Hybrid Color Model for Image Retrieval Based on Fuzzy Histograms

Vedran Ljubovic*
University of Sarajevo
Faculty of Electrical Engineering Sarajevo

Haris Supic†
University of Sarajevo
Faculty of Electrical Engineering Sarajevo

Abstract

A hybrid color model is a color descriptor formed by combining different channels from several other color models. In computer graphics applications such models are rarely used due to redundancy. However, hybrid color models may be of interest for Content-Based Image Retrieval (CBIR). Best features of each color model can be combined to obtain optimum retrieval performance.

In this paper, a novel algorithm is proposed for selection of channels for a hybrid color model used in construction of a fuzzy color histogram. This algorithm is elaborated and implemented for use with several common reference datasets consisting of photographs of natural scenes. Result of this experimental procedure is a new hybrid color model named HSY. Using standard datasets and a standard metric for retrieval performance (ANMRR), it is shown that this new model can give an improved retrieval performance. In addition, this model is of interest for use in JPEG compressed domain due to simpler calculation.

CR Categories: H.2.4 [Systems]: Multimedia databases—; H.3.3 [Information Search and Retrieval]: Retrieval models—; I.4.7 [Feature Measurement]: Feature representation—Invariants

Keywords: content-based image retrieval, fuzzy histograms, color spaces, query-by-example

1 Introduction

Content-Based Image Retrieval (CBIR) is an area of science that deals with searching large image databases based on visual similarity of images, rather than just image descriptions and tags. Certain visual feature is programmatically extracted from each image and used for comparison.

One of the oldest and best known features of the kind are color histograms. While histograms give excellent results in retrieval experiments, in practical applications they show a number of well known shortcomings [Pavlidis 2008; Rubner et al. 2000; Stricker and Orengo 1995]. In recent years those problems were successfully addressed by using fuzzy processed color [Chatzichristofis et al. 2010]. Each pixel is fuzzy-matched into one of a small number of colors, and then a histogram of such colors is formed. More recently, these fuzzy color histograms have successfully been used in context of the "visual-bag-of-words" approach [Chatzichristofis et al. 2013].

A large improvement in performance of retrieval based on color histograms can be obtained by first converting the image into an appropriate color model [Ljubovic and Supic 2013b]. CIE L*a*b* model is perceptually uniform and therefore should be optimal for extracting visual features such as color histograms [Konstantinidis et al. 2005]. However, its calculation is costly and therefore most authors use HSV model as it gives the best results [Chatzichristofis et al. 2010; Ljubovic and Supic 2013b]. HSL color model may have some benefits over HSV [Ljubovic and Supic 2013a].

This opens the question: is it possible to construct a novel color model specifically for image retrieval by using the best features of existing models? Over the years, authors have proposed a number of color models specifically intended for the use in image retrieval, including HMMD [Manjunath et al. 2001] and IHLS [Hanbury and Serra 2002].

When using the fuzzy matching approach described by [Chatzichristofis et al. 2010], the problem with HSV model is that pixels cannot successfully be classified as "white" by observing just one component (V channel), one needs to observe both S and V. HSL removes this problem by introducing the lightness (L) channel. However, the modified formula for S channel presents a new problem for classification of gray pixels [Ljubovic and Supic 2013a].

In this paper we will present an approach to constructing a hybrid color model for image retrieval. We will give an algorithm for training and selection of components of such color model. Through a number of experiments we will then construct an optimal color model for image retrieval based on fuzzy color histograms through a selection of standard image collections and, using standard retrieval metrics, demonstrate that this model gives good retrieval performance.

Remainder of this paper is structured as follows: Section 2 gives a description of image retrieval based on fuzzy color histograms. Section 3 describes the used experimental setup and comparison metrics. In Section 4 we propose an algorithm for construction of a hybrid color model and implement this algorithm. Section 5 contains an experimental comparison of our new retrieval feature with state-of-the-art retrieval methods found in literature with a brief discussion. Finally, Section 6 gives some concluding remarks and directions for future work.

