

IDENTIFYING COMPUTER GRAPHICS USING HSV COLOR MODEL AND STATISTICAL MOMENTS OF CHARACTERISTIC FUNCTIONS

Wen Chen¹, Yun Q. Shi¹, Guorong Xuan²

¹New Jersey Institute of Technology, Newark, NJ, USA

²Dept. of Computer Science, Tongji University, Shanghai, China
{wc47@njit.edu, shi@njit.edu}

ABSTRACT

Computer graphics generated by advanced rendering software come to appear so photorealistic that it has become difficult for people to visually differentiate them from photographic images. Consequently, modern computer graphics may be used as a convincing form of image forgery. Therefore, identifying computer graphics has become an important issue in image forgery detection. In this paper, a novel approach to distinguishing computer graphics from photographic images is introduced. The statistical moments of characteristic function of the image and wavelet subbands are used as the distinguishing features. In addition, we investigate the influence of different image color representations on the feature effectiveness. Specifically, the efficiency of using RGB and HSV color models is investigated. The experiments have shown that the features extracted from HSV color space, which decouples brightness from chromatic components, have demonstrated better performance than that from RGB color model.

1. INTRODUCTION

This paper aims at the development of a novel method to automatically separate computer graphics from photographic images. In this paper, computer graphics will be used to refer to images which are created by a variety of rendering software, and photographic images are the output of imaging acquisition devices such as digital camera.

The identification of computer graphics is challenging. Computer graphics can now achieve such a photorealistic level that people cannot identify it with high confidence. From a practical point of view, an automatic classification system is useful to deal with this issue. On the one hand, the breakthrough made in this research area will defeat the image forgery in the following areas: criminal investigation, journalism, intelligence, etc. On the other hand, it will help to improve the rendering technology to generate more photorealistic computer graphics used in movie industry.

The detection of computer graphics can be considered as a two-class pattern recognition problem. The goal is to

determine appropriate features and develop a classification system to differentiate computer graphics from photographic images. In the prior works, several classification systems based on different types of image features have been developed. In [1] the classification problem was addressed by modeling the characteristics of cartoon. The features are extracted from the color saturation, color histogram, edge histogram, compression ratio, pattern spectrum and the ratio of image pixels with brightness greater than a threshold. The total number of features is 108. The prior work [2] uses image wavelet decomposition to capture regularities inherent to photographic images. The approach collects features from the first four order statistics – mean, variance, skewness and kurtosis – of horizontal, vertical and diagonal subbands. The same four statistics are computed from the linear prediction error for the wavelet coefficients. The dimension of feature vector is 216. In [3] the geometry features are extracted by analyzing physical differences between the generative process of computer graphics and photographic images. The 192 geometry features are characterized by differential geometry, fractal geometry and local patch statistics. In addition, there are other related works on broad image classification such as city images vs. landscapes images [4], and photographs vs. paintings [5].

A color image can be represented in different color space for different applications. In the previous works, RGB (Red, Green, Blue) color space was used to extract all [2] or part [3] of the features. In [2], the wavelet-based image statistics are computed for RGB three channels. In [3], the joint spatial-color patch statistics, fractal geometry and differential geometry are formed in RGB space. In [1], some features are collected in HSV (Hue, Saturation, Value (brightness)) color space. In this paper, the proposed method constructs all features in the HSV color space. The distinguishing features have been successfully used in our previous work on steganalysis for gray-scale image [6]. The features are the moments of characteristic function of wavelet subbands. We investigate the effect of image color representation on the classification performance. Specifically, the classification performance of HSV color models is compared with that of RGB model.

To evaluate the proposed approach, we use images from the Columbia Image Dataset [7]. The contents of

photographic images contain both natural and artificial scenes. For computer graphics, only those with high level of photorealism are considered. All images are restricted to color images.

The rest of this paper is organized as follows. In Section 2, the proposed features for differentiation of computer graphics from photographic images are described. The selection of color model is discussed. Section 3 contains experimental results. Finally, the conclusions are drawn and some future works are presented in Section 4.

2. IMAGE FEATURES

2.1. Selection of color model

In color image processing, there are various color models in use today. The RGB model is mostly used in hardware-oriented application such as color monitor. In the RGB model, images are represented by three components, one for each primary color – red, green and blue. Although human eye is strongly perceptive to red, green, and blue, the RGB representation is not well suited for describing color image from human perception point of view. Moreover, a color is not simply formed by these three primary colors.

