

可伸缩性视频编码的码率分配算法

王 琪¹, 赵 黎¹, 吴 枫², 李世鹏², 钟玉琢¹

(1. 清华大学计算机科学与技术系媒体研究所, 北京 100084; 2. 微软中国研究院, 北京 100080)

摘 要: 本文研究了渐进的精细的可伸缩性(Progressive Fine Granular Scalable, PFGS)视频编码的增强层码流的码率分配问题, 它是由 PFGS 各帧的增强层码流长度不同引起的. 我们首先建立了 PFGS 增强层的多帧联合的率失真(rate distortion, R-D)函数, 并通过 MPEG-4 标准测试序列检验了它的正确性. 然后利用联合率失真函数给出了局部最优的码率分配算法, 并进行了简化使得它的复杂性极低, 非常适合流化视频(streaming video)服务器使用. 实验结果显示本文提出的算法不仅使得各帧图像的解码质量比较平衡, 而且将 PFGS 的编码效率提高了将近 0.5dB.

关键词: 可伸缩性编码; 误差传播; 率失真函数; 码率分配

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2002)02-0205-05

A Rate Allocation Scheme for Progressive Fine Granular Scalable Video Coding

WANG Qi¹, ZHAO Li¹, WU Feng², LI Shripeng², ZHONG Yuzhuo¹

(1. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China;

2. Microsoft Research China, Beijing 100080, China)

Abstract: The enhancement layer rate allocation problem is examined in progressive fine granular scalable (PFGS) video coding. The problem arises from the fact that different frames in the enhancement layer have different rates in PFGS coding. A rate distortion (R-D) function for a multi frame group is first established for enhancement layer PFGS coding, followed by experiments to verify its validity using real test sequences. Optimal rate allocation among frames in the group is then given based on the R-D function, together with a simple implementation that is suitable for applications like streaming video. Experiments show that, compared with uniform bit allocation, optimal bit allocation not only makes the quality variation in decoded video much smoother, but also improves the average PSNR of PFGS coding by 0.3~0.5dB.

Key words: scalable coding; error propagation; rate distortion function; rate allocation

1 引言

视频编码的最初目标是为了在给定的码率下达到最优的解码质量. 然而由于近几年 Internet 上视频服务的增加, 使得视频编码的目标不仅是追求最佳的视频质量, 而是越来越注重视频编码的可伸缩性(scalable). 一方面基于 3 维小波变换的视频编码以其固有的多分辨率特性取得了一定的成功^[1~3]; 另一方面以 DCT 为基础的精细的可伸缩性(Fine Granular Scalable, FGS)视频编码^[4]已经被 MPEG-4 作为流式(streaming)应用的标准而采纳. 为了提高 FGS 的编码效率, 渐进的精细的可伸缩性(Progressive Fine Granular Scalable, PFGS)视频编码算法^[5](编码结构见图 1)在增强层图像编码时使用了前一帧重建的某个增强层图像为参考进行运动补偿, 使得运动补偿更有效, 从而提高了编码效率.

然而, 如果同 FGS 一样把增强层可用的比特数平均地分给各帧, PFGS 相邻帧之间的解码质量会有严重的波动, 整个视频序列的平均质量也会下降. 由图 1(a)所示, 第 2 帧使用

第 1 帧的质量较高的第 3 个增强层来作运动补偿, 所以增强层要编码的 DCT 残差较小, 于是产生的增强层码流较短; 而第 3 帧使用第 2 帧质量较差的第 2 个增强层来作运动补偿, 再加上为了避免误差传播第 2 帧生成的增强层参考图像是从第 1 帧的基本层恢复而来的, 所以第 3 帧的增强层要编码的 DCT 残差就比较大, 于是产生的增强层码流比较长. 于是使用均匀的码率分配时, 奇数帧的解码质量就比较低, 而偶数帧的解码质量比较高, 人的视觉系统对这种周期性的质量波动是非常敏感的. 另外, 当增强层解码可用的总比特数比较低时, 如果均匀地分配码率, 在图 1(a)所示的 PFGS 结构中, 第 4 帧可能已经解码到第 3 个增强层, 而第 3 帧可能才解码第 2 个增强层, 第 4 帧由于得不到高质量的参考, 它的解码质量会受到影响. 亦即如果我们能多分一些比特数给奇数帧, 使其也能够解码到第 3 个增强层, 不仅可以提高奇数帧的解码质量, 也能提高偶数帧的解码质量. 由图 1(b)也可以得到相应结论, 即应该为帧号不能被 3 整除的帧多分配一些比特, 而给帧号

