

# 基于动态阈值分块算法的屏幕图象压缩技术

朱芳宇 朱光喜 王 曜

(华中科技大学电子与信息工程系宽带无线与多媒体系统研究中心, 武汉 430074)

**摘 要** 计算机屏幕图象压缩一直是屏幕共享技术中的关键问题, 而现有算法一般是直接应用传统的 RLE, LZW 或 JPEG 算法, 但由于这几种算法在对计算机屏幕图象进行压缩时, 均无法兼顾恢复质量与压缩比的要求. 为此提出了一种基于动态阈值分块的高效压缩算法, 该算法首先将屏幕图象按照某种规则划分为纯色块、文字块、图标块和图象块 4 种特征块, 然后针对特征块特点采用相应的压缩算法, 以保证压缩算法在恢复质量与压缩比上能同时达到最佳. 实验结果表明, 该算法图象恢复质量较好, 压缩比较大, 压缩速度较快, 整体性能优于当前现有的屏幕压缩算法, 具有广泛的应用前景.

**关键词** 图象处理(510·4050) 屏幕共享 动态阈值 分块压缩

**中图分类号**: TN919.81 TN911.73 **文献标识码**: A **文章编号**: 1006-8961(2003)09-1084-05

## A Screen Image Compression Technology Based on Sub-blocking with Dynamic Threshold

ZHU Fang-yu, ZHU Guang-xi, WANG Yao

(Department of Electronics & Information Engineering, R&D Center of Broad Band Wireless Communication & Multimedia System, Huazhong University of Science & Technology, Wuhan 430074)

**Abstract** Compression algorithm for screen images, which comes from the Computer Screen, is an important task in the Screen Share technology. Traditional methods are based on the algorithms such as RLE, LZW or JPEG directly, which cannot meet the requirements of both the quality of recovery image and compression ratio to compress screen image respectively. High compression ratio algorithms (JPEG) result in a low quality of image, while no-distortion algorithm (RLE, LZW) leads to a low compression ratio. In order to resolve this contradiction, an effective compression algorithm based on dynamic threshold sub-blocking is proposed in this paper. Before the image compression, Screen image is divided into four sorts of blocks by certain provided rules first, which are named as Pure-color Blocks, Text Blocks, Icon Blocks, and Image Blocks. Then these blocks are compressed separately depended on their characteristic so as to ensure that the quality of recovery image is the overall best and the compression ratio is the highest. Pure-color Block, Text Block, and Icon Block are compressed with no-distortion compression algorithm, and Image Block is compressed with JPEG algorithm. Result of experiments shows that our algorithm has improved quality of recovery image, higher compression ratio, and higher compressed speed than traditional algorithm.

**Keywords** Screen share, Dynamic threshold, Sub-block compression

## 0 引言

随着计算机网络应用技术的发展, 人们开始利用网络去实现诸如远程控制、监视、计算机支持的协同工作 (Computer Supported Cooperative Work, CSCW) 等方面的应用, 这方面的需求直接促使了屏

幕共享技术的出现. 屏幕共享的目的是为了在远端客户机上重现主机屏幕的显示内容, 或是以一种虚拟终端的概念出现, 以便使得远端的用户感觉如同在主机旁边观察屏幕一样, 但在当前的计算机网络条件下, 实时屏幕共享并不是一件很容易的事, 如以典型的屏幕图象为例, 在没有任何压缩的情况下, 每秒传输的数据量最低也有 46M bit, 这么大的数据量在当前的

