Image Retrieval Based On Semantics of Intra-Region Color Properties

Atul Sajjanhar[1†], Guojun Lu[2], Dengsheng Zhang[2], Wanlei Zhou[1], Yi-Ping Phoebe Chen[1]

[1]School of Engineering and Information Technology
Deakin University
221 Burwood Highway
Burwood, VIC 3125
Australia
{atuls, wanlei, phoebe}@deakin.edu.au

[2]Gippsland School of Information Technology
Monash University
Northways Road
Churchill, VIC 3842
Australia
{guojun.lu, dengsheng.zhang}@infotech.monash.edu.au

Abstract

Traditional image retrieval systems are content based image retrieval systems which rely on low-level features for indexing and retrieval of images. CBIR systems fail to meet user expectations because of the gap between the low level features used by such systems and the high level perception of images by humans. Semantics based methods have been used to describe images according to their high level features. In this paper, we performed experiments to identify the failure of existing semantics-based methods to retrieve images in a particular semantic category. We have proposed a new semantic category to describe the intra-region color feature. The proposed semantic category complements the existing high level descriptions. Experimental results confirm the effectiveness of the proposed method.

1. Introduction

Many methods have been proposed for content based image retrieval (CBIR). CBIR systems rely on low-level features which can be automatically extracted and used to index images. State of the art CBIR systems fail to meet user expectations because such systems are unable to index images according to the high level features as perceived by the user. Semantic gap is the gap between human perception and low-level features. The gap is bridged using high-level semantics to index and retrieve images.

Semantics have been used to describe low-level image features such as color, texture and shapes of objects. The focus of this paper is in the area of semantics for color based features of images. Semantics for color have been proposed for regions and for whole-image. Othman et al. [1] proposed color names for the whole scene based on CIE*LAB color histograms. Liu et al. have proposed semantic color names for region based image retrieval (RBIR) [3]; their system uses the HSV color space. The system used by Liu et al. is in the category of RBIR which is based on intra-region properties i.e. inter-region properties have not been utilized for image semantics. Corridoni et al. [2] have used semantics to describe intra-region and inter-region properties. They have used hue, saturation, luminance, warmth, size and position as intra-region properties. Contrasts of hue, saturation, luminance and warmth have been used as inter-region properties. We focus on semantics for RBIR based on color because RBIR is closer to human perception than features for the whole-image. We focus on the intra-region color properties only. A key component of systems utilizing semantics for colors is

† PhD candidate, GSIT, Monash University, Australia

Authorized licensed use limited to: Monash University. Downloaded on January 30, 2009 at 22:08 from IEEE Xplore. Restrictions apply.

to quantize colors into color names from a predefined set of color names.

In Section 2, we present an overview of RGB and HSV color spaces. In Section 3, we discuss the high-level semantic color names proposed by Liu et al. and propose a modification to the method. In Section 4, results of experiments are shown. It is shown that HSV which is perceptual color space performs better than RGB for RBIR. Effectiveness of the proposed color-names is also tested. In Section 5, we present the conclusions and directions for future work.

2. 3D Color Spaces

Color is widely used in image retrieval. Region Based Image Retrieval (RBIR) is close to human perception. The whole-image characteristics might be used as an additional descriptor in the semantics of an image; however, it is not suitable to be used on its own. Images are segmented into regions of uniform color (and texture) and the features of the regions are used to index images for RBIR.

Color can be represented in a number of color spaces [4]. High-level semantic color names have been used to represent colors [5]. We discuss two commonly used color models, namely, RGB model and HSV model. RGB color space is a uniform color space represented as a cube. Colors are constituted from the three primary colors, namely, red, green and blue. RGB model is suitable for hardware implementations. The drawback of the RGB model is that it is does not perceptive, for instance, humans do not perceive colors as being constituted by red, green and blue components. HSV color space [6] is a color space which defines perceptual colors based on hue, saturation and value (luminance). Hue is a color attribute that describes a pure color i.e. the dominant wavelength. Saturation is the amount of hue present in the color. Value (or luminance) gives the amount of light in the color i.e. the brightness. Average RGB color may be computed for each region; HSV is computed for each region based on the average RGB.