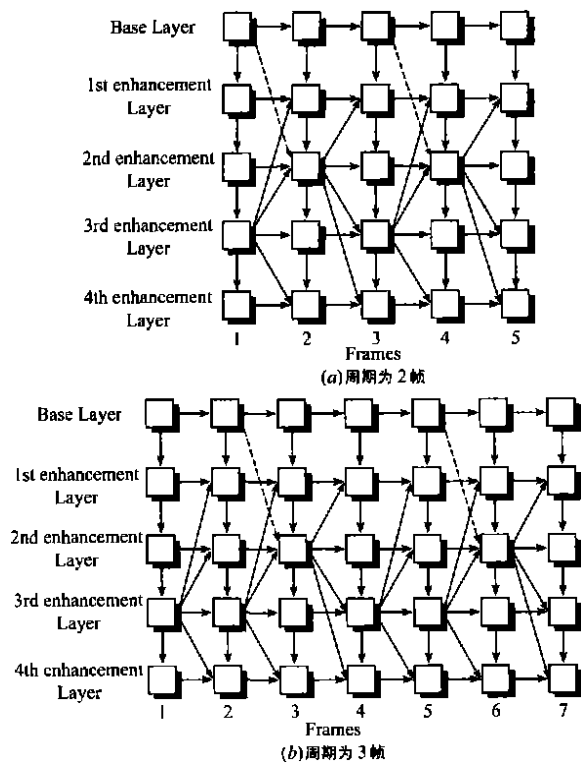


图 1 PFGS 编码框架

能被 3 整除的帧少分配一些比特。

综上所述,均匀的码率分配算法不仅会造成相邻帧之间的图像质量的波动,而且在码率较低时还会降低整个视频序列的解码质量。好在 PFGS 编码产生的增强层码流是完全可伸缩的,解码器可以从每帧的增强层码流的开始解码到任意位置,所以 PFGS 的增强层的码率分配不是在编码过程中完成,而是服务器在传输前根据信道状况来进行。因而一个简单地切分 PFGS 的增强层码流的方法对应用十分重要,它可以作为在视频流化中使用 PFGS 技术的一个工具。

2 问题的描述和简化

PFGS 的增强层码率分配算法不同于视频编码过程中的码率控制(rate control)算法,它是在码流被传输前根据信道的反馈信息来进行,也就是说它能独立于 PFGS 的编码过程。因此它有一些独特的要求:首先应该能提高整个视频的解码质量,这是根本目标;其次应该消除人眼十分敏感的相邻帧之间解码图像质量的抖动问题,尽可能地使其平滑;第三种方法必须十分简单,因为 PFGS 的编码技术将主要用于服务器和客户模式的应用,这样一个视频服务器会响应许多客户的请求,服务器需要根据用户的请求迅速分配码率,如果这个码率分配方法太复杂,那么服务器很难实时响应大量用户的请求。用 R_i 表示给第 i 帧分配的比特数,而用 D_i 表示第 i 帧的均方误差,那么码率分配其实质是一个求全局优化的问题,它可以用下面的式子来描述:

$$\min(D_1(R_1) + D_2(R_2) + \dots + D_N(R_N)) \quad (1)$$

满足条件 $R_1 + R_2 + \dots + R_N = S$ (常数)。这里 N 表示整个视频序

列的总帧数, S 为给定的总比特数。在求解上面的目标函数时,还需要满足一个附加条件就是相邻帧的均方误差不能出现很大的跳变。

实际上,要想求解目标函数(1)的全局最优解是十分困难的,首先因为整个序列的帧数 N 是一个比较大的值,求解全局最优解计算量非常大,而且需要大量的缓冲区和造成较大的延迟,这在实际中是不可接受的;其次每一帧总有一个增强层重构后作为后面一帧的参考图像,这种运动补偿形成了一种相邻帧间的依赖关系,使得全局优化的模型非常复杂;第三由于网络的带宽是随时变化的,也就是视频序列在传输过程中,不同的时刻可用的码率是变化的,因此服务器完全不必求解上面的目标函数的全局最优值,而只要对每一个瞬间假设网络近似恒定时求解局部最优即可。所以我们可以用局部优化的方法来求解 PFGS 的码率分配这个全局的问题,如把连续的几帧图像作为一个码率分配的基本单元,假设在这些帧传输时可用的码率是固定的,然后把这些可用的比特数分配给这些帧。注意到在图 1 所示的 PFGS 编码结构中,每帧的增强层参考图像的质量是周期变化的(a 和 b 分别对应了周期为 2 帧和 3 帧的情形),因此在码率分配过程中可以按若干帧为一个基本单位,将整个视频分成许多个组,然后将总码率平均分配到每一个组,再把每组可用的比特数优化到每帧图像。因此这个局部优化方案可以用下式表示:

$$\min(D_1(R_1) + D_2(R_2) + \dots + D_n(R_n)) \quad (2)$$

满足条件 $R_1 + R_2 + \dots + R_n = S$ (常数),其中 n 为组内的帧数,是每组可用的比特数。通过求解上面的极值问题,就可以确定在各组的各个帧中如何分配码率,从而完成整个序列的码率分配。

3 率失真函数的指数模型

码率控制是视频序列编码中的常用技术,它在编码过程中通过调整量化参数来控制生成码流的速率。码率控制的基础是率失真(rate distortion, R-D)函数^[6,7];同样我们的码率分配算法的基础也是 PFGS 的增强层的率失真函数,即增强层解码的比特数与解码的失真度之间的函数关系。对码率控制的研究表明,除了均匀分布之外,对信号的其它一般分布,率失真函数不存在准确的表达式;然而在中高码率下,即量化步长较小时,DCT 残差信号的率失真函数具有以下形式^[8]:

$$D(R) = \sigma^2 e^{-\alpha R} \quad (3)$$

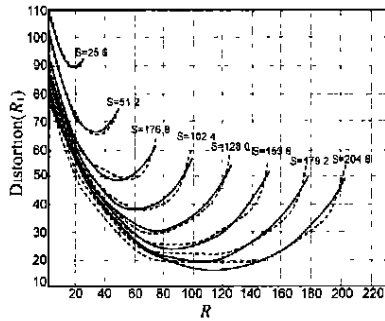
位平面编码是一个逐步量化(每次量化步长减半)的过程,它是一种特殊的标量量化;所以 PFGS 增强层的率失真函数也应该具有和式(3)相同的形式。这里是解码图像与原图像的误差的均方值, σ 是当增强层码率为 0 时解码图像(也就是基本层重构图像)与原图像的均方误差的平方根,它正比于增强层要编码图像信号(DCT 残差信号)的方差的平方根, R 是增强层解码的比特数, α 是一个常数。

由于 PFGS 的增强层码率分配是以若干帧图像为一个分配单位,因此我们要研究的不是单帧图像的增强层的率失真函数,而是多帧图像的增强层的联合率失真函数。将式(3)的率失真函数代入局部的码率分配目标函数,得到:

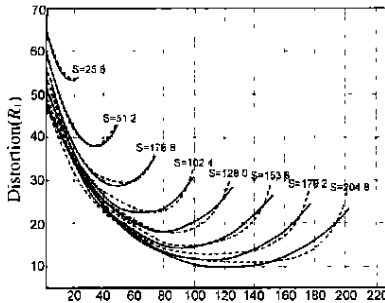
$$f(R_1, R_2, \dots, R_n) = \sigma_1^2 e^{-\alpha R_1} + \sigma_2^2 e^{-\alpha R_2} + \dots + \sigma_n^2 e^{-\alpha R_n} \quad (4)$$

这里 $f(R_1, R_2, \dots, R_n)$ 是组内各帧解码图像的失真度之和, σ_i 表示了第 i 帧基本层重构图像与原图像的均方误差的平方根, R_i 是解码第 i 帧的增强层所用的比特数.

