemes are presented in Figure 3.

Overall results demonstrating the proportion of people who gave positive feedback on certain color combinations are presented in Figure 4. Average result for harmonies proposed in this paper is 0.45 , for the ones proposed by colorstudio.com – 0.16 , for the traditional ones – 0.11. So, traditional rules for combing colors are not actual now.

5 System Description and Results

Most e-commerce web sites use textual descriptions to provide the color information of clothing items. Although this represent quite useful technique, the descriptions must be made by humans. Labeling items manually is very expensive, time-consuming and seems to be infeasible for many applications. It requires huge amount of human labor in order to manually annotate large-scale image databases. Developed system aims to tackle these problems.

5.1 System Architecture

Our system combines two modalities - content (color scheme of the image) and text (linguistic query given by the user). User can form its request to a system either in a form of linguistic query (just textual description) or by uploading an image (query by example). For the combinational query, we need to consider the semantics of the linguistic query and query image provided by the user as well.

The software has two logical phases – indexing and retrieval. The first one is intended for the admin and it is mainly connected with the treatment and addition of the images to the database. This involves precomputation of fuzzy dominant colors. The second part deals with the exploitation of the database through natural query processing, which can be of 3 types.

User can provide an exemplar image and perhaps specify other matching details (similarity or harmony measures). At the time image is uploaded into a database, fuzzy color scheme of the image is extracted and stored.

Today we have Windows Forms project for desktop written on C# programming language, where all the key-components we describe in this paper are fully operating. In the nearest future, we are planning to transform it into a web-based application using ASP.NET platform as a base. The web-application will have three-tier (multilayered) architecture. The presentation tier (layer) will be exposed to the end user as a usual web application, which includes

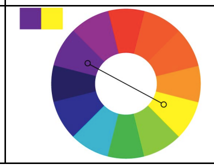
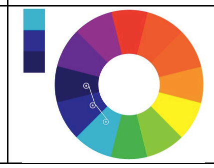
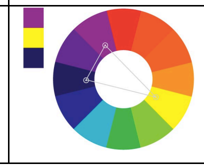
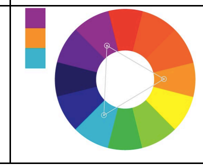
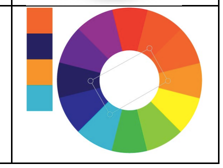
	Complementary	Analogous	Split complements	Triadic	Tetradic
Visual explanation					
Illustration					

Table 2. Conventional ways of combining colors [Itten, 1973]

Group ID	Colors in the group
1	
2	
3	
4	
5	
…	…

Table 3. Proposed groups of harmonical colors

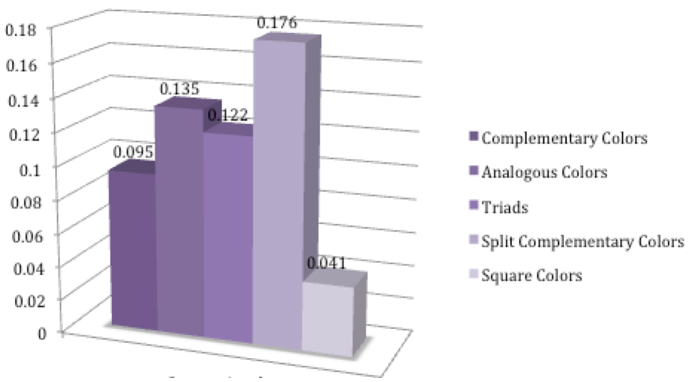


Figure 3. Results of a survey: users' opinion on conventional ways of combining colors.

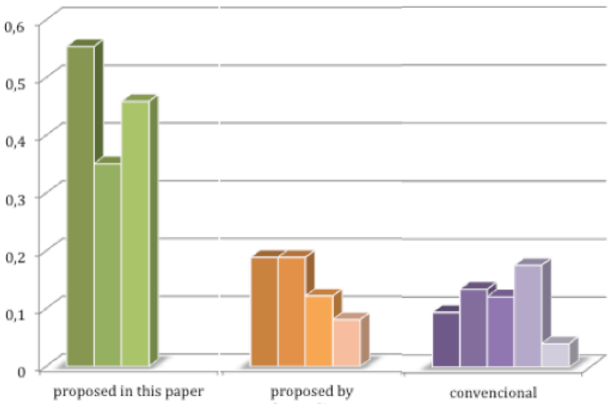


Figure 4. Results of a survey.

ASP.NET Webforms, Scripts (client-side JavaScipts), Styles (CSS). The already existing C# application will compose the Logic Tier of our project, though some additions and improvements can be made. Lastly, Data Tier and Data itself will be developed using SQL (see Figure 5). The high-level system architecture is presented on Figure 6.