When viewing a color object, human visual system characterizes it by its brightness and chromaticity. The latter is defined by hue and saturation. Brightness is a subjective measure of luminous intensity. It embodies the achromatic notion of intensity. Hue is a color attribute and represents a dominant color. Saturation is an expression of the relative purity or the degree to which a pure color is diluted by white light. The HSV model is motivated by the human visual system. In the HSV model, the luminous component (brightness) is decoupled from color-carrying information (hue and saturation). The HSV color model is defined as follows [8]:

$$H = \begin{cases} 60\left(\frac{G-B}{\delta}\right) & \text{if MAX} = R \\ 60\left(\frac{B-R}{\delta} + 2\right) & \text{if MAX} = G \\ 60\left(\frac{R-G}{\delta} + 4\right) & \text{if MAX} = B \\ \text{not defined} & \text{if MAX} = 0 \end{cases} \quad (1)$$

$$S = \begin{cases} \frac{\delta}{MAX} & \text{if MAX} \neq 0 \\ 0 & \text{if MAX} = 0 \end{cases}$$

$$V = MAX$$

where $\delta = (MAX - MIN)$, $MAX = \max(R, G, B)$, and $MIN = \min(R, G, B)$. Note that the R, G, B values in Equation (1) are scaled to $[0, 1]$. In order to confine H within the range of $[0, 360]$,

$$H = H + 360, \text{ if } H < 0.$$

It is more natural for human visual system to describe a color image by the HSV model than by the RGB model.

Intuitively, the features extracted in the HSV color space can capture the distinct characteristics of computer graphics better. For example, computer graphics is more color smooth than photographic images in the texture area. Fewer colors are contained in computer graphics. Intensity of computer graphics reveals different characteristic of edge and shade. These differences between computer graphics and photographic images are best described by decoupling the intensity from chromatic information, say, hue and saturation.

Inspired by the way human visual system perceives the color object, we propose to construct features from the HSV color space. For the purpose of performance comparison, the features are also extracted in the RGB color space. As shown in the next section, HSV features have better classification performance than RGB features.

2.2. Moments of wavelet characteristic function

To accurately separate computer graphics from photographic images, determination of distinguishing features is a critical step. In the following, we describe the features proposed in this approach which are the statistical moments of wavelet characteristic function.

Image histogram has been widely used in image analysis. It is well-known that the histogram of a digital image or its wavelet subband is essentially the probability mass function (pmf), if the image grayscale values or the wavelet coefficient values are treated as a random variable. Any pmf p_x may be expressed as a probability density function (pdf) f_x by using the relation

$$f_x(x_0) = \sum_a p_x(a) \delta(x_0 - a) \quad (2)$$

That is, if each component of the histogram is multiplied by a correspondingly shifted unit impulse, we then have the pdf. The pmf and pdf is exchangeable in the context of digital signal processing, e.g., in that of discrete Fourier transform (DFT). Thus the pdf can be treated as the normalized version of a histogram. According to [9, pp. 145-148], the characteristic function (CF) is simply the Fourier transform of the pdf (with a reversal in the sign of the exponent).

Denote histogram and its CF by $h(f_i)$ and $H(f_k)$, respectively, we propose to use the statistical moments of the CFs of both a test image and its wavelet subbands as features, which are defined as follows.

$$M_n = \sum_{j=1}^{(N/2)} f_j^n |H(f_j)| \bigg/ \sum_{j=1}^{(N/2)} |H(f_j)| \quad (3)$$

where $H(f_i)$ is the CF component at frequency f_i , n is the moment order, and N is the total number of points in the horizontal axis of the histogram. These moments are called as features generated from the test image for short.

In addition to the moments of the test image, we also extract features in the same manner from the prediction-error

image. The prediction-error image is the difference between the test image and its predicted version. The prediction algorithm used here is given by [10]

$$\hat{x} = \begin{cases} \max(a, b) & c \leq \min(a, b) \\ \min(a, b) & c \geq \max(a, b) \\ a + b - c & \text{otherwise} \end{cases} \quad (4)$$

where a , b , c are the context of the pixel x under considerations. \hat{x} is the prediction value of x . The locations of a , b , c are illustrated as in Fig. 1.

