

# 基于颜色聚类的计算机桌面图像压缩算法

卢毓海<sup>1,2</sup>, 沈燕飞<sup>2,3</sup>, 王春洁<sup>1,2</sup>, 朱珍民<sup>2,3</sup>

(1. 湘潭大学信息工程学院, 湖南 湘潭 411105; 2. 中国科学院计算技术研究所, 北京 100190;

3. 移动计算与新型终端北京市重点实验室, 北京 100190)

**摘 要:** 针对计算机桌面图像压缩问题, 提出一种基于颜色聚类的图像压缩算法。将桌面图像划分成  $16 \times 16$  的非重叠块, 归为文本/图形块、自然图像块及混合块 3 类。对色彩丰富文本/图形块进行颜色聚类, 以降低块的颜色种类数, 并做无损压缩。对自然图像块采用 H.264 帧内预测编码方法, 对混合块采用混合编码方法。实验结果表明, 该算法所得图像的峰值信噪比和结构相似度均高于传统算法。

**关键词:** 计算机桌面图像; 混合图像; 块分类; 颜色聚类; 图像压缩; 图像编码

## Compression Algorithm for Computer Desktop Image Based on Color Clustering

LU Yu-hai<sup>1,2</sup>, SHEN Yan-fei<sup>2,3</sup>, WANG Chun-jie<sup>1,2</sup>, ZHU Zhen-min<sup>2,3</sup>

(1. College of Information Engineering, Xiangtan University, Xiangtan 411105, China;

2. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;

3. Beijing Key Laboratory of Mobile Computing and New Device, Beijing 100190, China)

**【Abstract】** This paper proposes an image compression algorithm based on color clustering to compress computer desktop image. It divides computer screen image into  $16 \times 16$  non-overlapping blocks, then each block is classified into text/graphic block, hybrid block or picture block. For text/graphic blocks with rich color, a color clustering method is used to reduce the number of colors, then text/graphic block is coded by lossless compression. Natural image block is coded by H.264 intra prediction coding method. Hybrid block is coded by hybrid coding method. Experimental results show that the image by proposed algorithm has higher PSNR and SSIM than traditional algorithms.

**【Key words】** computer desktop image; compound image; block classification; color clustering; image compression; image coding

DOI: 10.3969/j.issn.1000-3428.2012.21.059

### 1 概述

随着计算机应用技术的不断提高和计算机网络的迅速发展, 人们开始利用网络去实现远程控制、监视计算机支持的协同工作, 并且得到越来越广泛的应用, 主要有网络多媒体会议、远程办公、远程教学、产品演示、股票分析系统以及基于桌面共享的云计算应用<sup>[1]</sup>等。这些应用直接促使了桌面共享技术的产生和发展, 但由于计算机桌面图像的信息量巨大, 实时传输前必须对桌面图像进行压缩。

计算机桌面图像包含了文本、图形和自然图像信息, 与扫描文档图像一起统称为混合图像。传统的图像和视频压缩标准 JPEG 系列、H.26X 系列和 MPEG-X 系列等, 是基于人类视觉的敏感特性和自然图像色调连续的特征而制定的, 它们对于自然图像都能达到一个良好的压缩效

果, 但对于同时包含文本/图形信息与自然图像信息的混合图像, 直接应用这些传统压缩标准来压缩都不能达到理想的压缩效率与压缩质量, 如解码图像容易产生明显的振铃效应, 因此, 传统压缩标准并不适合于计算机桌面图像的压缩。

为此, 本文对计算机桌面图像编码方法进行改进, 提出一种基于颜色聚类的新的计算机桌面图像压缩算法。采用简单、低复杂度的块结构, 根据图像直方图统计信息将图像分为文本/图形块、自然图像块以及混合块 3 类, 对复杂的文本图形块进行颜色聚类, 降低块的颜色种类数, 然后对不同类别的块分别应用不同的压缩方法。

