高效区分实拍图像与生成图像的方法

李　宗，徐家品
（四川大学 电子信息学院，成都 610064）

摘　要：提出一种高效的图像区分方法，专注于区分实拍图像与生成图像。该方法在图像预处理过程中，利用了图像尺度及颜色数等几个高效的判断项，将那些容易区分的生成图像召回，减少了后期垃圾图像与文字图像判断项的判断量，减少了图像筛选运行时间。

关键词：图像预处理；Sobel 算子；垃圾图像；文字图像
中图分类号：TP391.41　　**文献标志码**：A　　**文章编号**：1001-3695(2012)09-3540-03
doi:10.3969/j.issn.1001-3695.2012.09.090

Effective method for distinguishing photograph and graphics

LI Zong, XU Jia-pin
(*College of Electronics & Information Engineering*, *Sichuan University*, *Chengdu* 610064, *China*)

Abstract: This paper proposed an efficient graphic distinction algorithm which focused on distinguishing graphics versus photographs. By utilizing several highly efficient entries such as scale and color, the computer-generated graphics could be recalled so that the postprocessing stage of spam images and text images had less workload which significantly reduced the runtime of image filtering.

Key words: image preprocessing; Sobel operator; spam image; text image

由于图像具有有效的信息携带和感性的信息表达能力，同时图像的运用范围更加广泛，使得当今互联网上图片数量剧增。互联网上的这些图像质量参差不齐，目前处于一种杂乱无序的状态，使得图像处理或图像存储变得非常复杂。图像处理技术的发展使得大规模的图像处理成为可能，人们普遍希望对众多图像进行一定的质量分级、内容分类等处理，从而建立规范的图像数据库，降低图像存储成本，提高图像处理的有效性，改善图像搜索的效率。

互联网上的图像大致可以分为两类：a）由相机等图像采集设备实拍得到的图像；b）由计算机生成的图像。在实际的商业应用中，存在着自动将大量图像高效区分为实拍图和生成图的必要，这也将为图像的后续处理提供便利。本文提出在图像预处理过程中，利用图像尺度及颜色数等几个高效的判断项，将那些容易区分的生成图像召回，这样就减少了之后垃圾图像与文字图像判断项的判断量，减少了图像筛选运行时间。

1　相关工作

目前已有的实拍照片与生成图像区分算法，要么准确率和召回率偏低，要么算法效率较低。当前一种基于图像颜色数、主颜色率、色彩饱和度、颜色直方图相关性这几个特性综合判断的方法[1~4]，认为生成图像的颜色比较单一，主颜色率比较高，色彩饱和度高的像素点较多，并且计算机生成图的颜色直方图和统计所得的计算机生成图像的平均颜色直方图的相关性大于与统计所得的相机拍摄图像的平均颜色直方图的相关性。但是该方法的召回率和准确率都偏低。

图像在应用于网站显示之前，一般会对图像的尺度及颜色进行插值或采样处理，这些处理改变了原本图像的一些性质，如颜色数、主颜色率等，使得这些特征对生成图像和实拍图像的区分力度变弱，甚至不能区分。

另外，网络上显示的一些商品图像（如一件置于黑色背景上的白色衬衫），由于其本身的特性，图像本身的颜色数就比较低、颜色饱和度较高，故之前的判断方法不能很好地处理这类图像。

基于以上原因，本文加入了作用于颜色空间的垃圾图判断项，该判断项对图像进行分块处理，提取出各颜色分块的颜色关系特征进行判定，而不是简单地对整幅图像的主颜色率、饱和度进行判断。另外，本文分析图像包含内容，提出了根据文字区域所占比例来判断图像类型，与颜色空间的判断项具有很好的互补性。

2　技术路线

本文先利用图像的物理尺度与颜色数这两个简单的判断项，将那些容易区分的生成图像召回；其次，利用垃圾图像判断项，将颜色单一的生成图像召回；最后，利用文字图像判断项，将包含文字较多的图像召回。以上几个判断项，分别作用于图像的颜色空间与非颜色空间，故具有良好的互补性。经过如上几个步骤，大部分的生成图像已被召回，准确率约为 97.43%，时间效率约为 5.62 ms/p。下面对本方法作具体介绍。分析实拍照片与生成图像，可以发现两者在以下三个方面存在比较明显的区别，如图 1、2 所示。

1）图像物理尺寸　实拍照片尺寸较大，并且符合一定长宽比；生成图像一般尺寸较小，长宽比不一定。

收稿日期：2012-01-05；**修回日期**：2012-02-12

作者简介：李宗（1983-），男，新疆克拉玛依人，硕士，主要研究方向为多媒体通信（li_zong@foxmail.com）；徐家品（1957-），男，四川成都人，教授，主要研究方向为多媒体通信.

