



Distinguishing Cartoons Images from Real-Life Images

Mustafa M. Abd Zaid, Loay E. Gerorge, Ghadah Al-Khafaji

College of Science, University of Baghdad,
Iraq

Abstract— Distinguishing cartoon images from real-life images is important for many applications, such as web browser, or when the user search for images in multimedia documents, he/she often has in mind specific image types that is search for. In this paper, automatically method is introduced for classifying images as real-life or cartoons; it is based on extracting features from (RGB) and (HSV) colors space, the first (RGB) feature depend on the difference of current pixel and its neighbors, the cartoon images have more zero difference than photographs, the second (RGB) feature based on the extracting brightness from (RGB) color space for the image, since the cartoon images more brightness than photos, and the last features extracted from (HSV) color space the average saturation and brightness of images are extracted based on the cartoon images have large saturation and brightness than photos. The test results are indicated high performance for the suggested method.

Keywords: Real-Life Images, Cartoon Images, HSV and RGB Features, K-Means Applications, Image Classification.

I. INTRODUCTION

In general, images can classify as photographs, portraits, maps, buttons, charts, or several other types, make it considerably easier for people to specify and get back the kind of images they are interested in[1]. Today, the automated system classification become an urgent requirement, where for web images included that of (V. Athitsos, M.J. Swain, and C. Frankel [1]) that utilized the image contents to classify between photographs or graphics images, also adopted (A. Vailaya, A.K. Jain, H.-J. Zhang [2]) the classification between city images vs. landscapes based on low-level features, and (F. Cutzu, R.Hammoud and A.Leykin [3]) differentiating photographs of real scenes from photographs of paintings according color, edge, and texture properties.

In this paper, an efficient automated classification technique is introduced that able to distinguish the cartoon images from the real-world images according to some assumptions that related to content-based. The reset of this paper is organized as follows; section (II) is devoted to describe the suggested distinguish system scheme in details, section (III) mainly to describe using image sets which are used for testing the performance of proposed system, while in section (IV) some of the tests' results are given.

II. THE PROPOSED SYSTEM

The proposed system structure shown in figure (1), the following described steps illustrated the implementation clearly. Basically, the main concerns were taken into account are:

- Cartoon images contain few, simple and strong colors.
- Saturation in cartoon images is higher than photographic.
- Cartoon images have strong black edges [3].

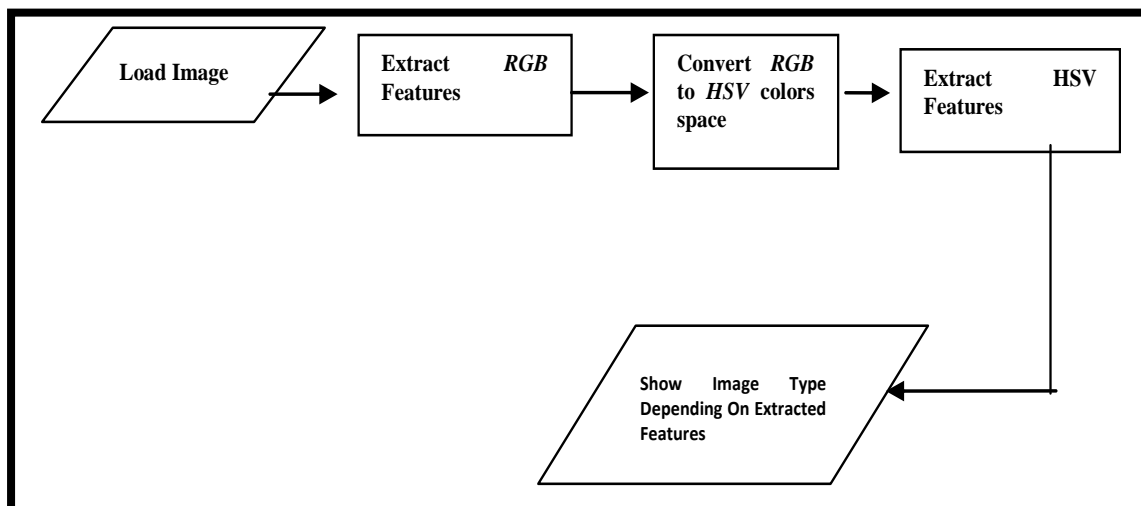


Fig. (1): The layout of proposed system

Step-1: Load the input image I of size $W \times H$.