### 2 相关研究

目前计算机桌面图像的压缩编码主要分为基于层和基于块 2 类。基于层的算法根据 3 层混合光栅内容(Mixed

**基金项目:** 国家自然科学基金资助项目(61001123); 广东省中国科学院全面战略合作基金资助项目(2011B090300070)

**作者简介:** 卢毓海(1987—), 男, 硕士、CCF 会员, 主研方向: 图像编码, 视频编码; 沈燕飞, 助理研究员、博士; 王春洁, 硕士; 朱珍民, 正研级高级工程师

**收稿日期:** 2012-01-10

**修回日期:** 2012-02-29

**E-mail:** luyuhai@ict.ac.cn

Raster Content, MRC)图像模型<sup>[2-3]</sup>,将混合图像分解为前景层、遮罩层和背景层,再分别编码。基于块的算法则是将图像分成不重叠的大小相同的块,并基于块的特征进行分类编码。两者相比,前者的分类更精确。对于高质量要求的扫描文档图像压缩,商用软件 DjVu<sup>[4]</sup>、Digipaper<sup>[5]</sup>、Lura-Document<sup>[6]</sup>都属于这类方法。但也正是由于这种精确复杂的分类方法极大地降低了运算速率,因此在实时性较强的桌面图像压缩中,低复杂度的基于块的方法更加适合。

现有的计算机桌面图像压缩算法都是基于块进行的。AT&T 的虚拟网络计算(Virtual Network Computing, VNC)<sup>[7-8]</sup>采用了基于子矩形分类的无损压缩算法,但其应用于自然图像时,压缩效率过低。作为改进,文献[9]提出一种混合的桌面图像压缩算法,将  $8 \times 8$  的块分为文本/图形块和图像块,并分别采用 JPEG-LS 和动态的 JPEG 压缩。该算法对于纯文本或图像均有较好的压缩效果,但当两者混合时, JPEG 压缩在文本处引起了强烈的振铃效应。另一种分类更精细的混合压缩算法由文献[10]提出,其在图像块中进行了文本/图形元素的提取和分类编码,极大地提高了编码图像的质量。但当该算法应用于边缘渐变的文本如 PPT 图像时,仍会产生较强的振铃效应,而且对于实时传输系统来说,编码速度较慢,处理的帧数较低。

### 3 计算机桌面图像压缩算法

#### 3.1 总体过程

为能对计算机桌面图像进行实时有效压缩,本文根据桌面图像中文本/图形和图像各部分在色彩和纹理上的不同特性,提出新的压缩方法,其总体过程如图1所示。

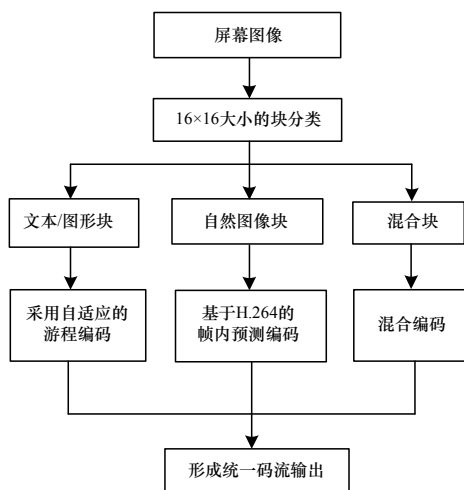


图1 计算机桌面图像压缩算法的总体过程

本文算法首先将计算机桌面图像分割成  $16 \times 16$  大小的非重叠块,然后根据每个块不同的色彩和纹理统计特征,将不同的块分为文本/图形块、混合块和图像块3类,其中,文本/图形块采用无损压缩编码方式,图像块采用类似于 H.264 帧内预测的有损编码方式,对混合块进行进

一步的分析,采用混合编码方式。

#### 3.2 块分类算法

对于待压缩的计算机桌面图像,先将其分割成  $16 \times 16$  大小,对于  $16 \times 16$  块分类采取的是基于直方图信息的,因为通常文本/图形块的色彩简单,纹理变化剧烈;而自然图像块的色彩丰富,纹理变化较平缓。所以基于直方图统计分析进行,算法简单高效。统计它们各自的直方图信息,主要有以下2点区别:

(1)文本/图形块的颜色种类数较少,而自然图像块的颜色种类数较丰富;