2)图像颜色空间　实拍照片的颜色较丰富;生成图像的颜色较单一。

3)图像内容　实拍照片包含的内容比较丰富,信息量大;生成图像包含的内容单一,很多包含文字。

图1　实拍照片

图2　生成图像

针对以上三个方面特征的区别,提出了下面的区分步骤,如图 3 所示。

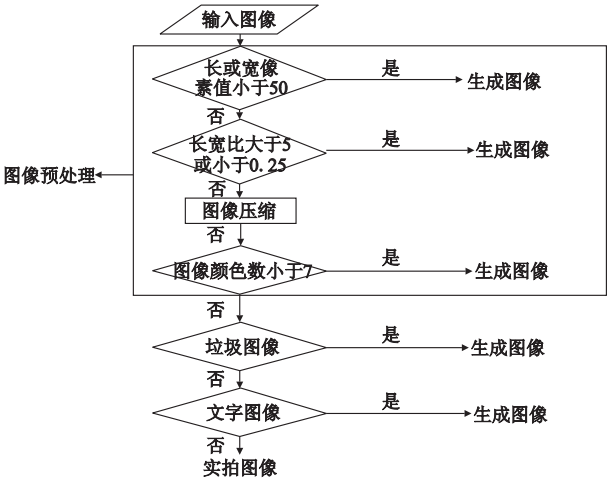


图3　算法流程

2.1　图像预处理

先利用一些简单的特征将那些在物理尺度、颜色空间上比较容易区分的生成图像召回,并对图像进行压缩。

1)图像长宽像素判断　长或宽小于一定像素的图像判断为生成图像。

2)长宽比判断　对于实拍照片,由于相机图像采集设备的形状一定,长宽比是有一定范围的。将图像长宽比大于或小于一定值的图像判断为生成图像。

3)图像压缩　对长宽像素大于一定值的图像按原始的长宽比进行压缩。图像压缩的目的是为了减少两个判断项的运行时间。从实验效果来看,图像压缩并未对算法效果产生明显的影响,但是大大减少了运行时间。

4)图像颜色数判断　将图像颜色空间深度进行压缩,统计图像包含的颜色数,对颜色数小于一定值的图像判断为生成图像。

将近 50% 的生成图像在这一步骤中已经被召回,该步骤准确率很高,约为 99.94%。同时,该步骤时间效率较高,经过该步骤处理后,减少了后续相对较耗时的判断的判断量,有效地提高了算法的效率。

2.2　垃圾图像判断

通过观察互联网上的生成图像可知,有很大一部分是颜色单一的图像,这些图像的主体颜色是白色、灰色或黑色。但是如果只通过整幅图像的颜色数、主颜色率等判断项来判断,则可能会导致严重的误判,故将图像进行分块处理,根据各个分块主颜色的关系来判断是否是生成图像。下面是该方法的具体步骤。

首先,对图像的颜色空间进行压缩,判断图像的主颜色是否为白色或黑色;如果是,则进一步将图像分成九个块,分别求取每个块中的主颜色。如果满足以下条件,则判定该图像为生成图像(图 4):a)图像主颜色为黑色,九个块的主颜色也为黑色;b)图像主颜色为白色,九个色块中有多个块的主颜色也偏白色,且图像的有效颜色数比较少,实现时,需根据具体的图像集和需求调整颜色数阈值;c)图像主颜色为白色,中间块的主颜色也为白色,且对角上存在主颜色为白色的块。

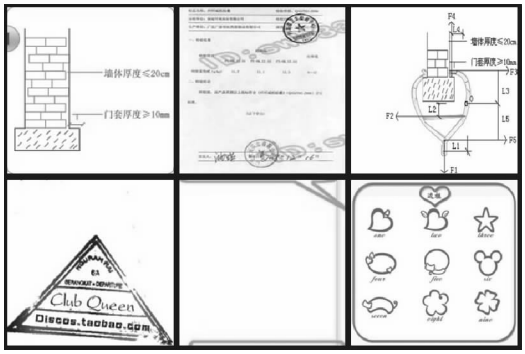


图4　垃圾图像判断效果

2.3　文字图像判断

生成图中有很大一部分是文字图像,这类图像在颜色空间上比较难将其与实拍照片区分开来。下面通过判断图像中文字区域所占比例来区分这类图像。具体步骤如下:

a)将图像转换为灰度图。对于文字图像,图像对比度的特征远大于颜色特征,故以下的操作均在灰度空间进行,这样也提高了运行效率。

b)利用 Sobel 算子[5~7]作边缘检测。文字图像具有比较明显的边缘特征,使用左对角和右对角 Sobel 算子对图像的边缘区域进行检测,可以增强文字图像的效果。之所以选用 Sobel 算子,一方面是因为该算子是各向同性的,因此对各个方向的笔画都能检测;另一方面,Sobel 算子能产生双重边缘,这就更加凸显了文字图像笔画的边缘特征。图 5 左边为原图,右边为边缘检测所得图像。