Step-2: Extract first (RGB) feature (difference feature), where this (RGB) feature depends on the difference of current pixel and the neighbors pixels, because of the cartoon images have little change in colors the zero difference for cartoon would riddance more than zero for photographs, according to equations (1-7), and then followed by finding the thresholds.

$$D1 = (Img(x, y)_{RGB} - Img(x + 1, y)_{RGB}) \dots \dots \dots (1)$$

$$D2 = (Img(x, y)_{RGB} - Img(x, y + 1)_{RGB}) \dots \dots \dots (2)$$

$$D3 = (Img(x, y)_{RGB} - Img(x + 1, y + 1)_{RGB}) \dots \dots \dots (3)$$

$$Img(x, y)^D = \begin{cases} D1 & \text{if } D1 > D2, D1 > D3 \\ D2 & \text{if } D2 > D1, D2 > D3 \\ D3 & \text{if } D3 > D1, D3 > D2 \end{cases} \dots \dots \dots (4)$$

$$F1_{RGB} = \begin{cases} Count_{RGB} + 1, \text{next } x, y & \text{if } Img(x, y)^D_{RGB} = 0 \\ \text{next } x, y & \text{if } Img(x, y)^D_{RGB} <> 0 \end{cases} \dots \dots \dots (5)$$

$$F1_{RGB} = (Count_{RGB}) / (W \times H) \dots \dots \dots (6)$$

$$F1 = ((F1red + F1green + F1blue)/3) \dots \dots \dots (7)$$

Where Img is original image, D is selected different, $F1_{RGB}$ and $F1$ are the extracted features.

The thresholds for this feature is extracted based on the size of loaded image, because this feature vary according by the size of image, therefore stander dimensions for loading image is chose for thresholds and those dimensions are change automatically with dimensions of loaded image, as equations below shown.

$$Thrshold1=(r \times 0.8) \dots \dots \dots (8)$$

$$Thrshold2=(r \times 0.15) \dots \dots \dots (9)$$

$$Thrshlod3=(r \times 0.62) \dots \dots \dots (10)$$

Where r computes as in equation (10), if r less than 1.5 or more than 0.5, then using the equations (11, 12) respectively, to reduce the standard size, and reuse equation (10) until the condition is satisfied:

$$r = \left(\frac{Wratio + Hratio}{2} \right) \dots \dots \dots (11)$$

$$sz1 = sz1 + 10, sz2 = sz2 + 10, \text{if } r < 0.5 \dots \dots (12)$$

$$sz1 = sz1 - 10, sz2 = sz2 - 10, \text{if } r > 1.5 \dots \dots (13)$$

The $sz1$ and $sz2$ are the standard weight and height, which are $sz1=304$ and $sz2=304$ in the first step and then are modified if necessary using the $Wratio$ and $Hratio$ such as:

$$Wratio = \frac{W}{sz1} \dots \dots \dots (14)$$

$$Hratio = \frac{H}{sz1} \dots \dots \dots (15)$$

Figure (2) shows the flowchart for this feature in 100 cartoon images and 100 photos, where clearly shown the difference feature is higher for cartoon images than real-life images.