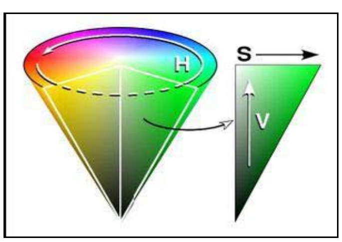


Figure 1. HSV - 3D color space

The 3D color space for HSV is shown in Figure 1; it is a non-uniform color space. Hue is the different colors represented along a circle. Saturation is the depth/richness of the color. V represents the luminance or the brightness. V=0 will give a "black" color (irrespective of the values of H and S). HSV color space is in fact a cylinder and not a cone as shown in Figure 1. The conical representation compensates the perception that when V is low (i.e. for dark colors) the color change is not noticeable for changes in S; this distortion results in a non-uniform color space. The effectiveness ofobtain in color names in RGB color space and HSV color space is evaluated in Section 4.

3. Color Names for Regions

Color names are obtained by assigning natural language names to a quantized color space [7][8][9]. The quantization of the color space is subjective as individuals will distinguish colors based on their experiences, vocabulary and cultural background. The general approach is to quantize the color map into a few distinct colors and then use additional properties of the color space to qualify the color name.

In Section 4, we have compared HSV and RGB color spaces. HSV color space outperforms RGB color space for retrieval of perceptually similar images based on color. Our investigations are based on the color names proposed by Liu et al. [3] since they have used the HSV color space. We propose a method to complement the color names used by Liu et al. [3] with additional semantic information. Liu et al. have proposed semantics based high-level color names by modifying the quantization levels used by Conway [7]. The average color is obtained for each region in RGB color space and subsequently converted into HSV space. A basic color name is obtained for quantized values of H. Adjectives are prefixed to the color name based on the quantized value of S and V. Liu et al. have quantized H into 8 basic colors. The quantization for H is performed non-uniformly; however, it is not clear how the quantization levels have been selected. Saturation and Luminance are quantized into 3 bins. Adjectives are assigned to each quantization level of Saturation and Luminance; the adjectives qualify the basic color names which are assigned to quantized values of Hue. There are a few special cases: when S=0 and V=1, the color obtained is 'grey'; when S=0 and V>80, the color obtained is 'white'; when V=0, the color is always 'black'. They have defined a total of 35 color names. It is found that the color names described above are unable to discriminate images of a certain semantic category. In order to identify the drawback of the method, consider the sample region shown in Figure 2.

Image regions are obtained from an image. We obtain regions using JSEG which is a segmentation

Authorized licensed use limited to: Monash University. Downloaded on January 30, 2009 at 22:08 from IEEE Xplore. Restrictions apply.

Figure 2. Query

tool based on homogeneity of color-texture patterns [10]. Using the image region in Figure 2 as a query, we perform an experiment to retrieve images from the dataset (as described in Section 4) which are perceptually similar. The method used is the same as in [3]; it is based on color names as described above. The results retrieved for the sample region are shown in Figure 3. The method is unable to handle the query because the retrieved images are not perceptually similar to the query image. There are other images within the database which are more similar to the query image. We propose a modification to the color names described above. The motivation to modify the color names is to be able to obtain better results for the query image in Figure 2 and for other queries images in the same semantic category. We attempt to answer the question: "Is it possible to complement color names with adjectives so that better results may be obtained?". We need to provide a semantic description to complement the existing color names based on low-level features.

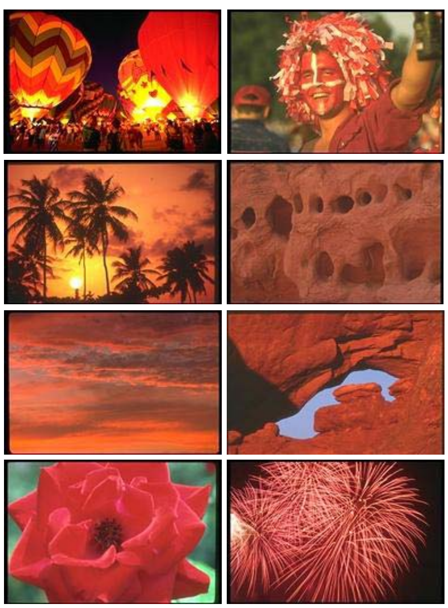
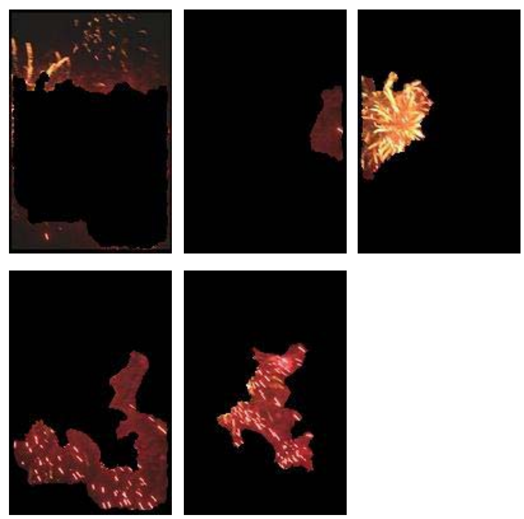