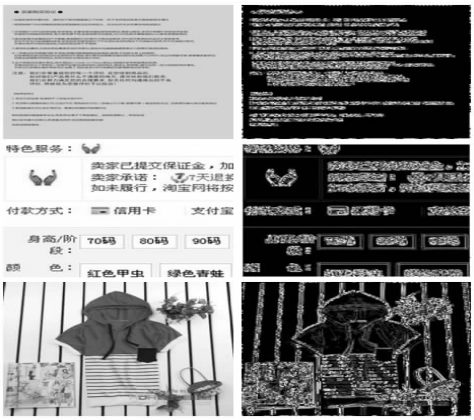


图5　Sobel边缘检测效果

c)文字区域检测。分析 Sobel 边缘检测所得图像,发现文字图像存在一个特征,文字的行与行之间存在明显的空白行。根据统计,未做压缩前,图像中文字的高度大多分布在 8 ~ 45 个像素之间,压缩后文字高度大多都低于 15 个像素。将边缘点图像分割为一个个 16 × 16 像素大小的窗口,分别对每一个窗口判断是否是文字区域:在窗口内作横向投影(建立在文字

是横向书写的假设之上),投影值小于一定值的行被认为是空白行,求取窗口内最大的连续空白行数目 Max_BL,之后根据 Max_BL 作判断:(a) Max_BL 小于一定的阈值,则认为在该窗口中没有文字存在,为非文字区域,因为如果有文字存在,不同行的文字之间应该会有间隔;(b) Max_BL 大于一定的阈值,同样认为该窗口为非文字区域。之后除去判断为非文字区域的窗口中的边缘点,保留大于一定阈值的判断为文字区域的窗口中的边缘点。图 6 为文字区域判断效果。

图6 文字区域判断效果

从图 6 中可以看出,文字图像边缘点基本都被保留了下来,而对于非文字区域的边缘点基本都被抑制。

d) 文字区域数目统计。经过以上两步处理,文字区域与非文字区域表现出了较明显的不同,文字区域更倾向于横向密集分布,而非文字区域倾向于无方向性的曲线分布。使用高 10 个像素、宽 40 个像素大小的矩形窗口,分别以 4 个像素和 8 个像素的移动量对整幅图像区域进行判断;在矩形窗口内,如边缘点个数大于一定阈值,则该窗口被认为是文字区域。统计整幅图像中文字区域所占比例,大于一定阈值的图像判断为文字图像。图 7 为文字区域统计效果示意图,左边为文字区域判断效果,右边为文字区域定位效果。

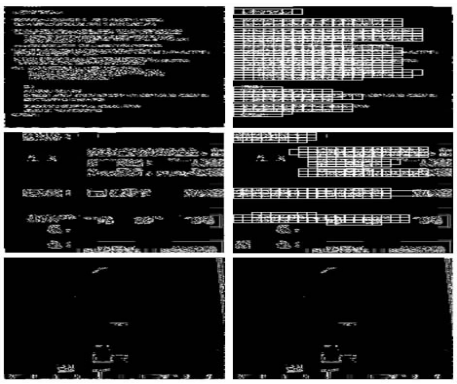


图7 文字区域统计效果

条纹状图案经过以上处理,表现出了与文字图像类似的分布特性,容易被误判为文字图像。细致分析条纹图案及文字图案,可发现条纹图案的边缘点倾向于呈线状分布,因此,计算 10×40 矩形窗内 10 行中相邻像素差值的绝对值之和 Sum_d,统计矩形窗中 Sum_d 小于一定值的行的数目,如大于一定阈值,则该矩形区域不记为文字区域。该判断项有效地解决了条纹状图案的误判,其效果如图 8 所示,图 8(a) 为原图,(b) 为 Sobel 边缘检测图,(c) 为文字区域判断效果图,(d) 为未加条纹判断项前得到的文字区域判断示意图,(e) 为加上条纹判断项后得到的文字区域统计示意图。从图 8(d)(e) 的对比可以看出,加入条纹判断项后,原先误判为文字的条纹图案基本都能正确判断。

