

一种以二值图像格式保存纯色信息的二值化简单彩色图像压缩方法*

A Simple Color Image Compression Method of Binary Image Format to Save the Color Information

黄孜祺¹, 廖宏¹, 卿旭², 吴彩荣²

HUANG Zi-qi¹, LIAO Hong¹, QING Xu², WU Cai-rong²

(1. 广西计算中心, 广西南宁 530022; 2. 南宁海蓝数据有限公司, 广西南宁 530022)

(1. Computing Center of Guangxi, Nanning, Guangxi, 530022, China; 2. Highland Digital Technology INC., Nanning, Guangxi, 530022, China)

摘要:二值化简单彩色图像压缩方法通过颜色分离、二值化处理等操作,以二值图像格式高效率压缩保存包含有类似蓝色信笺头、红色公章、黑色图文等单纯颜色信息的影像文件,而且还还原后能够原文原貌地恢复压缩前的图像和色彩内容,压缩效率比常用的8位以上彩色压缩方法提高8~10倍。二值化简单彩色图像压缩方法压缩7种以下单纯颜色图像文件的效果最为明显。

关键词:图像压缩 二值图像 颜色信息

中图分类号: TN911.73, TP317.4 **文献标识码:** A **文章编号:** 1002-7378(2009)04-0259-02

Abstract: A binarization simple color image compression method can preserve 6~7 kinds of pure color information of image documents by the way of color separation, binarization processing operations, such as blue head letter paper, red official stamp or black image characters. The compression efficiency ratio enhances 8~10 times by comparing with commonly 8 above colored compression method. In the color type is short in 7 kinds of situations reduces the effect to be most obvious.

Key words: image compression, binary image, color information

随着多媒体应用的普及,数字影像技术的发展,以及网络图像传输的增多,对图像压缩的研究已经变得越来越重要。因为图像的数据量庞大,存储和传输都比较困难。如果单纯靠扩大存储器容量、增加网络带宽的办法来解决问题是不现实的,通过数据压缩技术可以大大降低数据量。图像压缩技术使图像以压缩的形式存储和传输,既节约了存储空间,又提高了通信的传输效率。目前通常以二值图像压缩方法(如 CCITT TIFF Group4等)来保存纸质文件上的黑白信息,以8位(256色)以上的彩色图像压缩方法(如 JPEG)来保存彩色信息。但是,在日常的数字化加工工作中,政府部门、档案馆、税务、银行、保险、电信等单位资料中有相当多的,除黑色字符外还包含红(公章)、蓝(信笺头)、黄、绿色等简单彩色的

纸质文件,而且这些文件资料又必须原文原貌显示,有严格的数字化处理要求。如果采用通用的彩色图像压缩方法,大规模加工处理后会有较大的冗余,压缩存储效果也并不明显。因此,我们研究了二值化简单彩色图像压缩方法,高效率地压缩只有7种以下单纯颜色的图像文件,以便于提高压缩效率,节省图像文件的存储空间^[1]。

1 图像颜色分离

二值化简单彩色图像压缩方法压缩一个包含有红、蓝、黑、白(背景色)4种色彩信息的图像文件,首先是对通过扫描仪扫描获得的24位包含多色信息的彩色位图文件进行颜色分离操作。在该位图文件中,每一像素点的颜色信息都是以 RGB 三基色来表示的,需要先进行颜色空间转换,再进行颜色分离。所以,颜色分离操作可以分为颜色空间转换、色彩分离及颜色抖动二值化处理3个基本步骤(图1)^[2,3]。