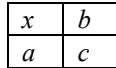


Fig.1: Prediction context

2.3. Feature extraction

To collect image features, i.e. moments of characteristic functions, a test image of either a computer graphic or a photographic image is represented by the H, S and V components. Each component image is first decomposed into three levels based on, say, Haar wavelet. At each level i , $i = 1, 2, 3$, there are four subbands (approximation, horizontal, vertical and diagonal). Totally, there are 13 subbands involved in the feature extraction if the component image itself is considered as a subband at level 0. For each subband, the first three moments are derived according to Equation (3), resulting in 39 features. For the prediction-error image of each component image, the same wavelet decomposition is employed and another 39 features are collected. Therefore 78 features are extracted from each component image and its prediction-error image. Since there are three component images and three prediction-error component images, the total number of features is 234. These features will be referred to as HSV-based features. To compare the performance, the test image is also represented by the R, G, B component images and another set of 234 features is collected following the same procedure as described above. We call them as RGB-based features.

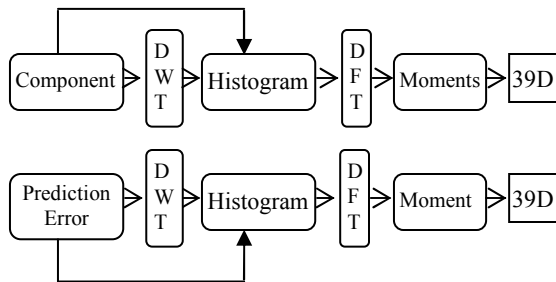


Fig.2: Feature extraction from each image component

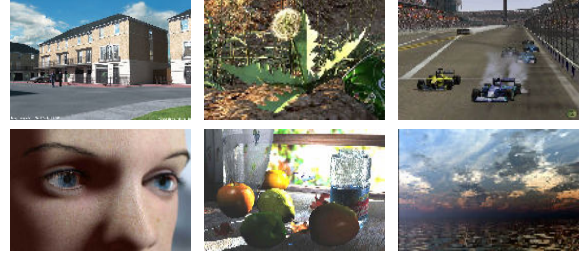


Fig.3: Some examples of computer graphics

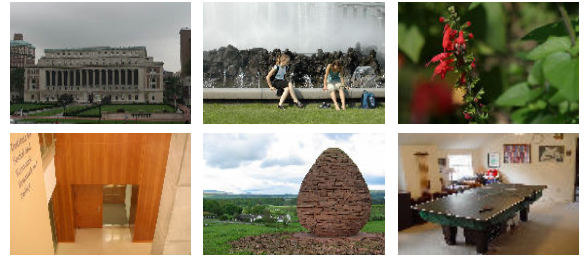


Fig.4: Some examples of photographic images

The block diagram of feature generation procedure is shown in Fig.2 where the component are H, S, V for HSV model and R, G, B for RGB model.

3. EXPERIMENT

3.1. Image database

In the experiment, we used 1900 photographic images and 800 computer graphics in the Columbia Image Dataset [7]. As introduced in [7], the computer graphics were obtained from a variety of 3D graphics websites. The highly photorealistic computer graphics contain diverse content in various artistic styles. The photographic images are obtained from three sources: eight hundred images were taken by the authors of [7], four hundred were obtained from Philip Greenspun’s personal collection [11], and the rest were downloaded from Google Image Search. Some examples are illustrated in Fig.3 and Fig.4.

3.2. Experimental results

The Support Vector Machine (SVM) classifier with RBF kernel was employed in the classification experiment. We used the “grid-search” method of LIBSVM [12] to find the optimal penalty parameter C and kernel parameter γ of RBF kernel. The classification performance is obtained by the 20 runs of experiment with the best parameters C and γ . To train the classifier, we randomly selected the training samples to include 5/6 of image set (1580 photographic images and 665 computer graphics). The testing samples from the rest 1/6 of image set contain 135 photographic images and 135 computer graphics which are not involved in the training stage. The average detection rate is shown in Table 1 where TP (true positive) represents the detection rate of computer graphics, TN (true negative) represents the

detection rate of photographic images, and the accuracy is the arithmetic average of TP and TN. The receiver operating characteristics (ROC) curves are shown in Fig.5. The experimental results in Table I show that the accuracy is 82.1% for HSV-based features, which is 5.2% higher than the accuracy of RGB-based features. The results indicate that the color model has an obvious influence on the effectiveness of image features. We also found that the proposed HSV-based features outperform the 216 wavelet features proposed in [2] which collects features in RGB space. However, the performance of [2] is better than that of our RGB-based features. This can be explained as follows. That is, the correlation between color channels is exploited to derive statistics as part of features in [2], which has enhanced the distinguishing power of features; while our RGB-based features are collected from each color channel independently. The observation of performance can also be verified by the ROC curves in Fig.5. Here we did not compare the proposed approach with [3] for the reason that the source code of algorithm in [3] is not publicly available.