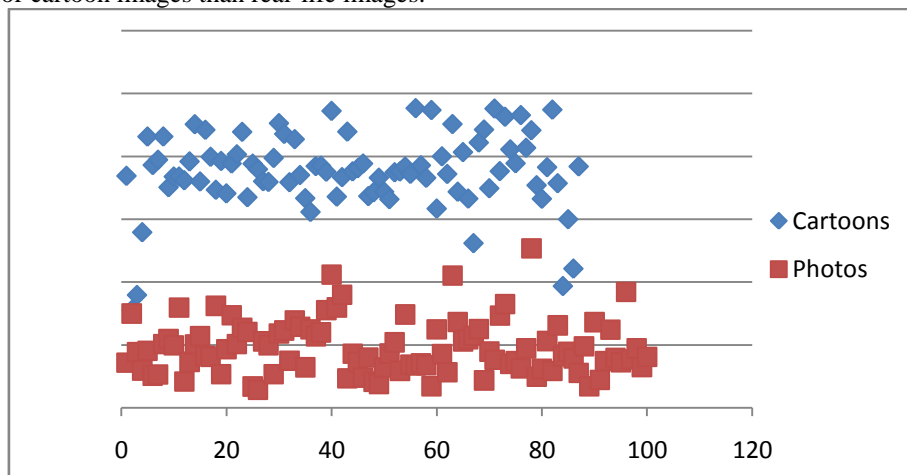


Fig. (2): Difference feature graph for 100 cartoon images and 100 photos

Step-3: Extract the second (RGB) features depends on the brightness of the image since the cartoon images more brightness than photographs, use equations (16-18). Figure (3) illustrates this feature is high for cartoons than of real-life images clearly.

$$Img(x, y)^B_{RGB} = (Img(x, y)_{RGB} + Img(x + 1, y)_{RGB} - Img(x, y + 1)_{RGB} + Img(x + 1, y + 1)_{RGB}) \dots (16)$$

$$F2_{RGB} = \left(\sum_{x=0}^h \sum_{y=0}^w \left((Img(x, y)^B_{RGB}) \right) \right) / (W * H) \dots \dots \dots (17)$$

$$F2 = ((F2red + F2green + F2blue)/3) \dots \dots (18)$$

Where $Img(x,y)^B_{RGB}$ represents brightness of the image, $Img(x,y)_{RGB}$ is original loaded image, $F2_{RGB}$ and $F2$ are second extracted features.

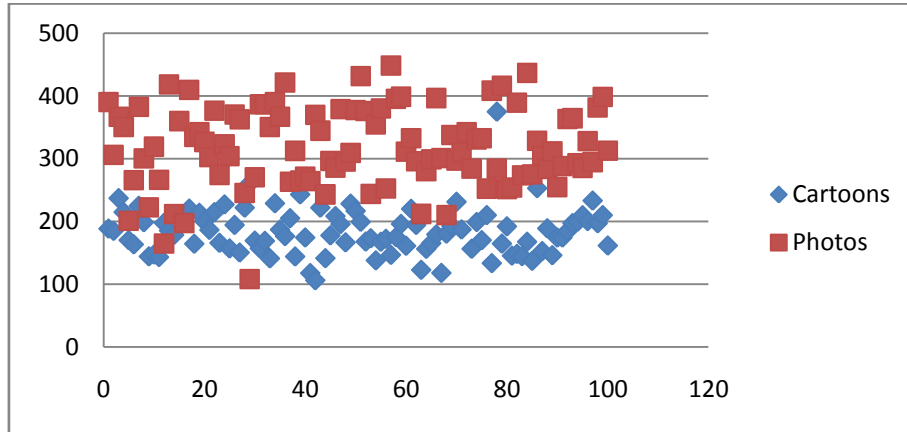


Fig. (3): Graph of brightness feature for 100 cartoon images and 100 real-life images

Step-4: Extract HSV features, where Hue refers to the color of red, blue and yellow and has the range from 0 to 360. Saturation means purity of the color and takes the value from 0 to 100%. Value refers to the brightness of the color and provides the achromatic idea of the color [4-6]. First convert the image in RGB to HSV color space, and then use equations (19-26) below.

$$R^N = R / 255$$

$$G^N = G / 255 \dots\dots\dots (18)$$

$$B^N = B / 255$$

$$Cmax = MAX(R^N, G^N, B^N) \dots\dots\dots (19)$$

$$Cmin = MIN(R, G^N, B^N) \dots\dots\dots (20)$$

$$Chorma = (Cmax - Cmin) \dots\dots\dots (21)$$

$$H = \begin{cases} 60 * (\frac{G'-B'}{Chorma} \text{ mod } 6) & , Cmax = R^N \\ 60 * (\frac{R'-B'}{Chorma} + 2) & , Cmax = G^N \\ 60 * (\frac{R'-B'}{Chorma} + 4) & , Cmax = B^N \end{cases} \dots\dots\dots (22)$$

$$S = \begin{cases} 0 & , Chorma = 0 \\ \frac{Chorma}{Cmax} & , Chorma <> 0 \end{cases} \dots\dots\dots (23)$$

$$V = CMAX \dots\dots\dots (24)$$

$$AvgS = (\sum_{i=0}^h \sum_{j=0}^w S_{i,j}) / (Width * Height) \dots\dots\dots (25)$$

$$AvgV = (\sum_{i=0}^h \sum_{j=0}^w V_{i,j}) / (Width * Height) \dots\dots\dots (26)$$