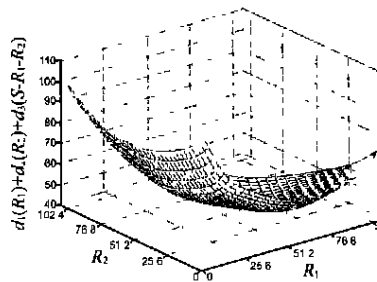
这样就得到了 PFGS 的多帧图像的增强层的联合率失真函数, 其中的各项之间并不是独立的, 因为 σ_i 是与 σ_{i-1} 有关 (由运动补偿关系产生) 的常数. 由于这个率失真函数是在一定的假设条件下得到的, 所以通过一个实验来验证上面的率失真函数在增强层的任何一个位置都成立. 实验中选取 PFGS 的基本层的码率是 128kb/s, 采样率为 10Hz, 使用 TM5 码率控制技术. 对图 1(a) 中的周期为 2 帧的 PFGS 编码结构, 用



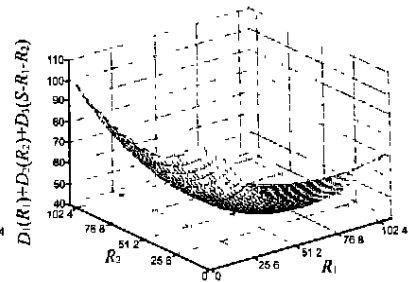
(a)Coastguard



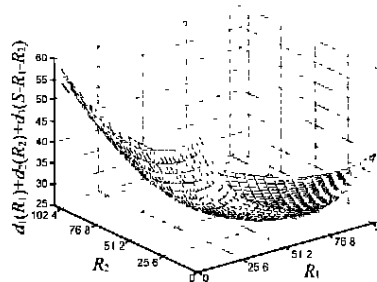
(b)Foreman



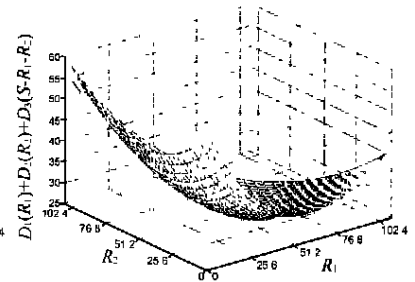
(a)Coastguard 的实际率失真曲面



(b)Coastguard 的指数形式的率失真曲面



(c)Foreman 的实际率失真曲面



(d)Foreman 的指数形式的率失真曲面

图 2 实际率失真函数曲线和指数函数曲线的比较

图 3 实际率失真函数曲面和指数函数曲面的比较

4 码率分配算法

有了 PFGS 的增强层联合率失真函数, 码率分配就比较简单了. 显然在码率分配算法中, 除了要求整个序列的解码质量要高以外, 同时要求相邻帧之间图像的解码质量不能出现大的跳变, 也就是说相邻帧图像的增强层最好是解码到相同位平面和相同的宏块位置. 如果给定的码率 S 能使组内各帧图像都解码无误差为 2^{k+1} (最高的增强层对应误差为 2^0 , 倒数第二个增强层对应误差为 2^1 , 依此类推) 的增强层, 但不能让组内各帧图像的增强层也都解码无误差为 2^k 的增强层, 如果码率分配算法已经知道了编码器在编码完每帧图像的每个位平面时所生成的比特数, 不妨假设第 i 帧编码无误差为 2^{k+1} 的增强层时所生成的比特数为 R'_i , 第 i 帧编码无误差为 2^k 的增强层时所生成的比特数为 R''_i , 分配给第 i 帧误差为 2^k 的增强层的比特数为 R_i , 那么增强层联合率失真函数式(4)变成

$$f(R_1, R_2, \dots, R_n) = \sigma_1^2 e^{-\alpha R_1} + \sigma_2^2 e^{-\alpha R_2} + \dots + \sigma_n^2 e^{-\alpha R_n} \quad (5)$$

满足 $R_1 + R_2 + \dots + R_n = S$, 其中 $\sigma_i^2 = \sigma_i^2 e^{-\alpha R'_i}$, σ_i 表示第 i 帧