Figure 3. Images Retrieved

The sample region in Figure 2, is obtained as a region of homogeneous color-texture pattern from the image in Figure 4a. In Figure 4b, we show all the regions obtained from the image using JSEG [10]. In Ref [3] the regions are assigned a semantic color name which is obtained by quantizing the average color of each region. The color space used in Ref [3] is HSV, hence, color names inherently contain information about hue, saturation and luminance. It is evident that the information contained in the color names based on hue, saturation and luminance is not sufficient to retrieve relevant images for the query in Figure 2.

(a)



(b)

Figure 4: (a) Sample Image (b) Segmented Regions

Table 1 : Intra-region Properties of the Segmented Regions

	Region 1	Region 2	Region 3	Region 4	Region 5
Area	43.42	12.08	14.27	4.31	25.92
Mean R	51.6308	198.7015	164.5205	85.6811	108.4416
Mean G	32.1772	127.4865	61.8899	34.2274	40.5576
Mean B	27.6915	78.1587	57.6087	32.4575	39.5019
STD R	7.8118	***11.5813***	8.5025	5.7896	7.7762
STD G	5.5029	***10.5899***	7.7906	3.3430	5.2649
STD B	5.4276	***9.7343***	7.5521	3.7615	5.2725

Authorized licensed use limited to: Monash University. Downloaded on January 30, 2009 at 22:08 from IEEE Xplore. Restrictions apply.

The image in Figure 2, has a noticeably high standard deviation compared with other regions of the parent image (Figure 4a). We compare the intra-region properties of the regions, including standard deviations of the color components. The properties for each region in Figure 4b are shown in Table 1. We examine the properties of the regions to identify a suitable semantic category.

From the results in Table 1, it is clear that the second region in Figure 2 has distinctly higher standard deviations for R, G and B. We propose to qualify the color names of regions with a suitable adjective when the standard deviation is high.

4. Results

We performed two sets of experiments. A subset of the Coral image dataset consisting of 5100 images is used. The database comprises of categories of images such as cars, roses, buildings, apes, etc. The images are segmented into regions based on their color and texture using JSEG [10]. JSEG does segmentation based on homogeneity of color-texture pattern. Regions with area less than 3% are ignored for indexing purposes.

In the first set of experiments, content based image retrieval is performed using the color feature. Evaluation is done for the two color spaces, namely, RGB and HSV. Target images for a query image are ranked based on EMD [11]. EMD is used to compute the distance between the query image and the target image. It determines the effort required to transform one distribution into another based on a traditional transportation problem from linear optimization. Consider a query image q which has n regions and a target image d which has m regions. The feature vector in RGB color space for an image q is as shown below.

$$F_q = \{R_q^i, G_q^i, B_q^i, w_q^i\}$$

where, i=1 to n, n being the number of regions in image q

R_q^i, G_q^i, B_q^i are the red, green and blue components of the average color for region i of image q

w_q^i is the weight assigned to region i of image q

Each feature is scaled to the range [0, 1]. The Euclidean distance between region x of the query image (q) and region y of the target image (d) is computed as shown below.

$$d_{qx,dy} = \sqrt{\left(R_q^x - R_d^y\right)^2 + \left(G_q^x - G_d^y\right)^2 + \left(B_q^x - B_d^y\right)^2}$$

A cost function is built by computing the Euclidean distance between each pair of regions. If there are n regions in the query image and m region in the target image, the cost function will have $n*m$ terms. Weights assigned to each region (shown in the feature vector) are also supplied to the EMD function. Weights are computed as the percentage area of the region. Cost function and weights are supplied to the EMD function and the distance between two images is computed. Distance between images is computed similarly in HSV color space. The average value of Recall-Precision for 20 queries is shown in Figure 6. The performance of HSV color space is significantly better than RGB color space. Hence, HSV is chosen as the color space for quantizing colors into high-level color names.

(a)

Authorized licensed use limited to: Monash University. Downloaded on January 30, 2009 at 22:08 from IEEE Xplore. Restrictions apply.