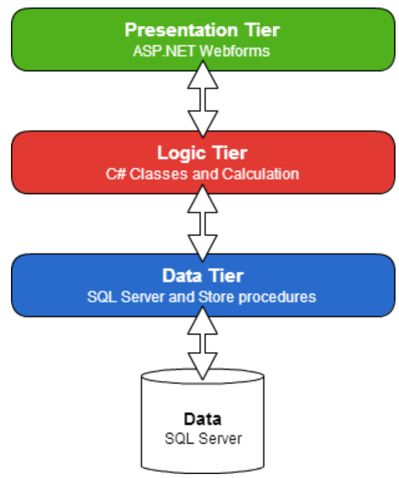


Figure 5. Technologies employed in a system

5.2 Examples

As it was mentioned, the aim of the system is to retrieve best matching apparels corresponding to the complex query posed by user based on color scheme.

Our system differs from traditional image retrieval systems in a number of aspects, like automated item description based on color schemes and natural query language.

We use the proposed method in the matching engine used for the query processing. Currently, the system supports the processing of 3 types of queries represented in Table 4 below. Note that in case of exemplar or combinational query, the given RGB image is first converted into HSI model. Next, based on the histogram, we identify the dominant color in the image and find similar apparels or apparels that fit to it. The harmony between a query image

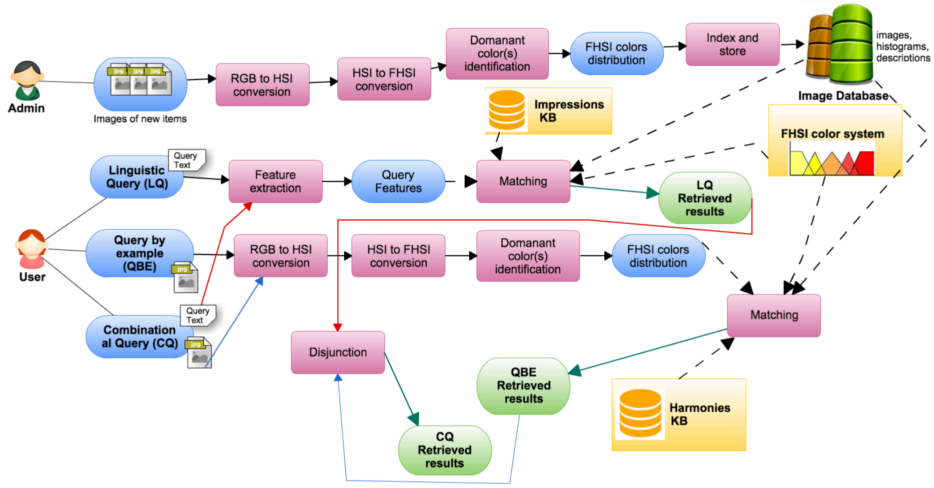


Figure 6. System Architecture

Type of query	Example
Linguistic query	*Something formal and not dark, Pale blue dress*
Query by example	Allows image input to a system. *Find similar dresses, find shoes that are combined with the dress on the uploaded photo.*
Combinational query	Has properties of both linguistic and exemplar queries. *Find apparels of dark hues that are perfectly combined with jacket on the photo.*

Table 4. Types of queries in the system

and database image is computed from the dominant color(s), using the table of color harmonies selections.

Example 1. "Deep red dress". This is a simple linguistic query. It works fast due to initial offline precomputation of apparels' color schemes. Using formulas defined above we find the constraint relations: Hue<17 or Hue>335, 49>Int.>95 (Figure 7).

Figure 7. Application screenshot for Example 1

Example 2. Query by example based on similarity metric we proposed. For example, a client has a photo with apparel (taken from fashion site or even real life) and want to find something similar similar. Figure 8 below demonstrates this example.

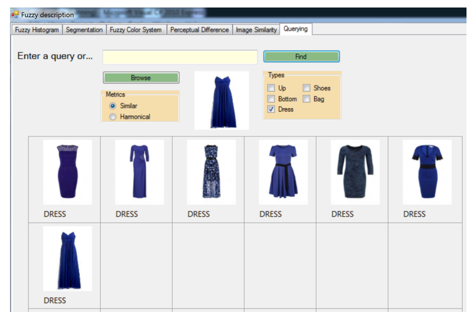


Figure 8. Application screenshot for Example 2

Example 3. Query by example based on a harmony metric, e.g. a client already has a skirt and wants to buy the remaining apparels – blouse and shoes. User needs to upload the skirt image and the system will extract the its dominant colors and fetch such apparels whose dominant colors are in a harmony with the skirt's one (Figure 9).