MPEG 4 的标准测试序列 Coastguard 和 Foreman, 增强层的码率从 128kb/s, 256kb/s, ... 直到 1024kb/s, 即两帧图像的增强层可用的比特数 S 分别是 25.6kbits, 51.2kbits, ..., 直到 204.8kbits. 以 1.6Kbits 为间隔覆盖奇数帧码率的变化范围, 即奇数帧的比特数分别选择从 1.6kbits, 3.2kbits, ..., 直到 $S - 1.6$ kbits, 实际的率失真曲线(虚线)和指数的率失真曲线(实线)见图 2. 对图 1(b) 中周期为 3 帧的 PFGS 编码结构, 选取每个单元总码率为 115.2kbits, 实际率失真曲面和指数形式的率失真曲面的示意图见图 3. 可以看出用指数的率失真函数可以比较精确地描述了码率和失真之间的关系.

解码无误差为 2^{k+1} 的增强层时重构图像与原图像的均方误差的平方根.

这是一个条件极值问题, 使用 Lagrange 乘子法, 首先定义下面的函数

$$g(R_1, R_2, \dots, R_n) = f(R_1, R_2, \dots, R_n) - \lambda(S - (R'_1 + R'_2 + \dots + R'_n) - (R_1 + R_2 + \dots + R_n)) \quad (6)$$

对上式求导, 得

$$\frac{\partial g}{\partial R_1} = \frac{\partial g}{\partial R_2} = \dots = \frac{\partial g}{\partial R_n} = \frac{\partial g}{\partial \lambda} = 0 \quad (7)$$

化简得

$$R_i = \frac{S - \sum_j R'_j}{n} + \frac{2}{n\alpha} \sum_{j \neq i} \ln \frac{\sigma_j}{\sigma_i} \quad (8)$$

另一方面, 码率分配还要降低各帧之间的解码图像质量的波动, 这要满足:

$$\sigma_i^2 e^{-\alpha R_i} = \sigma_j^2 e^{-\alpha R_j} \quad (9)$$

不难验证式(8)满足式(9), 也就是说提高整个视频的解码质量和避免各帧之间解码质量的波动是一致的, 本文提出的算

法同时解决了这两个问题。

但是在码率分配过程中要计算 $\frac{1}{\alpha} \ln \frac{\sigma_k}{\sigma_l}$ 运算量很大, 而且传输中服务器不可能保存一份原始视频的拷贝, 所以必须找到简单的方法估算 $\frac{1}{\alpha} \ln \frac{\sigma_k}{\sigma_l}$. 由于 PFGS 的增强层是采用位平面技术编码的, 某一个增强层编码完后剩下的误差都是小于这个位平面所描述的值; 从统计的角度看, 可以认为每一个位平面编码完毕后, 解码图像的失真近似相等的, 那么当组内各帧图像解码完误差为 2^k 的增强层时, 下面的等式近似成立

$$\sigma_k^2 e^{-\alpha(R'_k - R''_k)} = \sigma_l^2 e^{-\alpha(R'_l - R''_l)} \quad (10)$$

化简得:

$$\frac{1}{\alpha} \ln \frac{\sigma_k}{\sigma_l} = \frac{(R'_k - R''_k) - (R'_l - R''_l)}{2} \quad (11)$$

如果能够得到 PFGS 增强层编码过程中每层编码所产生的比特数, 那么就可以用编码产生的比特数的有关信息 $\frac{(R'_k - R''_k) - (R'_l - R''_l)}{2}$ 来估计 $\frac{1}{\alpha} \ln \frac{\sigma_k}{\sigma_l}$, 于是得到下面的码率分配公式

$$R_i = \frac{S}{n} + \frac{\sum_{j \neq i} (R'_i - R'_j)}{n} \quad (12)$$

在这个算法中, 只需要知道编码器编码每一帧的每一个增强层使用了多少比特, 而这个信息可以在编码过程中很容易地生成。

5 实验结果

图 4 显示了序列“akiyo”的解码图像的 PSNR 随码率变化的曲线, 其中横坐标是码率, 单位是 kbits/s, 纵坐标是信噪比, 单位是 dB. 编码参数分别为: 视频序列格式是 CIF; 编码帧率为每秒 10 帧; 第 1 帧为 I 帧, 而其余帧均为 P 帧; 量化方法采用 H. 263 的量化; 基本层码流为 128 kbits/s. 增强层的码流分配采用两种方法, 第一种方法平均分配码率到每一帧 (虚线), 第二种方法用本文所提出的算法 (实线). 实验结果显示本文提出的算法能提高 PFGS