(b)

Figure 5: Retrieval Results in RGB and HSV Color Space

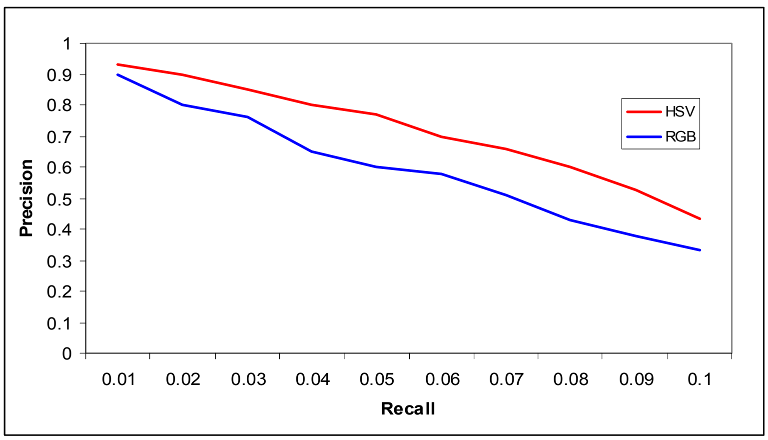


Figure 6. Recall-Precision plot for RGB vs. HSV color-based image retrieval

In the second experiment, we show the effect of including high-level semantics description of the region color. A color name is used to describe the hue of the region. Adjectives based on saturation and luminance and standard deviation are used to qualify the color name. Color names are obtained by quantizing H into 8 values. The quantization levels are uniformly created. S and V are quantized into bins, similar to that used in [1]. The results of retrieval for the image region in Figure 7a are shown in Figure 7b. Comparing the results with the results in Figure 3, it is apparent that the new color names are effective in retrieving semantically similar images. The new color names obtained by complementing existing color names, improve the performance of some images.

(a)

Authorized licensed use limited to: Monash University. Downloaded on January 30, 2009 at 22:08 from IEEE Xplore. Restrictions apply.

(b)

Figure 7: (a) Region used as Query (b) Retrieved Images using proposed Color Names

5. Conclusion

In this paper, we evaluated RGB and HSV color spaces for feature extraction from images. It is found that HSV color space is suitable to describe colors according to human perception. We have used high-level semantic color names to represent image regions based on the mean H, S and V of the region. The proposed semantic color names are based on the standard deviation of constituent color components of the region. It is shown that the proposed semantic names are effective for some types of images which were not handled well previously. In this paper, standard deviation of the color has been quantized into two values. In the future, we will study the effect of more levels of quantization and its effect on the performance. For instance, very low values of standard deviations might indicate regions which make up the image background and hence, these regions may be ignored during the retrieval process based on the assumption that the foreground is more relevant.

References

[1] Aniza Othman, Kirk Martinez, "Colour Appearance Descriptors for Image Browsing and Retrieval", *SPIE Electronic Imaging: Multimedia Content Access: Algorithms and Systems*, 27 January, San Jose, USA.

[2] J. M. Corridoni, A. Del Bimbo, P. Pala, "Image retreival by color semantics", *Multimedia Systems,* 1999, Volume 7: pp. 175-183.

[3] Y. Liu, D. S. Zhang, G. Lu, W. Y. Ma, "Region-Based Image Retrieval with High-Level Semantic Color Names", *International Multimedia Modeling Conference (MMM05)*, Melbourne, Australia, January 2005

[4] K. N. Plataniotis, A. N. Venetsanopoulos, *Color Image Processing and Applications*, Springer, Berlin, 2000.

[5] Ying Liu et.al. "A survey of content-based image retrieval with high-level semantics", *Pattern Recognition*, 2007

[6] E. B. Goldstein, *Sensation and Perception*, 5th ed. 2003: Thomson Wadsworth.

[7] D. M. Conway. "An Experimental Comparison of Three Natural Language Color Naming Models," *East-West Inter. Conf. on Human-Computer Interactions*. 1992. St. Petersburg, Russia: pp. 328-339

[8] P. L. Stanchev, D. Green Jr. and B. Dimitrov, "High Level Color Similarity Retrieval", *International Journal of Information Theories & Applications*, 2003. 10(3): pp. 363-369

[9] J. Ren, Y. Shen and L. Guo, "A Novel Image Retrieval Based on Representative Colors", *Image and Vision Computing*. Nov. 2003. N.Z.: pp. 102-107.

[10] Y. Deng, B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video", *IEEE Trans. Pattern Anal. Mach. Learn. (PAMI)* 23 (8), 2001. pp. 800–810.

[11] Y. Rubner, C. Tomasi, L. Guibas, "A metric for distributions with applications to image databases", *IEEE International Conference on Computer Vision (ICCV'98)*, January 1998, pp. 59–67.

Authorized licensed use limited to: Monash University. Downloaded on January 30, 2009 at 22:08 from IEEE Xplore. Restrictions apply.