10M/100M的局域网中是难以实现的。

因此,要想实现屏幕图象共享,高效的压缩算法是必须的前提条件。在绝大多数情况下,计算机屏幕图象存在着很大的空间相关性,这表现在一幅屏幕图象的连续像素之间有着很强的相似性,尤其在 windows 操作平台中,在多数情况下,其基于窗口的 GUI(Graphical User Interface)界面,经常出现大片颜色相同的背景块,如单一颜色的桌面、应用程序的背景、按钮及窗口的边框等。与此同时,由于操作系统的限制,致使屏幕上出现的颜色数目是离散,且有限的。屏幕图象的这些特性为实现高效率、高质量压缩算法提供了理论上依据。目前有很多的产品已实现了这一功能,比如 XWindow、微软 Windows 2000 的远程终端服务、NetMeeting、SYMANTEC 的 PC Anywhere、Remote Administrator、VNC (Visual Network Computing)<sup>[1]</sup>以及国内的红蜘蛛教学系统等。但在这些已有的系统中,使用的压缩算法基本上可以分为无失真压缩编码 RLE (Run Length Encoding),LZW (Lempel-zir Walch) 等及有失真编码 (JPEG, GIF 等) 两类,但这两类算法各有优点:无失真压缩编码速度快,解码后的图象质量好,尤其对于大块纯色区域或文字区域;而有失真算法虽压缩率高,尤其对于色彩复杂的图象区域,压缩比可以达到很高,但是,它们的缺点也相对明显,如无失真算法的高速度,高质量是以降低压缩比为代价而得到的,当图象较为复杂时,压缩比很低,这将影响传输效率;而 JPEG 算法虽然压缩比较高,但编码速度较慢,而且对于文字区域的恢复效果不好,也不太适用于屏幕压缩。

不过,通过比较发现,这两种算法的优缺点恰好互补,不难想象,通过某种方法将它们有机结合起来,取长补短,应该可以组成一种正好适应屏幕图象特点的新型压缩算法。本着这种思路,在充分研究屏幕图象特点的基础上,提出了一种“多类分块压缩”(Multi-Type Sub-Blocking Compression, MTSB) 算法,试验数据表明,该算法在恢复图象的质量,压缩效率和编码速度等 3 个方面都取得了较好的结果。

## 1 MTSB 算法

### 1.1 MTSB 算法原理

通过对屏幕图象的充分研究发现,计算机屏幕图象具有以下特点<sup>[2]</sup>:

(1) 在绝大多数情况下具有色彩简单、以文字为主的特性;

(2) 少数部位具有相关性小,颜色丰富的特性;

(3) 如果有图片,则总可以找到一个最小的矩形使得图片完全居于其中;

(4) 如果有复杂背景,且其背景和前景对比明显,则可以通过一定算法分离。

根据这些特点,可以认为屏幕图象主要由以下几种特征区域组成:第 1 类是纯色部分,如单色桌面,资源管理器的白色背景等;第 2 类是文字部分,也就是单色背景上加上某一颜色的文本形成的区域,最典型的代表就是 Word 文档;第 3 类是图标部分,这种由 Windows 系统创建的显示标识,具有色彩简单,边缘锐利的特点,各种工具栏按钮也属于这一部分;最后一类是自然图象部分,这些图象大部分拍摄自现实世界,具有色彩复杂,边缘模糊等特点。这 4 种区域的交错排列,就构成了计算机屏幕上显示的各种图象。在本算法中,将这 4 种区域分别称为纯色区域,文字区域,图标区域,图象区域。针对各个区域的不同特点,可以采用不同的压缩算法进行压缩,以达到整体最佳的压缩比。

(1) 对于纯色区域,仅传输其任意一个点的颜色值便可以无失真恢复该区域,对多个连续的纯色区域,还可用块间简单预测的方法来减少信息的冗余。

(2) 对于文字区域,由于数据块中存在局部二值化的特点,因此,可以仅传输文字区域的前景色和背景色,并按照前景色和背景色将该区域二值化,然后传输以 bit 为单位二值化的图象。在解压缩的时候,首先解码出二值化的图象,然后用前景色和背景色代替 1 和 0,即可无失真地恢复图象。对于二值图象的压缩,目前已经有非常成熟的算法可以利用<sup>[3]</sup>。

(3) 对于图标区域,该类区域具有纹理简单且有规律、颜色过渡少、高频成分多的特点。实验证明,如果采用 JPEG 等变换编码方式,其编码效率还不如无失真编码方式,且图象恢复质量较差,如图 1 所示。因此,该区域不适合使用变换编码或类似的失真算法来处理。对于图标区域,通过研究发现,虽然图标区域的颜色不止一种,但却很有限(一般为 16