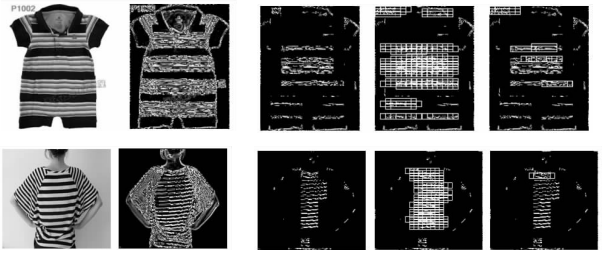
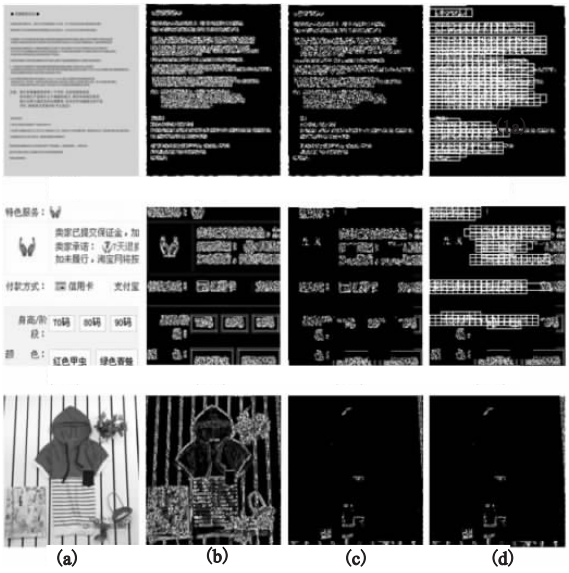


图8 条纹误判解决效果

图 9 为整个文字判断算法过程示意图,图 9(a) 为原图,(b) 为 Sobel 边缘检测结果,(c) 为文字区域检测结果,(d) 为文字区域定位结果。



(a) (b) (c) (d)

图9 文字图像判断过程

3 测试结果

通过以上几个步骤的判断,已经能够有效地区分出实拍照片与生成图像,对互联网上 26 000 多张图像的测试结果如表 1 所示。

表 1 测试结果

	准确率/%	召回率/%	运行效率/ms/p
预处理	99.94	53.95	0.61
垃圾图像判断项	93.92	13.85	2.56
文字图像判断项	92.46	18.11	3.12
完整	97.43	85.91	5.62

从测试结果可看出,本文方法在网络图像的测试中取得了优异的结果,准确率达到 97% 以上,召回率也在 85% 以上;其运行效率也很高,每秒能实现 170 多张图像的处理。相较于王永攀等人[4] 的方法,本文方法的召回率、准确率都有了很大提高,而时间效率稍低(文献[4]中方法的召回率为 60.4%,准确率为 49.8%,时间效率为 2.3 ms/p。)。本文提出的方法,达到了高效准确地区分实拍照片与生成图像的目的,从方法的召回率和准确率上来说都是简单有效的,具有很强的实用性。

参考文献:

[1] VASSILIS A, MICHAEL J S, CHARLES F. Distinguishing photographs and graphics on the world wide Web [C]//Proc of IEEE Workshop on Content-Based Access of Image and Video Libraries. 1997:10-17.

(下转第 3550 页)

(上接第 3542 页)

[2] 陆宗骐,梁诚. 用 Sobel 算子细化边缘[J]. 中国图象图形学报, 2000,5(6):516-520.

[3] NG T T, CHANG S F, HSU J, *et al*. Physics-motivated features for distinguishing photographic images and computer graphics[C]//Proc of the 13th Annual ACM International Conference on Multimedia. 2005: 239-248.

[4] 王永攀,吴诗娟,冯华君,等. 区分商品图中计算机生成图和相机拍摄图的方法[J]. 光电工程,2010,37(2):113-121.

[5] CHEN Yuan-hao, LI Zhi-wei, LI Ming-jing, *et al*. Automatic classification of photographs and graphics[C]//Proc of IEEE International Conference on Multimedia and Expo. 2006:973-976.

[6] ZHANG Jin-yu, CHEN Yan, HUANG Xian-xiang. Edge detection of images based on improved Sobel operator and genetic algorithms [C]//Proc of International Conference on Image Analysis and Signal Processing. 2009: 31-35.

[7] QU Ying-dong, CUI Cheng-song, CHEN San-ben, *et al*. A fast sub-pixel edge detection method using Sobel-Zernike moments operator [J]. Image and Vision Computing,2005,23(1): 11-17.

[8] CUTZU F, HAMMOUD R, LEYKIN A. Distinguishing paintings from photographs[J]. Computer Vision and Image Understanding, 2005,100(3):249-273.