2 On Fuzzy Color Histograms

The basic approach to the construction of fuzzy color histograms outlined in [Chatzichristofis et al. 2010] is as follows: For each pixel, a trapezoid fuzzy matching function is applied to one of its color channels to determine membership of that pixel to one of the color bins. For example, in HSV model, a fuzzy membership function similar to the one on Fig. 1 would be applied to V channel to determine that pixels' membership to black bin. The output of this function is a value in range [0,1] where 1 means that a pixel is fully black while 0 means that it's not at all black. In case that this value is less than 1, the matching continues to white, gray and other colors. Finally, the histogram bin that corresponds to each color will

*e-mail:vljubovic@etf.unsa.ba
†e-mail:hsupic@etf.unsa.ba

be incremented by this colors' membership value.

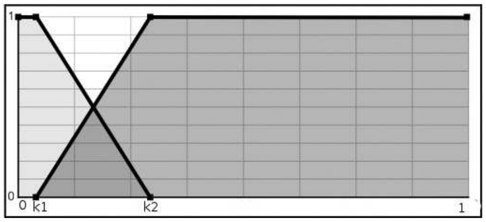

Figure 1: Simple trapezoid fuzzy matching function used in construction of fuzzy color histogram [Chatzichristofis et al. 2010].

The ultimate result of this process is a color histogram with a very small number of bins, which therefore uses a low amount of memory and resolves a class of problems that are commonly referred to as "the curse of dimensionality". Fuzzy processing ensures that pixels that have only slightly different color e.g. due to changes in lightning are matched to the same bins. Paper [Ljubovic and Supic 2013a] describes a fuzzy color histogram in HSL color model with 10 bins that correspond to colors given in Fig. 2.

Figure 2: Palette of colors used in a fuzzy histogram with 10 bins [Ljubovic and Supic 2013a].

When observing Fig. 1 we note that the given trapezoid function is characterized by two values we can call k_1 and k_2. For each of the fuzzy functions in our retrieval system, values k_1 and k_2 can be varied to obtain optimum retrieval performance.

2.1 Description of Chosen Retrieval Feature

The starting point of this paper will be a fuzzy color histogram in HSL model with 10 bins, where 3 bins correspond to black, white and gray, while remaining 7 bins correspond to illumination and viewpoint invariant colors. Referenced literature suggests that such feature results in optimal retrieval given its small size.

A large number of experiments with color histograms [Ljubovic and Supic 2013b] show that color models based on polar coordinates such as HSL are optimal for a histogram with small number of bins. The explanation is that a suitable quantization of Hue channel encodes color that is invariant to illumination and viewpoint, while saturation and brightness can be used to detect black, white and gray pixels. The issue examined in this paper is whether such histogram can be further improved by combining channels from several color models.

Through analysis of literature on image retrieval based on histograms we have selected the following commonly used color models: RGB, HSV, HSL, HSI, HMMD, IHLS, YUV and CIE L*a*b*. It should be noted that all of those models have formulas for efficient conversion from the usual RGB model except for CIE L*a*b* (usually referred to as Lab) which requires complex calculations. Those formulas are given in an appendix to this paper.

Observation of these formulas shows that HSV, HSL, HSI, HMMD and IHLS use the same formula for calculation of hue i.e. the hue channel in all those models is the same. Meanwhile, RGB, YUV and Lab lack a hue channel and therefore it is impossible to calculate color bins in a described way. Research described in paper [Chatzichristofis et al. 2010] resulted with optimal fuzzy classification for hue that cannot be further improved. Therefore the remainder of this paper will focus on classification of black, white and gray, which in our experiments were shown to be crucial for retrieval performance, while hue will be classified as in [Chatzichristofis et al. 2010].

As in [Ljubovic and Supic 2013b], two fuzzy color histograms are compared using Matsushita distance with formula:

$$d = \sqrt{\sum_i \left(\sqrt{H_i} - \sqrt{Q_i}\right)^2}$$

2.2 Classification of Black, White and Gray

The problem of classification of pixels that belong to objects whose natural color (i.e. color under normal illumination) is black, white, or gray is ill-posed. For example, given a black pixel on an image, it isn't possible to know wheter this pixel belongs to a colored object under poor illumination or an object whose natural color is black (Fig. 3). This information can't be obtained through algorithmic manipulation of pixels, it is lost. Retrieval methods should be robust with respect to this classification error, for example through use of color histograms with an appropriate distance metric [Deselaers et al. 2008; Rubner et al. 2000] such as Matshushita distance [Ljubovic and Supic 2013b].