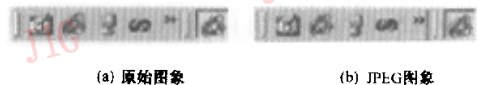


图 1 人工图象 JPEG 压缩效果对比图

色),因此可以根据具体编码目标动态地建立一个调色板去保存该区域使用的颜色值,再将数据颜色值转换为调色板中的颜色索引,然后按照一般无失真数据压缩的办法处理。

对于图象区域,由于 JPEG 算法能够取得很好的编码效率和压缩效果,因此,对于屏幕图象中的图象区域,使用 JPEG 算法压缩是最佳的选择。

## 1.2 算法流程

根据上面的论述,MTSB 屏幕图象编码流程如图 2 所示。

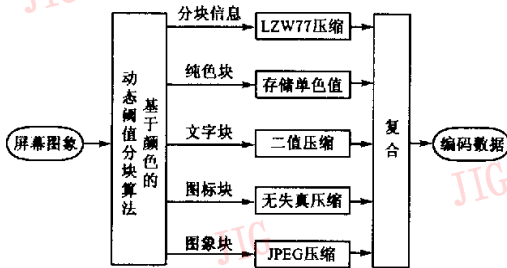


图2 MTSB 屏幕图象编码流程

用 MTSB 算法编码时,屏幕图象首先是被划分为  $16 \times 16$  的数据块;接着,根据一定算法将每一个数据块划分到纯色块、文字块、图标块、图象块等 4 种类型块中;同时保存并采用 LZW77 算法压缩分块信息;然后对每一个数据块按照不同类型块的特点,采用相应的压缩算法进行编码;最后将各个数据块的压缩数据复合,即得到编码后数据。

## 2 基于颜色的动态阈值分块算法

由于 MTSB 压缩算法的基本思想是根据屏幕图象不同区域的特征来利用不同方法进行压缩,因此,根据数据块的特点将其判定为某一类型块就成为整个压缩算法的关键所在,这将是算法是否有效及效果是否最佳的主要影响因素。

从提高压缩率的角度来看,数据块是适用无失真算法,还是运用有失真算法处理,关键在于数据块中包含的数据值在统计上的集中和重复程度,若集中程度越大,就越适合于使用 Huffman、RLE 或者基于字典的 LZW 算法进行压缩处理。在屏幕图象中,数据块的数据就是像素颜色值,而通过扫描得到的数据块像素颜色值序列,可以视为一维信号。由于图标区域颜色简单,像素取值只占所有颜色可能值

的极少一部分,故而基本不存在颜色过渡区域;而图象区域则存在大量过渡区域,颜色复杂,且在取值范围中分布很广。由此可以猜想,如果某一数据块颜色简单,则可以将其判断为图标区域,如果颜色复杂,则视为图象区域。

如果上述推断正确,那么对于颜色数目少于某一阈值的块,就应该选择无失真算法去处理。为证明这一结论,因此进行了一系列验证实验,实验方法是:每次先选取一个阈值,然后用该数值与实验数据块 ( $16 \times 16$ ) 包含的颜色数目进行比较,如果数据块颜色数目小于或等于该阈值,并且该数据块无失真压缩后的尺寸小于有失真压缩后的尺寸(DCT 量化系数选择 60%),则认为判断成功,否则作为判断失误;相反,如果判定结果为自然图象,而且有失真算法获得了更高的压缩率,那么也认为是判定成功。在选取不同的数值作为颜色数目判断阈值(简称判断阈值)进行实验后,就得到了如图 3 所示的曲线。

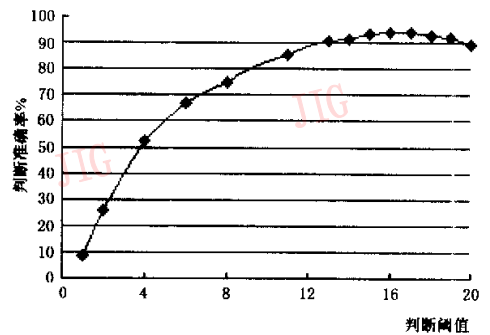


图3 分块算法中不同判断阈值取值的试验结果

由实验结果可以看出,当判断阈值为 16 时,算法取得了最高的正确判断率,然而如果判断阈值恒定为 16,那么从对实际图象的压缩过程中发现,解码图象在大面积的自然图象区域内部有时会出现部分马赛克现象,这就严重地影响了视觉效果。经过分析得知,这些色块均是因为在识别时被认为是人工图象,而使用了无失真算法所造成;由于周围被认为是自然图象的数据块使用的是 JPEG 算法压缩,从而在两类区域的边缘上产生了明显的块效应,以致出现了马赛克现象。之所以将自然图象中的一部分识别为人工图象,则是因为自然图象中也有部分区域颜色简单的缘故。由此可见,仅仅通过固定阈值这种简单的判断是无法较好地地区分数据块类型的。

为此提出了一种动态阈值的分块算法,即根据

屏幕图象的特点来进行压缩的算法,由于多数图象显示区域都是矩形,因此如果当前数据块被判断为自然图象时,则可以认为扫描已经进入矩形图象区域.此时,可进行如下处理:

(1) 降低判断阈值.阈值的降低将使判断过程包容图象区域中随机出现的其他类型块,而不致于将一块图象区域识别为几个交错的图标块和图象块的组合.