(2)文本/图形块的直方图分布为离散态,而自然图像块的直方图分布为连续态。

块分类的处理过程如图2所示。

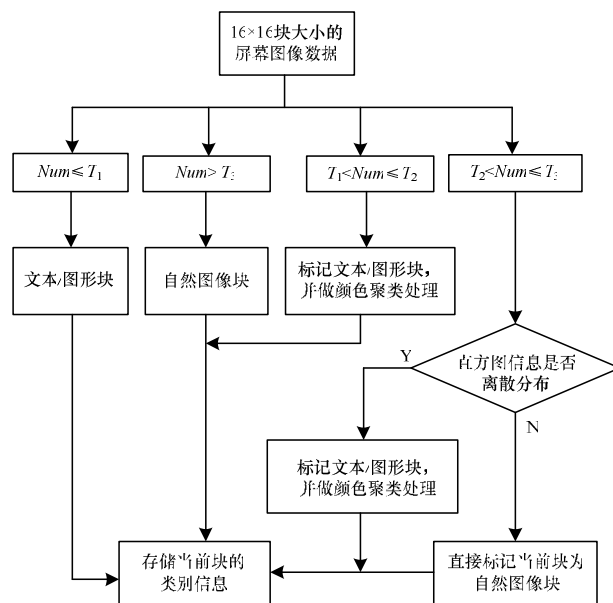


图2 块分类的处理过程

$T_1$ 、 $T_2$ 与 $T_3$ 分别为3个颜色种类数的阈值(本文中 $T_1$ 、 $T_2$ 、 $T_3$ 分别设为4、9、20),  $Num$ 为颜色种类数。对于简单的文本/图形块的颜色种类数不会超过 $T_1$ ,因此,将颜色种类数不超过 $T_1$ (含 $T_1$ )的块直接划分为文本/图形块;而即使有些复杂的文本/图形块的颜色种类数较多,但也不超过 $T_3$ ,因此,将颜色种类数大于 $T_3$ 的直接划分为自然图像块;对于颜色种类数大于 $T_1$ 且小于 $T_2$ 的,也划分为文本/图形块,但此种情况需要对此类块的颜色数据做聚类分析处理,使得经过聚类处理后的颜色种类数也降到 $T_1$ ,以便下一步对文本/图形块进行无损压缩编码;对于颜色种类数大于 $T_2$ 且小于 $T_3$ 的块,由于自然图像块的直方图分布是非离散的,因此,先对块直方图信息进行分析,若直方图分布是非离散的,则划分为自然图像块;若直方图分布是离散的,则划分为文本/图形块,并对其颜色数据做聚类分析处理。

#### 3.3 颜色聚类方法

聚类<sup>[11-12]</sup>方法目前已广泛的应用于机器学习与数据挖掘等计算机领域,而本文则首次将聚类的方法应用于计

计算机桌面图像压缩领域,使得对于计算机桌面图像的压缩得到了一定的改善。

对于颜色种类数超过  $T_1$  的文本/图形块,由于其颜色种类数已经较多,若直接对这类文本/图形块采用无损的自适应游程编码,则将较大的增加压缩的复杂度以及降低图像的压缩比。通过对这类文本/图形块进行颜色聚类处理之后,使得其颜色种类数降低到  $T_1$ ,这样既降低了无损压缩的复杂度,又提高了图像的压缩比。

对颜色的聚类采用的是非层次聚类算法中的经典算法——K-Means<sup>[13]</sup>算法。其主要思想为:随机地选择  $k$  个对象,每个对象初始地代表了一个簇的平均值或中心。对剩余的每个对象,根据其其与各个簇中心的距离,把它赋给最近的簇。然后重新计算每个簇的平均值。这个过程不断重复,直到准则函数收敛。K-means 算法的具体步骤如下:

**Step1** 从  $n$  个数据对象中随机选取  $k$  个对象作为初始聚类中心。

**Step2** 根据每个聚类对象的均值(中心对象),计算每个对象与这些中心对象的距离,并根据最小距离重新对相应对象进行划分。

**Step3** 重新计算每个(有变化)聚类的均值(中心对象)。

**Step4** 循环 Step2~Step3,直到每个聚类不再发生变化为止。

## 4 块压缩算法

### 4.1 文本/图形块编码

由于文本/图形块具有强烈的边缘与形状特征,人眼对文本/图形信息敏感,对其的解码恢复质量要求较高,因而无损压缩方法是最佳选择。经过颜色聚类处理后,其颜色种类数不超过  $T_1$ ,因此,本文提出自适应游程编码与索引表编码<sup>[12]</sup>相结合的无损压缩算法,如图3所示。

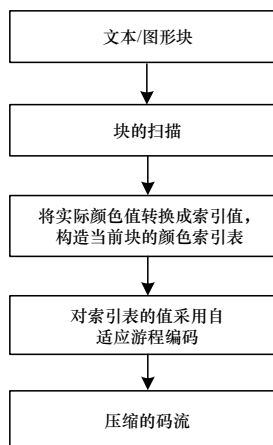


图3 文本/图形块的无损压缩算法流程

文本/图形块的无损压缩算法首先将块的实际颜色值转换成索引值并用相应的索引值对文本/图形块构造索引表,然后对索引表进行实际编码。这样可以较大地节省码流,从而提高压缩比。颜色值到索引表的转换过程如图4所示,其中,从左至右分别为块原始像素分布、颜色数据转换成索引数据和转换后块的索引数据分布。