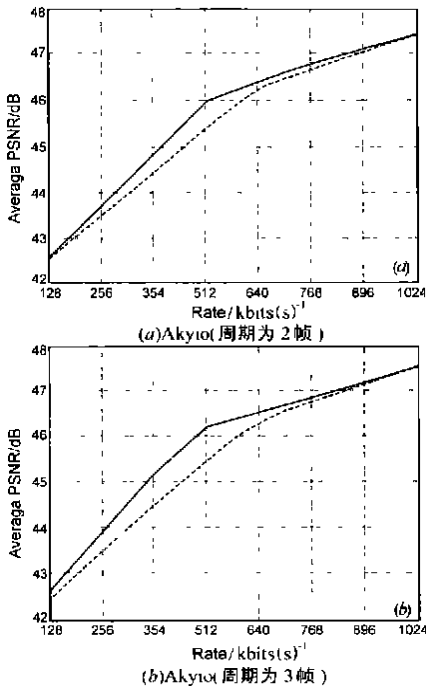


图 4 整个视频序列的 PSNR 随码率变化的曲线

的编码效率达 0.3~0.5dB. 这里要注意两点: 第一, 在我们的算法完全解码高质量参考图像的码率下, 算法和均匀分配相比达到最佳 (见实线曲线中的拐点); 第二, 在均匀分配也完全解码高质量参考图像的码率下, 算法和均匀分配相比非常接近 (见两条曲线的重合部分).

图 5 显示了 Akiyo 序列在 512kb/s 时解码图像的 PSNR 变化曲线. 其中横坐标是帧数, 纵坐标是信噪比, 单位是 dB. 可以明显的看出平均的码率分配方法 (虚线) 使得相邻帧图像的解码质量有很大的波动, 这种波动最大可以达到 1.5dB. 而采用我们提出的码率分配算法 (实线), 不仅提高了整个视频的质量, 而且减少了相邻帧之间的图像质量波动.

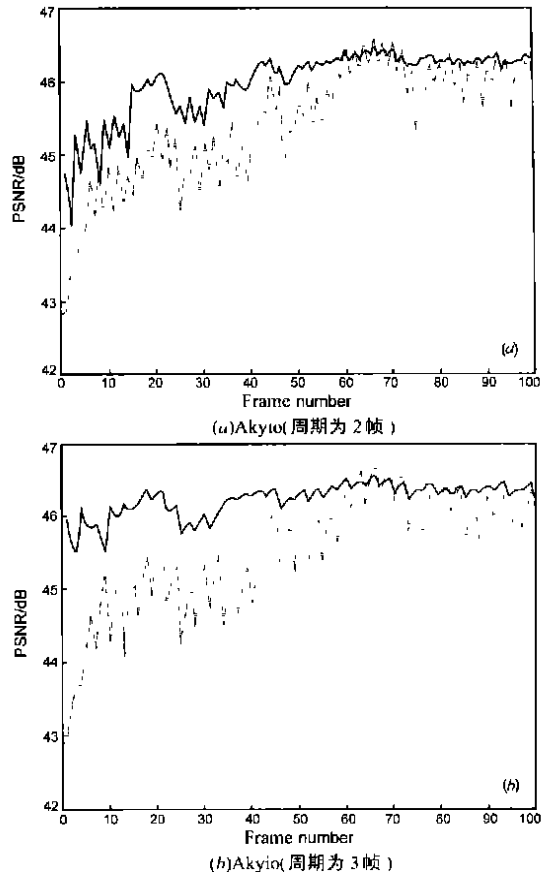


图 5 解码图像的 PSNR 随帧数的变化曲线

6 结论

本文根据 PFGS 编码方法的结构特点, 提出了一个改善 PFGS 整个视频的质量和减少相邻帧之间图像质量波动的码率分配算法. 首先建立了一个多帧图像增强层的联合率失真模型, 在该模型的基础上给出了一个简单而又有效的码率分配方法. 结果表明, 本文提出的算法一方面使整个视频的平均质量提高了 0.3dB~0.5dB, 另一方面使 PSNR 随帧数的变化曲线更加平滑, 提高了解码图像的主观质量.