Where R^N , G^N and B^N are the new values of current pixel after divided the current pixel value by 255 for each band. Cmax and Cmin are the max and the min values among these new values. Chorma represents the difference between the max and the min. H describes the huge of image. S represents the saturation value. V refers to the variance. AvgS and AvgV are the average saturation and variance, in these two averages, values considered as distinguish features between cartoons and real-life images. See figure (4) to illustrates the high value for cartoon compared to real-life images for 100 of cartoon images and 100 of photos, which show

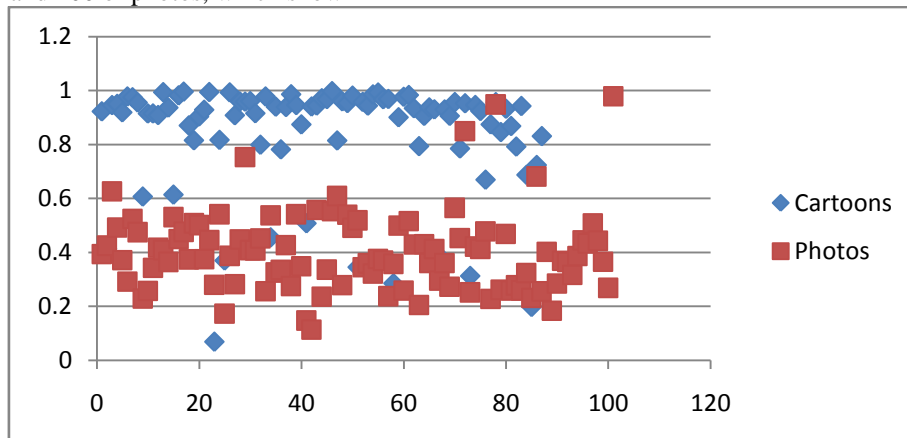


Fig. (4): Graph of value feature for 100 cartoon images and 100 real-life images, cartoons have more highly valued pixels

Step5: Show image type; determine the type of image depending on extracted thresholds for each feature.

1. Cartoon if $F1 > \text{Threshold1}$ or $F2 > 350$ or ($F2_{\text{red}}$, $F2_{\text{green}}$ and $F2_{\text{blue}} < 80$)
2. Photograph if $F1 < \text{Threshold2}$ or $F2_{\text{red}}$, $F2_{\text{green}}$ and $F2_{\text{blue}} < 160$
3. Else if

$\text{AvgV} > 0.85$ Then $\text{count} = \text{count} + 1$

If $(\text{AvgS} + \text{AvgV}) / 2 > 0.6$ Then $\text{count} = \text{count} + 1$

If $(F1_{\text{red}} + F1_{\text{green}} + F1_{\text{blue}}) / 3 > \text{threshold3}$ Then $\text{count} = \text{count} + 1$

If $\text{count} > 0$ then cartoon else Photograph

III. THE IMAGE SET

For testing the proposed system performance, it is applied on 900 real-life images of different types (Corel images database), the class “photograph” included 9 different classes of Corel images database (see figure (5)), and more than 850 cartoon images of different objects, the classes “cartoon” included Line (pencil or ink) drawings and computer-generated images, 10 different classes of cartoon images contain different colors and characters of different size used in this paper (see figure (6)), which are calculated from web and cartoon-11k [7]. The testing images have different sizes.



Fig. (5): One image example from each real-life images class (Corel database).