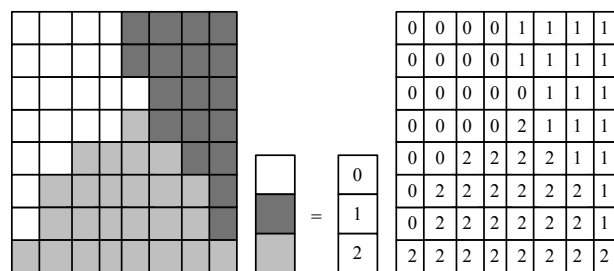


图4 颜色值到索引表的转换过程

对转换后的数据采用自适应游程编码算法对其进行压缩。其主要原理为,根据图像的统计特性预先设计一张变长码表,对每个码字进行编号,编号由小到大对应着码字的由短到长。编码过程中对游程分配一个相应的码字编号,根据游程的出现频率,动态地更新游程所对应的码字编号,使得出现频率高的游程对应较小的编号,即对应较短的码字。游程对应码字编号的更新过程如图5所示。其中,游程10原来对应的码字编号是8,先通过码字编号8搜索码表找到相应的码字,用此码字对游程10进行编码,然后更新游程10对应的码字编号,更新后游程10对应的码字编号为7,下一次再出现游程10时,就用码字编号为7的对应码字对其进行编码。

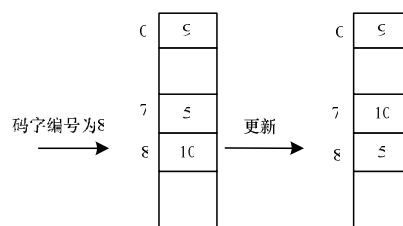


图5 游程对应码字编号的更新过程

### 4.2 自然图像块编码

自然图像的色彩信息丰富,纹理较平滑,人眼对自然图像的解码恢复质量相对较低,即使有部分失真人眼也不会有太大觉察,因而对自然图像块采用有损压缩算法可以实现有效压缩。

本文对图像块的编码采用的是类似于 H.264 的帧内预测编码算法,对亮度分量采用  $16 \times 16$ 、 $8 \times 8$  和  $4 \times 4$  这3种帧内编码模式,每种模式采用9种预测方向,色差分量采用  $8 \times 8$  帧内预测编码模式,每种模式4种预测方向,熵编码采用基于上下文自适应的可变长编码算法。

### 4.3 混合块的编码

混合块同时包含文本/图形信息以及自然图像的信息,因此,在前期的块分类中,首先将其分为自然图像块。在对自然图像块进行实际编码之前先检测该图像块是否包含有文本或图形信息,若检测到含有文本或图形信息,则该块属于混合块,并对此类块采用近似无损压缩方法进行编码。

H.264 的帧内预测编码算法通过量化步长来控制解码图像的失真度,因此,本文对混合块的编码采用减小量化步长来实现近似无损压缩,这就简单有效了实现了对混合块的压缩编码。

## 5 实验与结果分析

### 5.1 实验环境

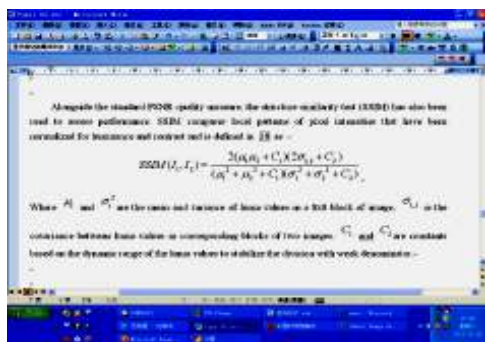
为检验本文算法对计算机桌面图像的压缩性能,实验选取了4张1024×768像素的计算机桌面图像进行测试,包括2张网页图像(Image1与Image2)、1张Word文档图像(Image3)以及1张PPT图像(Image4)。第1张网页图像包含大量文本内容;第2张网页图像同时包含文本内容与自然图像内容;第3张是1张Word文档截屏图像;第4张是一张PPT截屏图像。4张测试图像如图6所示。