Figure 3: In above image several regions are marked with red circles that are encoded with colors close to black, but correspond to objects of different natural colors under poor illumination.

This reasoning is well known since the earliest days of image retrieval.

However, a good classification of black, white and gray pixels is still important since such pixels represent noise in the color (hue) histogram. For example, in HSV color model we know that black pixels should have Value channel equal or close to zero. Conversion formulas (in appendix) show that for such pixels C is small, leading to a large rounding and quantization error in calculation of hue. In real-life bitmap images, color is quantized to integer (byte) values and further error is introduced due to inexact conversion from RGB to HSV model, image compression or other reasons. For black pixels (where $V \approx 0$), H in such images is essentially random and should be discarded (Fig 4.). Similar reasoning applies to gray and white pixels.

Figure 4: Image above is a photograph from Wang dataset with a large percentage of black pixels. Below the image is transformed by setting saturation and value to maximum while leaving hue unchanged. We see that in the black colored regions hue channel does not carry information useful for image retrieval.

If we observe a color histogram in HSV model with 32 bins formed by uniform quantization of Hue and Value to 4 levels, and Saturation to two levels, we see that 8 histogram bins correspond to pixels that are perceptually black or very dark, while there is no bin corresponding to white color and thus white (or very light) pixels will be randomly distributed into another 4 histogram bins. Similar problem is present in some form in all known color models.

One approach to black, white and gray classification is to form sharp cutoff levels in Saturation and Value channels (used e.g. in [Androutsos et al. 1998]). Pixels that are below or above those levels are classified separately as black, white or gray, while remaining pixels are further processed. This however introduces another problem: the relationship between saturation and value is nonlinear and is dependent on hue. For example, in HSV model, a certain set of (H,S,V) will be perceived as black and uncolored by a human observer, but changing just the hue value results in a pixel that is perceived as colored. Similar nonlinearities in HSL model are described in [Ljubovic and Supic 2013a].

A promising approach to resolving this problem is fuzzy color processing with Multiple Participant defuzzyfication algorithm. In this approach each of color bins is incremented by "pixel color clarity" which is a fuzzy value. This provides some degree of robustness against nonlinear relationships in the color space, but even better retrieval could be obtained by using a color model that doesn't have such nonlinear relationships.

3 Experimental methodology

Use of standardized, annotated datasets is an important issue in CBIR [Müller et al. 2002]. To this purpose we used three datasets that are often cited in literature: Wang SIMPLIcity [Wang et al. 2001], UW [uwd 2013] and UCID [Schaefer and Stich 2003]. Each of these datasets consists of a collection of images (photographs from travel and natural scenes) and a set of queries with expected results. Further, each of these datasets represent different usage scenarios in the broader class of photography, and have apparently been processed with different digital filters. This ensures that our training steps model a broad range of human perception.

To avoid the problem of overfitting, we have split these datasets into training set and evaluation set. Training set consists of 20 queries from Wang dataset (out of 1000), 10 queries from UCID dataset (out of 264) and 20 queries from UW dataset (out of 290). The training queries have been selected in a way to be broadly representative of the whole dataset. For example, Wang dataset consists of images in 10 categories so we have chosen two queries in each category such that the two queries are mutually dissimilar.

Here it should be noted that the UCID dataset suffers from a large problem of generality. Many sample queries have only one defined result or very rarely succeed in retrieving all results. The consequence is that minor changes in retrieval method can have apparently large effect on overall performance. This is the reason we use just 10 queries from this dataset, and those 10 queries were selected such that queries with a larger number of results are represented.

For comparison between retrieval methods we have used ANMRR metric [Manjunath et al. 2001] which is in recent times predominantly used in literature for this purpose.

Implementation of retrieval method and ANMRR metric used in this paper is released as an open-source tool implemented in C++. This implementation as well as the specification of training set can be downloaded from the website accompanying this paper.[1]

4 Construction of a Hybrid Color Model

In this section we will propose a high-level algorithm for construction of a hybrid color model for image retrieval based on fuzzy color histograms, based on the discussion given in Section 2, then we will implement this algorithm using experimental procedures given in Section 3.