(2) 继续扫描,在新的阈值下对后续数据块进行判定,如果数据块是纯色块、文本块时,则恢复初始阈值,并认为扫描已经离开了图象区域.

通过上述动态阈值的处理,图象区域已基本与屏幕其他部分分离,此时可进行校正处理.校正的准则为:

① 如果图标块上下左右4个相邻块中,有2个或2个以上被标记为图象,则该图标块应直接标记为图象块.

② 反之,如果图象块的相邻数据块中有2块或2个以上被标记为其他类型,则该图象块应直接标记为图标块,并应使用无失真压缩算法,以避免与相邻区域产生明显色差.

实验表明,经过校正的分块表,能够更好地区分屏幕图象中的各种不同数据块,这样就为下一步的压缩提供了最佳的基础.

### 3 算法性能测试

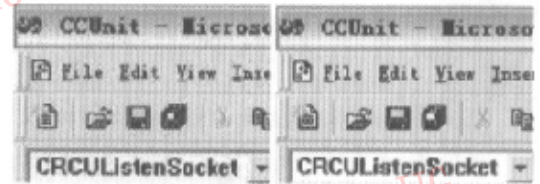
对算法的性能测试,主要包括压缩率和压缩后数据解码图象质量的测试.为使性能测试结果具有普遍意义,特别选取了几种典型的显示场合的计算机屏幕图象做实验,结果如下:

压缩率的测试,包括直接计算各种模式下的压缩比以及在相同条件下与当前最流行的JPEG算法的压缩效果进行比较两个方面,由于JPEG算法在不同量化系数下压缩比有很大的差别,为使压缩结果具有可比性,采用在相同信噪比下来比较不同算法的压缩率,测试和比较结果如表1所示,图4则是在相同模式下,MTSB算法与JPEG算法压缩后得到的解码图象效果图.

从试验结果可以看出,在相同的信噪比下,除了极为少见的全屏为自然图象外,对其他屏幕图象,MTSB算法都能比JPEG算法取得更好的压缩效率,压缩比平均提高了3~6倍,对多数屏幕图象的

表1 相同信噪比下MTSB算法与JPEG算法压缩比比较

	信噪比(dB)	压缩比	
		MTSB	JPEG
桌面	22.4	208:1	72:1
资源管理器	31.7	78:1	13:1
WORD	38.1	68:1	18.3:1
自然图象	18.6	41:1	5.3:1
网页	29.6	44:1	8:1



(a) JPEG(1024×768,79KB) (b) MTSB(1024×768,27KB)

图4 MTSB算法实验结果及与JPEG算法主观效果对比

解码效果也比JPEG算法要好得多,而对自然图象,压缩比稍差的原因主要是由于MTSB算法在处理自然图象的时候与JPEG的算法存在较大差别而造成的,因此算法在对自然图象处理方面尚存在较大的改进余地.实验数据证明,MTSB算法对压缩绝大多数情况的屏幕图象都能够取得很好的效果.

### 4 总结

在针对计算机屏幕图象的应用领域,MTSB算法与当前流行的JPEG算法相比有着较大的优势,其作为实现屏幕共享、课件制作的核心技术,有着广泛的应用前景.该算法已经在某公司产品中得到应用,已取得了较好的经济效益,并具有良好的市场前景.

本算法如果能在分块部分进一步推广,以使其在对自然图象的压缩方面也能达到JPEG算法的水平,那么很有机会将其推广到现在适用JPEG算法压缩的所有领域,而不是作为仅针对屏幕共享这一特定领域的专用算法.这方面将是在后续工作中重点研究的问题.

### 参考文献

- 1 WinVNC 3.3 (Visual Network Computing)[EB/OL]. <http://www.uk.research.att.com/vnc>, 1998-10.
- 2 陈琦,李华,朱光喜.一种新的应用于屏幕共享的图象块识别算法[J].电讯技术,2000,40(6):59~63.
- 3 任明武,曹雨龙,唐振武.黑白文档图象的高倍压缩[J].计算机研究与发展,1999,36(7):888~891.