(a)Image1



(b)Image2



(c)Image3



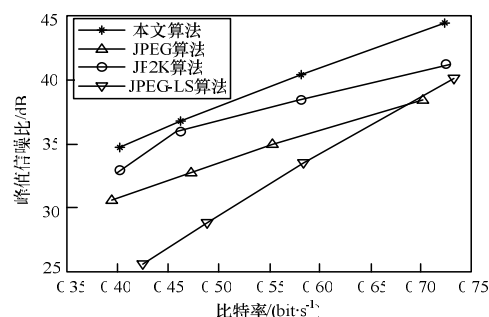
(d)Image4

图6 测试图像

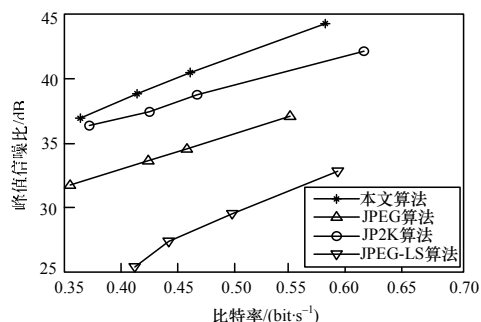
实验在 Inter Core2 3.0 GHz 个人计算机上进行,为验证本文算法在计算机桌面图像压缩方面的改进,选取另外3种编码方法与本文算法进行比较分析:(1)JPEG 有损压缩算法,采用 ILG'S JPEG 开源代码进行实验;(2)JPEG2000 有损压缩算法,采用 JasPer JPEG2000 开源代码进行实验;(3)JPEG-LS 近似无损压缩算法,采用 JPEG-LS 开源代码进行实验。

### 5.2 结果分析

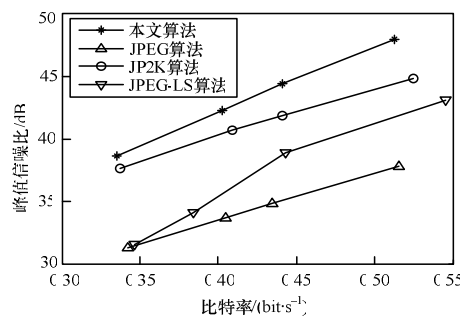
分别用本文算法、JPEG、JPEG2000、JPEG-LS 算法对4幅测试图像进行压缩编码,峰值信噪比(PSNR)与结构相似度(SSIM)<sup>[14-15]</sup>的比较结果分别如图7~图8所示。