The high-level algorithm for construction of a hybrid color model is as follows:

1. Form lists of candidate channels for classification of black, white and gray. A candidate channel in this context is a single channel from each of mentioned color models e.g. S_{HSV}.
2. Tune parameters k_1 and k_2 for each candidate channel by substituting this channel into the histogram based feature described in [Ljubovic and Supic 2013a].
3. Measure performance of each candidate channel and find the best performers for black, white and gray classification.
4. Construct a hybrid color model from previously selected channels and hue (H) channel.
5. Further refine k_1 and k_2 in context of this new hybrid color model.

We start with a list of candidate channels for each of black, white and gray. Those lists are given in Table 1. Of course, for black and gray trapezoid fuzzy functions are used with maximum value (1) in zero, whereas for white a trapezoid function is used with maximum

[1] `http://people.etf.unsa.ba/~vljubovic/etfis/`

Color	Candidate channels
Black	V_{HSV}, L_{HSL}, I_{HSI}, m_{HMMD}, Y_{YUV}, L_{LAB}
White	V_{HSV}, L_{HSL}, I_{HSI}, Y_{YUV}, L_{LAB}
Gray	S_{HSV}, S_{HSL}, S_{IHLS}
Other colors	H_{HSV}

Table 1: List of candidate channels for each color bin.

value in one. m_{HMMD} refers to "minimum" component of HMMD model.

Next step in our algorithm is tuning the parameters k_1 and k_2 for each of the above features. Each of the channels given in Table 1 is substituted into the feature described in [Ljubovic and Supic 2013a] such that, for example, V_{HSV} is used for classification of black pixels, while white, gray and color bins are classified using method described in cited paper.

Values k_1 and k_2 are optimized by performing grid search on interval [0,1] with step 0.01 (of course, $k_1 < k_2$).

Full experimental results in Excel and CSV format are available at the website accompanying this paper.[2] In Table 2 we summarize the best performing values k_1 and k_2 and the ANMRR score obtained with those values (smaller ANMRR means better retrieval).

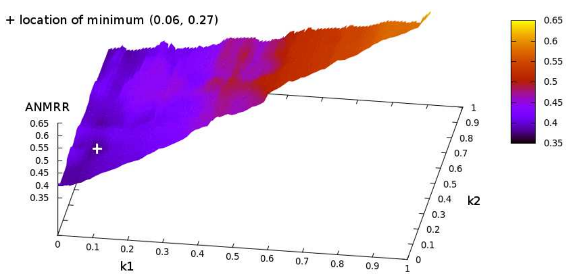

Figure 5: 3D plot of ANMRR for various values of k_1 and k_2 for black color matching using I_{HSI} channel

Here it should be noted that our experiments revealed that this function $f(k_1,k_2)$ almost always has multiple minima. In choosing the optimal values we observed not just the least compound ANMRR value but also the value where performance is best across all three datasets. This is illustrated on Fig. 5. The global minimum (0.06,0.27) is marked with a plus sign, however we notice other minima at around (0.05,0.4) and also (0,0.05) and (0.1,1.0).

From Table 2 it can be seen, as expected, that the HSV color model enables a very good classification of black and gray pixels, however it somewhat fails to classify white pixels. The reason for this is that neither maximum value of V channel nor minimum value of S channel guarantee that a pixel is white, both conditions must be satisfied. This is further illustrated in Fig. 4 that plots relationship between Saturation and Value when Hue is fixed to 0° (red). The region that can be described as "white" in a fuzzy sense is fairly small and has an irregular shape.

Table 2 shows that, compared to HSV model, a better classification of white pixels can be obtained using color models that have a "lightness" or "luminance" channel. This observed result models the fact that, as pixel lightness increases, the colors become "washed out" and the value of hue becomes less relevant for perception.

[2] http://people.etf.unsa.ba/~vljubovic/etfis/

Color	Channel	k_1	k_2	ANMRR
Black	Y_{YUV}	0.07	0.30	0.3738
	V_{HSV}	0.09	0.33	0.3743
	m_{HMMD}	0.00	0.43	0.3751
	I_{HSI}	0.06	0.27	0.3754
	L_{HSL}	0.06	0.28	0.3767
	L_{LAB}	0.06	0.34	0.3897
White	Y_{YUV}	0.60	0.88	0.3795
	I_{HSI}	0.57	0.82	0.3802
	L_{HSL}	0.67	0.86	0.3806
	V_{HSV}	0.64	0.86	0.3828
	L_{LAB}	0.54	0.84	0.3892
Gray	S_{HSV}	0.02	0.23	0.3661
	S_{HSL}	0.04	0.28	0.3711
	S_{IHLS}	0.01	0.10	0.3744

Table 2: Optimal values of k_1 and k_2 for various features.