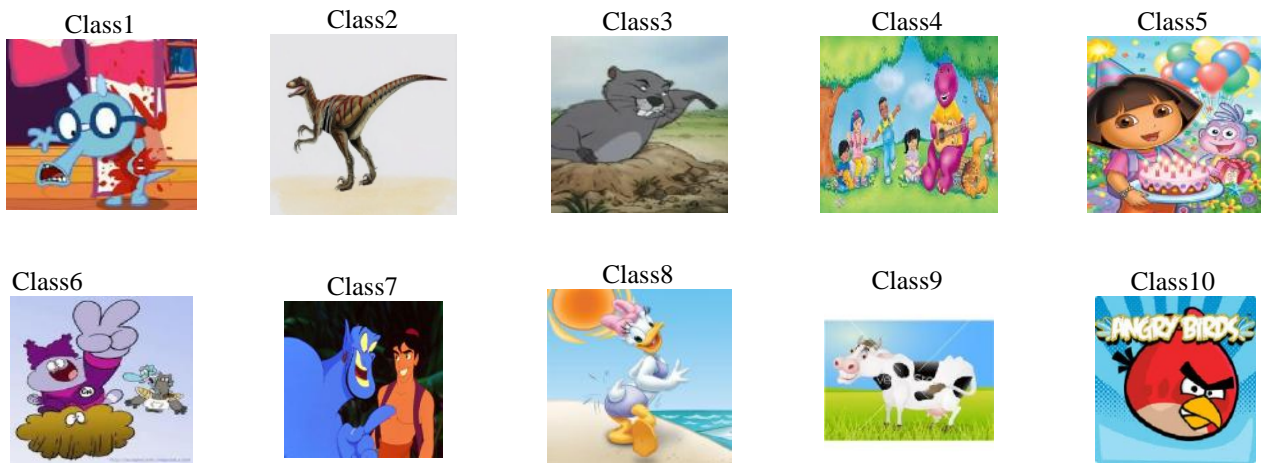


Fig. (6): One image example from each cartoon images class.

IV. EXPERIMENTS AND RESULT

The computed features of each image compared to extracted thresholds, and then consider the image to be a cartoon or real-life image. The results show that the system has well results in more natural images and cartoon images, the system is showed false results in the brightness images and simple color for real-life images (images have no long change in color), (see figure (7)). Tables (1&2) show the result of the proposed system on real-life images and on cartoon images

respectively; these cartoons sometimes content many details or have the saturation and brightness such as photographs, (see figure (8)). The change in the extracted thresholds will lead to an increase in the correct ratio of a particular type of images and a decrease in the other kind.

Table 1: The effectiveness of the proposed method in real-life images database

Class #	True ratio	False ratio
Class1	99%	1%
Class2	89%	11%
Class3	98%	2%
Class4	97%	3%
Class5	100%	0%
Class6	97%	3%
Class7	93%	7%
Class8	93%	7%
Class9	87%	13%

Table 2: The effectiveness of the proposed method in cartoon images database

Class#	True ratio	False ratio
Class 1	96.5%	3.5%
Class2	98%	2%
Class3	96.5%	3.5%
Class4	99%	1%
Class5	96%	4%
Class6	92.5%	7.5%
Class7	82%	18%
Class8	93.5%	6.5%
Class9	100%	0%
Class10	100%	0%



Fig. (7): Misclassified real life image classified as cartoon image

REFERENCE

- [1] V. Athitsos, M.J. Swain, and C. Frankel; "Distinguishing Photographs and Graphics on the World Wide Web", IEEE Workshop, On Content-Based Access of Image and Video Libraries(CBAIVL _97), Puerto Rico, Pp. 10-17, 1997.
- [2] A. Vailaya, A.K. Jain, H.-J. Zhang; "On Image Classification: City Vs. Landscapes, Int. J. Pattern Recogn". Pp. 1921-1936. , 1998.
- [3] F. Cutzu, R.Hammoud and A.Leykin; "Distinguishing Paintings from Photographs", Elsevier, Pp 249-273, 2005.
- [4] J.M Chaves-Gonzalez, M.A Vega-Rodriguez, J. A.Gomez-Pulido, and J. M Sanchez-perez; "Detecting Skin In Face Recognition Systems: A Color Spaces Study", Digital signal processing, Elsevier, Pp. 806-823, 2010.
- [5] S. Chitra, and G. Balakrishnan; "Comparative Study for Two Color Spaces HSCbCr and YCbCr in Skin Color Detection", Applied Mathematical Sciences, Vol. 6, No. 85, Pp. 4229 - 4238, 2012.
- [6] I. Aldasouqi, and M. Hassan; "Smart Human Face Detection System", International journal of computers, Vol 5, Pp.210-217, 2011.
- [7] T. Ahmed, M. Massudi, H. Husniza and E. Loay; "A Weighted Dominant Color Descriptor for Content-Based Image Retrieval", J. Vis. Commun. Image R., (Elsevier), Pp. 345-360, 2013.