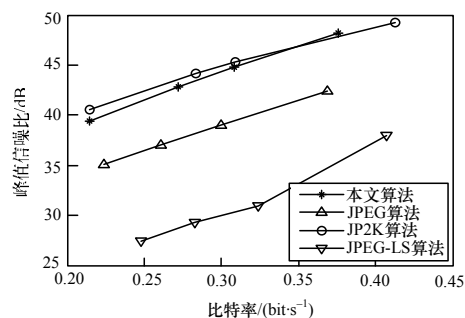
(a)Image1



(b)Image2



(c)Image3



(d)Image4

图7 不同算法的PSNR比较

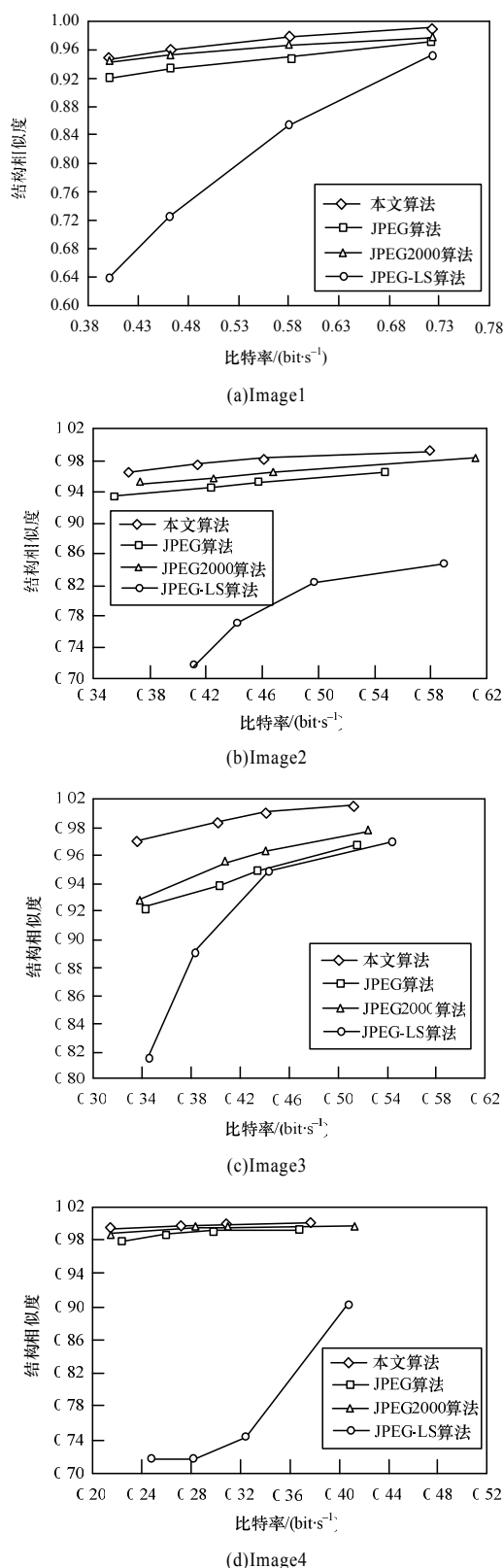


图8 不同算法的SSIM值比较

由图7可以看出,在各个比特率的情况下,本文算法对各种计算机桌面图像的压缩都要优于JPEG与JPEG-LIS压缩算法;在大多数情况下,本文算法压缩性能要优于JPEG2000算法,尤其对于包含大量文本/图形信息的网页类图像与Word文档类图像等。对于包含很多自然图像块与平滑块的Image4,本文算法的PSNR也与JPEG2000算

法非常接近。

由图8可以看出,本文算法的SSIM值比JPEG、JPEG2000以及JPEG-LIS算法的均要高。对于包含自然图像块与平滑块的桌面图像,如Image3, JPEG2000算法的PSNR值与本文算法的非常接近,但其SSIM值则要低于本文算法的SSIM值。

## 6 结束语

本文提出一种基于颜色聚类的计算机桌面图像压缩方法,该方法采用简单、低复杂度的块结构,首先计算机桌面图像分成 $16 \times 16$ 大小的非重叠图像块,并将块分为文本/图形、自然图像以及混合块3类,然后对不同的块分别应用不同的方法进行压缩。实验结果表明,本文提出的方法能够正确区分出3个块的类别,并能通过聚类方法有效减少复杂的文本/图形块的颜色个数以改善压缩效率,其压缩效率也要优于JPEG、JPEG2000以及JPEG-LIS算法。下一步将研究一种自适应的块分类方法,使得对于块的分类更加精确,并能实现不规则的块划分。

## 参考文献

- [1] Kochut A, Beaty K, Shaikh H. Desktop Workload Study with Implications for Desktop Cloud Resource Optimization[C]// Proceedings of IPDPSW'10. Atlanta, USA: IEEE Press, 2010.
- [2] ITU-T Study Group 16. TU-T T.44-2005 Mixed Raster Content (MRC)[S]. Telecommunication Standardization Sector of International Telecommunications Union, 2005.
- [3] Feng Guotong, Bouman C A. High-quality MRC Document Coding[J]. IEEE Transactions on Image Processing, 2006, 15(10): 3152-3169.
- [4] Haffner P, Bottou L, Howard P G, et al. DjVu: Analyzing and Compressing Scanned Documents for Internet Distribution[C]// Proceedings of ICDAR'99. Bangalore, India: [s. n.], 1999.
- [5] Huttenlocher D, Felzenszwalb P, Rucklidge W. DigiPaper: A Versatile Color Document Image Representation[C]// Proceedings of International Conference on Image Processing. Kobe, Japan: [s. n.], 1999.
- [6] Thierschmann M, Barthel K, Martin U. Scalable DSP Architecture for High-speed Color Document Compression[C]// Proceedings of Society of Photo Optical Instrumentation Engineers Conference. San Jose, USA: [s. n.], 2001.
- [7] AT&T Laboratories Cambridge. Virtual Network Computing (VNC)[EB/OL]. (2003-03-05). <http://www.cl.cam.ac.uk/research/dtg/attarchive/archive/vnc/>.
- [8] Tan Kheng-Joo, Gong Jiawei, Wu Bing-Tsung, et al. A Remote Thin Client System for Real Time Multimedia Streaming over VNC[C]// Proceedings of IEEE International Conference on Multimedia and Expo. Singapore: IEEE Press, 2010.
- [9] Said A. Compression of Compound Images and Video for Enabling Rich Media in Embedded Systems[C]// Proceedings of International Conference on Visual Communications and Image Processing. San Jose, USA: [s. n.], 2004. (下转第236页)