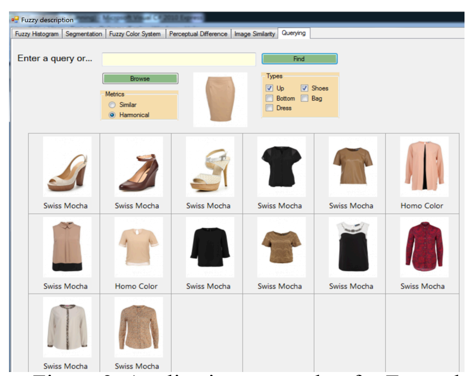


Figure 9. Application screenshot for Example 3

It is critical to note that generally, image retrieval which is founded on color analysis solely may bring too many false positives when database is large. That is why, usually, color-based features are integrated with other visual features, and the corresponding module can work as a subsystem within big retrieval system.

However, it is not critical particularly for this specific kind of system (apparel coordination), due to simple nature of images having light background with an apparel in the middle**.**

6 Conclusion

In this paper we have shown that fuzzy color processing can be very helpful for certain tasks that are beyond the capabilities of systems which are based on standard tag-based image database. Fuzzy approach allows us to define query conditions on the basis of linguistic terms, which is more natural way for a user to express his desire.

We believe that the proposed method can help to reduce the semantic gap between high-level concepts(e.g. elegant, harmonical, etc) and low-level features (colors). The preliminary experimental results we obtained from the survey on color harmony and in the prototype clothing search system have shown the strength of the proposed method.

As it was mentioned, color perception is usually very subjective, so in the future we plan to develop a Method for adapting the items retrieval to the user sensibility. We will implement this by collecting relevance judgments on the correspondingly retrieved clothing items and modelling of user's relevance feedback. System performance will be evaluated in terms of precision and recall values. Learning from users feedbacks will help us to better satisfy user needs.

The methodology can be applied to any fields in which matching color descriptions based on their fuzzy semantics is important.

References

[Ford, 2015] Dana Ford. The official color of the dress is ... 'cnn.com' News web site. February 27, 2015. [http://edition.cnn.com/2015/02/26/us/blue-black-white-gold-dress/]

[Goel, 2014] Nidhi Goel. Image Retrieval Using Fuzzy Color Histogram and Fuzzy String Matching: A Correlation-Based Scheme to Reduce the Semantic Gap. *Intelligent Computing, Networking and Informatics, Advances in Intelligent Systems and Computing*, DOI: 10.1007/978-81-322-1665-0_31, Springer India 2014.

[Hildebrand and Fathi, 1999] Hildebrand L. and Fathi M. Soft Computing as a Methodology for Color Processing. *Proceedings of the EUSFLAT-ESTYLF Joint Conference*, Palma de Mallorca, Spain, 1999.

[Itten, 1973] Johannes Itten. The Art of Color: the subjective experience and objective rationale of color. New York: Van Nostrand Reinhold. ISBN 0-442-24037-6, 1973.

[Nachtegael et al., 2007] M. Nachtegael, S. Schulte, V. De Witte, T. M´elange, and E.E. Kerre. Image Similarity – From Fuzzy Sets to Color Image Applications. *VISUAL 2007, LNCS 4781*, pp. 26–37, Springer-Verlag Berlin Heidelberg 2007.

[Norita, 1994] Toshio Norita. Fuzzy Theory in an Image Understanding Retrieval System. Applied Research in Fuzzy Technology International Series in Intelligent Technologies, Volume 1, pp 215-251, 1994.

[Schafer et al., 1999] Ben Schafer, Joseph Konstan and John Riedl. Recommender Systems in E-Commerce. *Proceedings of the 1st ACM conference on Electronic commerce*, pp.158-166, 1999.

[Shamoi et al., 2014] Pakizar Shamoi, Atsushi Inoue and Hiroharu Kawanaka. Perceptual Color Space: Motivatins, Methodology, Applications. *Proceedings of SCIS & ISIS 2014, Japan, December 2014.*

[Shamoi et al., 2014] Pakizar Shamoi, Atsushi Inoue and Hiroharu Kawanaka. Fuzzy Color Space for Apparel Coordination. Open Journal of Information Systems (OJIS), 1(2), Pages 20-28, 2014.

[Sherman and Price, 2003] Chris Sherman and Gary Price. The Invisible Web: Uncovering Sources Search Engines Can't See. *Library Trends* Vol. 52, NO. 2, pp.282–298, 2003

[Soto-Hidalgo et al., 2013] J. M. Soto-Hidalgo, J. Chamorro-Martínez, P. Martínez-Jiménez, Daniel Sánchez. Ranking Images Using Customized Fuzzy Dominant Color Descriptors. *Flexible Query Answering Systems*. Lecture Notes in Computer Science Volume 8132, 2013, pp 198-208.

[Younes et al., 2007] Amine Aït Younes, Isis Truck and Herman Akdag. Image retrieval using fuzzy representation of colors. *Soft Computing* (2007) 11: 287–298. Springer-Verlag 2006.