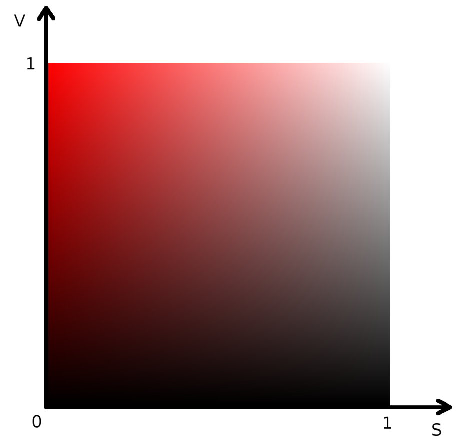

Figure 6: Plot of Saturation vs. Value (in HSV model) when Hue is fixed to 0° (red)

Given above, we conclude that the optimum classification for the three colors of interest is obtained using:

- Y_{YUV} for Black
- Y_{YUV} for White
- S_{HSV} for Gray

The next step in the described algorithm is to further tune the coefficients for each channel when used as a combined feature. As mentioned, the values in Table 2 were obtained by substitution of each feature separately into a HSL fuzzy color histogram. Now, another measurement is performed of the three selected channels in ensemble. We will again use the training set to optimize the coefficients by performing a grid search. This procedure yields the following coefficients:

- Y_{YUV} (black): $k_1 = 0.07$, $k_2 = 0.17$
- Y_{YUV} (white): $k_1 = 0.60$, $k_2 = 0.88$
- S_{HSV} (gray): $k_1 = 0.12$, $k_2 = 0.27$

We note that for each of the chosen components, the fuzzy range $k_2 - k_1$ is smaller than the one in Table 2. Classification of black, white and gray is more precise. This step can be repeated several times to obtain an even finer tuning.

Color model	Dataset (ANMRR)		
	Wang	UW	UCID
HSY	0.4219	0.2595	0.5590
HSL	0.4188	0.2601	0.5612
HSV	0.4240	0.2615	0.5701

Table 3: Comparison of the retrieval algorithm based on the new hybrid color model to the state of the art retrieval methods. HSL color model is used in paper [Ljubovic and Supic 2013a], HSV model in paper [Chatzichristofis et al. 2010], while this paper uses a hybrid of HSV and YUV color models named HSY.

Category	HSL	HSY	Change (%)
People	0.4169	0.4360	+1.91%
Beach	0.5841	0.5827	-0.14%
Monuments	0.5694	0.5777	+0.83%
Buses	0.3418	0.3275	-1.43%
Dinosaurs	0	0	0%
Elephants	0.6417	0.6399	-0.18%
Flowers	0.4758	0.4719	-0.44%
Horses	0.0934	0.0876	-0.58%
Mountains	0.5681	0.5945	+2.64%
Food	0.4878	0.5036	+1.58%

Table 4: Comparison of HSV and HSY model by categories of Wang1000 dataset. The values are ANMRR (less is better).

5 Experimental Results

In the previous section we have selected components that comprise a new hybrid color model for image retrieval based on fuzzy color histograms and determined the tuning coefficients k_1 and k_2. Next, we will test how this model performs as compared to other retrieval features known from literature. We will construct an image feature similar to the one described in [Ljubovic and Supic 2013a], however this time our new color model will be used. This image feature is a fuzzy color histogram with 10 bins corresponding to 10 basic colors, where black is classified using $\mathrm{I_{HSI}}$, white using $\mathrm{Y_{YUV}}$, gray using $\mathrm{S_{HSV}}$, and the remaining 7 colors are detected from hue ($\mathrm{H_{HSV}}$).

Experimental results are given in Table 3. This time, an exhaustive search experiment is performed on Wang, UW and UCID datasets. Lower ANMRR value means more accurate retrieval.

It can be seen that the new HSY model provides only a small improvement over other models, in particular compared to HSL model in Wang1000 dataset. While UW and UCID datasets consist exclusively of outdoors photographs, Wang dataset includes both outdoors photographs and other types of images. Therefore this result should be further examined with regards to types of queries that return good or bad results.