Finally, Table II contains the classification performance for three different component combinations – HV (hue and brightness), SV (saturation and brightness) and HS (hue and saturation). The 156 features from the hue and brightness components can achieve accuracy of 79.6% which is better than the 234 RGB-based features and comparable to the 216 wavelet features [2].

Table I: Detection rate

	TP	TN	Accuracy
HSV	71.9%	92.3%	82.1%
RGB	64.0%	89.9%	76.9%
[2]	68.6%	92.9%	80.8%

Table II: Performance of various component combinations

	TP	TN	Accuracy
HV	67.2%	91.0%	79.6%
SV	59.3%	89.0%	74.2%
HS	62.5%	90.2%	76.4%

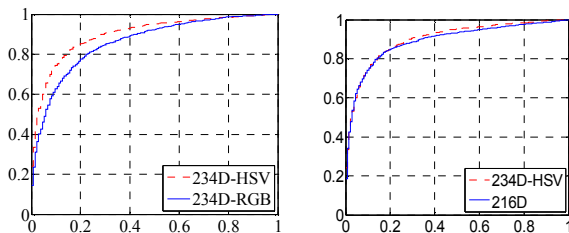


Fig. 5: ROC curves (x axial for FP, y axial for TP)

4. CONCLUSION AND FUTURE WORKS

A novel approach to the identification of computer graphics is proposed in this paper. The distinguishing features are

formed by using statistical moments of characteristic function of wavelet subbands and their prediction-errors. The effect of the color space representation on the classification performance has been investigated.

The proper selection of color image representations is important to extract effective features. There are many color models in use today in the area of color image processing. In this paper, we investigate only two commonly used color representation, i.e. HSV and RGB. It is highly expected that there exists an optimal color model which contributes to the most effective distinguishing features. One of our works in progress is to design an optimization algorithm to search for the best color model.

Furthermore, once the features have been derived, it is possible that there may be redundant or unnecessary features among the derived features. The results in Table II indicate that the classification accuracy can be achieved even only a subset of the image features is used. Hence, another future work is to design an algorithm to select a reduced feature set without significant degradation in classification performance.

REFERENCES

- [1] T. Ianeva, A. de Vries, and H. Rohrig “Detecting cartoons: a case study in automatic video-genre classification,” in *IEEE International Conference on Multimedia and Expro*, 1, pp. 449-452, 2003.
- [2] S. Lyu and H. Farid, “How realistic is photorealistic?,” *IEEE Transactions on Signal Processing*, 53, pp. 845-850, February 2005.
- [3] T.-T. Ng, S.-F. Chang, J. Hsu, L. Xie, and M.-P. Tsui, “Physics-motivated features for distinguishing photographic images and computer graphics,” in *ACM Multimedia*, Singapore, November 2005.
- [4] A. Vailaya, A. K. Jain, H.-J Zhang, “On image classification: city vs. landscapes,” *Intl. J. Pattern Recognition*, 31(1998), 1921-1936.
- [5] Florin Cutzu, Riad Hammoud, Alex Leykin, “Distinguishing paintings from photographs,” *Computer Vision and Image Understanding*, 100(2005), 249-273.
- [6] Y. Q. Shi, G. Xuan, D. Zou, J. Gao, C. Yang, Z. Zhang, P. Chai, W. Wen and C. H. Chen, “Steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image and neural networks,” *IEEE ICME*, July 2005.
- [7] Columbia University DVMM Research Lab: Columbia Photographic Images and Photorealistic Computer Graphics Dataset.
- [8] A. R. Smith, “Color gamut transform pairs,” *Comput. Graph.* 12(3) (1978) 12-19.
- [9] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd Edition, MA: Addison-Wesley Publishing Company, 1994.
- [10] M. Weinberger, G. Seroussi, and G. Sapiro, “LOCO-I: A low complexity content-based lossless image compression algorithm,” *Proc. of IEEE Data Compression Conf.* pp. 140-149, 1996.
- [11] Philip Greenspun’s collection, <http://philip.greenspun.com>
- [12] C. -W. Hsu, C. -C Chang, and C. -J Lin, “A practical guide to support vector classification,” July 2003.