收稿日期: 2009-09-21

作者简介: 黄孜祺(1982-),女,助理工程师,主要从事图文资料数字化技术和图像压缩技术研究工作。

* 广西科学基金项目(桂科基0731027)资助。

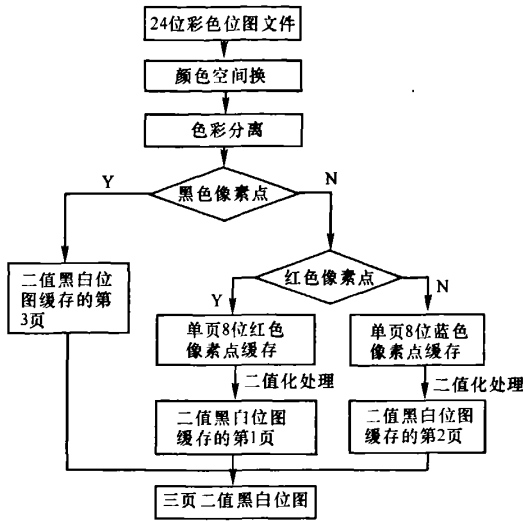


图1 色彩分离过程

颜色空间转换是将像素点的颜色表示方法由 RGB 色彩空间转换到 HSV 色彩空间表示,把原来用红色、绿色及蓝色表示的像素点颜色,转换成用色彩亮度、饱和度及色度来表示。颜色分离是在 HSV 颜色空间中对位图进行色彩分离,产生一个分别对应于红、蓝、黑三色的三页二值黑白位图,将被判断为黑色的像素点直接被放入到三页二值黑白位图缓存中,而被判断为红色及蓝色像素点则要经过贝叶斯抖动二值化处理后放入三页二值黑白位图的缓存中。位图中所占面积最大的颜色(比如白色)按背景色处理。如果背景色不是白色,而是黄色、绿色或其他颜色,则将其采用贝叶斯抖动方法二值化后存入背景色信息存储单元。如果样张图像上的纯色色彩超过4种,比如有 N 种色彩,则产生 $N-1$ 页二值黑白位图。二值化处理后将各色彩值存入各页对应的色彩信息存储单元,每一页可以任意采用一种通用或专门的二值图像压缩方法压缩,最终形成保存多种纯色信息的高压缩比图像文件^[4]。

2 图像解码还原

二值化简单彩色图像压缩方法压缩一个包含有红、蓝、黑、白(背景色)4种色彩信息的图像文件后还要对图像进行解码还原。解码还原先从背景色存储单元中提取色彩信息还原显示背景色,然后再对每一页被压缩的二值图像的位图解压缩,将其位图内容还原并赋予颜色信息存储单元中保存的彩色值后在背景色上叠加显示。第1页中的黑色像素点解码还原显示为红色,第2页中的黑色像素点解码还原显示为蓝色,第3页中的黑色像素点解码还原显示为黑

色。颜色多于4种的情形按此类推^[5]。

3 实际应用对比

对一份只含有黑色文字、红色印章、蓝色信笺头和领导签字的合同,分别采用二值化简单彩色图像压缩方法和 JPEG 压缩方法进行压缩的结果如图2所示。从图2结果可以看出,用二值化简单彩色图像压缩方法压缩图像形成的文件大小为111KB,用 JPEG 压缩方法压缩图像形成的文件大小为305KB。二值化简单彩色图像压缩方法的压缩率比通用的 JPEG 压缩方法提高了8~10倍,是一种颇具优势的彩色图像文件压缩工具。

名称	大小	类型	修改日期
样图 hls	32 KB	HL Image Format	2009-8-24 11:25
样图 jpg	111 KB	ACDSee 8.0 JPEG Image	2009-8-25 15:23
样图 bmp	305 KB	ACDSee 8.0 BMP Image	2009-8-25 15:26

图2 两种不同格式的样图文件大小对比

4 结束语

二值化简单彩色图像压缩方法既能保存图像的色彩又能拥有二值图像的高压缩率,充分利用多色图像的特点,用尽可能少的数据量来表示图像,尽可能消除图像冗余,节省图像的存储空间。目前该方法已经与作者单位自主研发的大规模图文资料数字化系统、大规模工程图纸数字化系统结合应用,既能大大节约图像的存储空间,又能保证图像的原文原貌显示。二值化简单彩色图像压缩方法对含有7种以下单纯颜色图像文件的效果最为明显,同时还适用于多种应用系统,在电子图书、电子期刊、数字档案馆以及电子政务、电子商务等领域具有广阔的推广应用前景。

参考文献:

- [1] 陈静. 静态图像压缩标准回眸[J]. 计算机应用与软件, 2005, 22(9): 130-132.
- [2] 傅德胜, 谢永华. 一种基于小波变换的分形图像编码压缩算法的研究[J]. 计算机应用与软件, 2004, 21(1): 66-68.
- [3] 杨波, 汪同庆, 叶俊勇, 等. 一种二值图像的有损压缩方法[J]. 计算机应用, 2002, 22(3): 37-38.
- [4] 颜彬, 陈传波. 基于四叉树的二值图像时域分层压缩算法[J]. 计算机工程与应用, 2002, 38(9): 69-71.
- [5] Chang Y C, Shyu B K. Adaptive post-processing for region-based fractal image compression [C]. IEEE Proceedings of the Data Compression Conference, 2000.

(责任编辑: 邓大玉)