Wang1000 dataset consists of 1000 images in 10 categories (Table 4.), where results from the same category as the query image are considered relevant, while results from other categories are considered not relevant. In 6 of those categories HSY model shows a slight improvement over HSL just as in other datasets. However, there is a sharp decrease in performance in categories People, Mountains and Food, and a slight decrease in performance in category Monuments.

All retrieval methods based on color histograms which don't contain an element of spatial distribution of color display a poor performance in People and Food categories. Images in both categories are characterized with intensely colored details and there are often many false positive results. Classification of such brightly colored patches into "white" bin actually gives better results in this particular experiment. However, such example is not relevant for real-life cases.

Category Mountains mostly consists of photos of snow-cowered mountain tops in sunlight as well as in dusk or dawn. In fuzzy histograms with low number of bins, most pixels in these photos will be classified as white or black, and therefore such methods will give very poor retrieval. This applies to both HSL and HSY model, however HSY gives slightly worse results for reasons described above.

Proper detection of gray colored pixels is crucial for correct retrieval in Horses category, due to false positive results from Elephants category, and so in this subset we see a marked improvement in retrieval when using HSY model as opposed to HSL model.

6 Conclusion

Fuzzy color histograms are a powerful and interesting new feature for image retrieval. Recently they have been used in the context of "visual-bag-of-words" (VBoW) retrieval concept [Chatzichristofis et al. 2013]. Most research so far has used fuzzy color histograms in HSV color model. We have shown that this model is less than ideal for the purpose. We have also shown that optimal retrieval may be obtained using a combination of features from several models.

In this paper an algorithm for construction of a hybrid color model was proposed. Then, this algorithm was used to construct a hybrid color model named HSY for use with fuzzy color histograms. HSY model consists of Hue and Saturation channels from HSV model and the Luminance (Y) channel from YUV model. Through a series of experiments it was shown that the newly proposed feature gives an improvement in retrieval performance compared to previous approach in HSV and HSL color model.

Even if performance for HSY model was roughly equal to HSL or HSV, such model would still be of interest in JPEG compressed domain. Since JPEG images use a form of YUV color model named $Y'C_BC_R$, calculation of V channel would not be neccessary as Y channel is already available.

It is well known in literature (e.g. [Pavlidis 2008]) that image features that are based exclusively on color may give a deceptively good retrieval in experimental settings, however in real life applications it will perform poorly. A full retrieval solution should use the aspect of texture, shape and other in addition to color. Fuzzy color histograms are especially useful for this purpose due to their small size.

In addition to combined retrieval features, future research should focus on approach to fuzzy histogram extraction in JPEG compressed domain. In this paper we have seen that YUV color model (used in JPEG compression) can be useful as component in a hybrid color model. Solution in JPEG compressed domain can offer improved feature extraction speed and use of specific texture features that can be obtained from DCT compressed domain.

References

ANDROUTSOS, D., PLATANIOTISS, K., AND VENETSANOPOULOS, A. N. 1998. Distance measures for color image retrieval. In *Image Processing (ICIP 98). Proceedings of the 1998 International Conference on*, vol. 2, IEEE, 770–774.

CHATZICHRISTOFIS, S. A., ZAGORIS, K., BOUTALIS, Y. S., AND PAPAMARKOS, N. 2010. Accurate image retrieval based on compact composite descriptors and relevance feedback information. *International Journal of Pattern Recognition and Artificial Intelligence 24*, 02, 207–244.

CHATZICHRISTOFIS, S. A., IAKOVIDOU, C., BOUTALIS, Y., AND MARQUES, O. 2013. Co. Vi. Wo.: Color visual words based on non-predefined size codebooks. *Cybernetics, IEEE Transactions on 43*, 1, 192–205.

DESELAERS, T., KEYSERS, D., AND NEY, H. 2008. Features for image retrieval: An experimental comparison. *Information Retrieval 11*, 2, 77–107.

HANBURY, A., AND SERRA, J. 2002. A 3D-polar coordinate colour representation suitable for image analysis. *Technical Report PRIP-TR-077, TU Wien*.

KONSTANTINIDIS, K., GASTERATOS, A., AND ANDREADIS, I. 2005. Image retrieval based on fuzzy color histogram processing. *Optics Communications 248*, 4, 375–386.