**朱芳宇** 1978年生,2001年毕业于华中科技大学电子与信息工程系,现为华中科技大学2001届通信与信息系统专业硕士研究生,研究方向为图象压缩及多媒体数据通信。



**王隳** 1976年生,华中科技大学电子与信息工程系2000级博士研究生,中国数字音视频编解码技术标准专家组成员,主要研究方向为多媒体通信中的音视频编码与传输。



**朱光喜** 1945年生,现为华中科技大学电子与信息系教授,博士生导师,系主任,中国数字音视频编解码技术标准专家组成员,长期从事计算机图象图形处理、多媒体通讯等领域的工作,获得多项研究成果,现主要从事 CSCW、多媒体通信等工作,发表论文近百篇。

## 专业用户的坚强后盾——戴尔 Precision 工作站

一台高性能、高品质、功能强劲、稳定可靠的工作站会给日常工作带来多大的改变?如果你从事的是3D设计、动画制作、建筑工程设计、金融分析、地质勘测,或者任何一种需要密集计算的工作,相信你对此会有着深刻的体会。正是基于这种体会,再附以特有的灵活性和卓越的性价比,使戴尔 Precision workstation 在国内专业用户之中赢得有口皆碑的一致赞许。

戴尔 Precision 360 就是近期颇受好评的一款性能价格比优异入门级工作站产品。从核心架构到外围设备, Precision 360 无不围绕着实际应用过程中用户可能对工作站提出的要求而设计。它采用了一系列最新技术,包括英特尔 875P 芯片组、奔腾 4 处理器、超线程、串行 ATA、双通道 D6DR 400 SDRAM 内存等等,由此带来的整体性能的全面提升,可显著改进大多数复杂计算的运行效能。由于相当一部分的工作站用于 2D 或 3D 设计等领域,因此图形性能至关重要, Precision 360 不但可支持最新的高端图形技术,更允许用户根据自身工作需要选择 AGP 8X 或 AGP 8X Pro50 图形卡,这就避免了用户的投资浪费。鉴于越来越多的企业开始架设千兆网络,为了解决大量数据传输中的瓶颈问题, Precision 360 特别集成了千兆网卡。全新设计的台式机机箱还使 Precision 360 拥有强大的系统扩展能力,从而为企业业务的变化增长预留了充足的空间。无论是综合指数还是单项成绩, Precision 360 都堪称专业工作站应用领域的一个理想选择。

对许多专业用户来说,一款普通的高端工作站也许已经足以满足需要,但要完成建筑设计、油田勘探、图形设计乃至现场演示等等无法固定在一个地点的工作任务,现有的大多数工作站产品就显得有些力不从心,而笔记本电脑又无法满足工作站级应用对系统性能的苛刻要求。从这种意义上讲,戴尔 Precision 移动工作站的出现可谓填补了这一空白。其最新产品 Precision M60 移动工作站将灵活的移动性和主流台式工作站的性能合二为一,在笔记本电脑的轻巧外观内提供了强大的运算能力和图形性能,这不仅有效解决了现场办公的问题,还使其他工作站用户也可以彻底摆脱地点的束缚,能随时随地享受到移动计算的便利,已成为提高工作效率的一条捷径。它配备了最新的英特尔迅驰技术,拥有先进的带有 128MB 显存的高端图形卡、高分辨率宽屏显示屏以及工作站级的 OpenGL 硬件加速功能,从而使用户在运行大型设计程序或进行多窗口应用时得以获得完美的应用体验。软件方面, Precision M60 也通过了 MCAD、DCC、AEC 和 MCAE 等业内领先的应用软件认证,不但进一步增强了图形性能,更为用户日常工作的稳定性和高效率奠定了牢固基础。

目前,戴尔工作站产品已广泛深入到国内计算机辅助设计、工程计算、数字内容创作、地理信息系统、电脑动画制作、软件开发和财务分析等各个专业领域。凭借着其特有的直销模式,戴尔可以利用用户群体的日益普及更加敏锐、准确的把握用户需求的发展趋势,并在第一时间灵活的予以满足和引导。同时,以获奖服务团队提供的戴尔服务与支持,也为专业用户构筑了一道坚强后盾,保障企业业务顺利、高效的运行。