

四值图像电子文档压缩算法

Quaternary image compression algorithm for electronic document

2017-09-04 发布

2018-01-01 实施

上海市质量技术监督局 发布

目 次

前言 III

引言 IV

1 范围 1

2 术语和定义 1

3 四值图像电子文档压缩的标准流程 2

4 四值压缩算法技术要求 2

5 四值压缩算法的应用 3

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准由上海市人民政府办公厅电子政务办公室提出并归口。

本标准由上海市人民政府办公厅电子政务办公室、上海金刚石数码研究中心负责起草。

本标准主要起草人：孙松涛、潘晓虹、王治君、姜执伟、陈黄海、白玉柱、潘海天、董瑜、李文、杜赫铭、姜毅、陶志锋。

引 言

为了满足四值图像电子文档压缩算法标准化需求,根据当前使用的电子文档(包括文本文件、电子表单、数字凭证),存在的体量较大、针对性压缩算法少、且已有的电子文档压缩算法的知识产权,可能存在“巨额使用费”陷阱的现状,针对中国文档只有四种颜色(四值)的特点,有必要提出一种需要针对电子文件的标准化的高效压缩算法。

采用本标准的 TIFF 图像文件,作为电子政务与电子商务的电子文档,可以绕过国外软件企业知识产权限制和“巨额使用费”陷阱。

四值图像电子文档压缩算法

1 范围

本标准规定了四值图像电子文档(以下简称四值图像)压缩的标准流程、技术要求和应用等。
本标准适用于上海市电子政务与电子商务的四值电子文档的产生与存档。

2 术语和定义

下列术语和定义适用于本文件。

2.1

无损压缩 **lossless compression**

一个比特位都不能丢失、修改与添加的压缩技术。

2.2

文本文件 **text documents**

一种由若干行字符构成的计算机文件作为一种容器存储特定编码格式的字符串。

2.3

电子表单 **electronic form**

电子政务领域的网上申报、审批、内部办公、政务信息采集、报送等形式的电子文档。

2.4

数字凭证 **digital document**

一种加载安全机制,采用多种安全技术的具有防伪、防篡改、可加载与读取信息的,由一张或者若干张尺寸不等的图片组成的电子文档。

2.5

电子文档 **Electronic document**

人们在社会活动中形成的,以计算机盘片、磁盘和光盘等化学磁性材料为载体的电子文档与档案。

2.6

四值图像 **Quaternary image**

由四种颜色组成的图像文件:白色底色、黑色文字、红头文字、红章盖在黑字上。

2.7

24 Bit 位图 **24-Bitmap**

由 RGB3 个字节的位图图像(Bitmap),可以表达为 2 的 24 次方,约 1 600 万种颜色。

2.8

复合压缩 **Composite compression**

通过两种及其两种以上的压缩算法,复合实施电子文档压缩处理的方法。

2.9

四值预压缩 **quaternary precompression**

将采用位图技术且其颜色不超过四种的电子文档,对其位图像素点用 2 个 Bit 数字表示(预处理)的压缩过程。

注:反之,解压缩过程也称之为四值解压缩。