LJUBOVIC, V., AND SUPIC, H. 2013. A compact color descriptor for image retrieval. In *Information, Communication and Automation Technologies (ICAT), 2013 XXIV International Symposium on*, IEEE, 1–5.

LJUBOVIC, V., AND SUPIC, H. 2013. Comparative study of color histograms as global feature for image retrieval. In *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2013 36th International Convention on*, IEEE, 1059–1063.

MANJUNATH, B. S., OHM, J.-R., VASUDEVAN, V. V., AND YAMADA, A. 2001. Color and texture descriptors. *Circuits and Systems for Video Technology, IEEE Transactions on 11*, 6, 703–715.

MÜLLER, H., MARCHAND-MAILLET, S., AND PUN, T. 2002. The truth about Corel-evaluation in image retrieval. In *Image and Video Retrieval. Lecture Notes in Computer Science*, vol. 2383. Springer, 38–49.

PAVLIDIS, T. 2008. Limitations of content-based image retrieval. In *Invited Plenary Talk at the 19th Internat. Conf. on Pattern Recognition, Tampa, Florida, December*, 8–11.

RUBNER, Y., TOMASI, C., AND GUIBAS, L. J. 2000. The Earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision 40*, 2, 99 – 121.

SCHAEFER, G., AND STICH, M. 2003. UCID: An uncompressed color image database. In *Proc. SPIE 5307. Storage and Retrieval Methods and Applications for Multimedia 2004*, International Society for Optics and Photonics, 472–480.

STRICKER, M. A., AND ORENGO, M. 1995. Similarity of color images. In *Proc. SPIE 2420. Storage and Retrieval for Image and Video Databases III*, International Society for Optics and Photonics, 381–392.

2013. Annotated groundtruth database. Department of Computer Science and Engineering, University of Washington. Retrieved on February 13th, 2013 from:. `http://www.cs.washington.edu/research/imagedatabase/`, 2.

WANG, J. Z., LI, J., AND WIEDERHOLD, G. 2001. SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 23*, 9, 947–963.

A Appendix: Color Model Conversion Formulas

In computer memory pixels are usually represented as 24-bit RGB values – 3 integers in range [0,255] corresponding to R, G and B channels. These values are converted into different color models using formulas given below. In addition, all values are rescaled to range [0,1] and represented as single-precision floating point values.

For shorter writing, we will introduce the following symbols:

$$\begin{aligned} M &= \max(R,G,B) \\ m &= \min(R,G,B) \\ C &= M - m \end{aligned} \tag{1}$$

Given above, hue (H) channel has the same meaning in HSV, HSL, HSI, HMMD and IHLS models, and that formula equals the following:

$$H = \begin{cases} \text{undefined}, & C = 0 \\ \frac{G-B}{C} \bmod 6, & M = R \\ \frac{B-R}{C} + 2, & M = G \\ \frac{R-G}{C} + 4, & M = B \end{cases}$$

Value (V) in HSV is simply equal M. Saturation (S) in HSV is defined as follows:

$$S_{HSV} = \begin{cases} 0, & C = 0 \\ \frac{C}{M}, & C \neq 0 \end{cases}$$

Lightness (L) in HSL model is an average of maximum and minimum given in (1):

$$L_{HSL} = \frac{1}{2}(M + m),$$

while intensity (I) in HSI is average of all three RGB channels:

$$I_{HSI} = \frac{1}{3}(R + G + B)$$

Saturation (S) in HSI and HSL is the same and given with the following equation:

$$S_{HSL} = \begin{cases} 0, & C = 0 \\ \frac{C}{1 - |2L_{HSL} - 1|}, & C \neq 0 \end{cases}$$

IHLS model uses the same value for lightness as HSL, while for saturation the value of C is used as given in (1).

In HMMD color model the color is described with 4 components which are Hue (calculated as above), maximum, minimum and difference, which are equal to M, m and C given in (1). Since M and C are already tested as $\mathrm{V_{HSV}}$ and $\mathrm{S_{IHLS}}$ respectively, in our experiments we included a value labeled $\mathrm{m_{HMMD}}$ which represents "minimum".

Finally, for calculation of $\mathrm{Y_{YUV}}$ we used the formula given in ITU CCITT T.81 standard:

$$Y = 0.299R + 0.587G + 0.114B$$