参考文献:

[1] Choi S J, Woods J W. Motion compensated 3-D sub-band coding of

video [J]. IEEE Trans. on IP, 1999, 8: 155- 167.

- [2] Shen K, Delp E J. Wavelet base rate scalable video compression [J]. IEEE Trans. on CSVT, 1999, 9: 109- 122.
- [3] Kim B J, Xiong Z, Pearlman W A. Low bit rate embedded video coding with 3D set partitioning in hierarchical trees (3D SPIHT) [J]. IEEE Trans. on CSVT, 2000, 10: 1365- 1374.
- [4] Li W. Overview of fine granularity scalability in MPEG-4 video standard [J]. IEEE Trans. on CSVT, 2001, 11(3): 301- 317.
- [5] Wu F, Li S, Zhang Y-Q. A framework for efficient progressive fine granularity scalable video coding [J]. IEEE Trans. on CSVT, 2001, 11(3): 332- 344.
- [6] Chiang T, Zhang Y-Q. A new rate control scheme using quadratic rate distortion model [J]. IEEE Trans. on CSVT, 1997, 7(1): 246- 250.
- [7] Mallat S, Falzon F. Analysis of low bit rate image transform coding [J]. IEEE Trans. on SP, 1998, 46(4): 1027- 1042.
- [8] Hang H M, Chen F J. Source model for transform video coder and its application Part I: Fundamental theory [J]. IEEE Trans. on CSVT, 1997, 7(2): 287- 298.

作者简介:



王 琪 男. 原籍北京, 1973 年出生于河北省深县. 1997 年毕业于清华大学计算机系, 获工学学士学位. 现为清华大学计算机系媒体研究所博士研究生, 主要研究领域为图像和视频的编码及传输, 发表论文 10 余篇.



赵 黎 男. 1975 年出生于湖北省沙市. 1999 年毕业于清华大学电机系, 获工学学士学位. 现为清华大学计算机系媒体研究所博士研究生, 主要从事多媒体方向的研究, 包括视频分析与传输, 多媒体数据编码算法, 以及视频信息基于内容的检索. 发表文章 10 多篇.

眼科数字裂隙灯图像处理系统

(北京工业大学信号与信息处理研究室, 北京 100022)

北京工业大学信号与信息处理研究室与北京市同仁医院合作开发的眼科数字裂隙灯图像处理系统目前已在同仁医院投入使用, 并在临床应用中得到了专家及用户的好评. 其中部分关键技术已申请国家专利.

该系统由图像采集、图像处理、图像分析、信息数据库等子系统组成. 图像采集系统主要通过改造国产裂隙灯上的普通相机摄像光路, 设计光电接口, 利用高分辨率的数码相机拍摄眼科图像, 采集到的高清晰度图像通过 USB 接口传送到计算机; 由 PC 机完成图像处理、分析、存储、以及有关信息的管理; 处理分析的结果可以在监视器上输出, 也可以直接显示或打印, 还可以通过局域网传送进行远程会诊或教学. 采集到的图像要进行处理, 才能更好地为临床诊断服务. 裂隙灯图像处理系统应包括图像的处理、分析和管理模块, 各模块应完成的功能为:

1. 图像处理模块

主要包括图像亮度、对比度调整; 图像旋转、无级缩放、漫游等功能.

2. 图像分析模块

(1) 角膜类图像分析

角膜病灶分布和面积测量; 角膜血管分布和面积测量;

角膜浑浊度程度标定; 测量裂隙宽度, 评估角膜

厚度相对变.

(2) 结膜类图像分析

球结膜及睑结膜的充血程度分析; 球结膜及睑结膜病灶分析.

3. 信息数据库模块

存储中国人常见的典型外眼病图像(目前已有 2000 幅); 电子病历, 图像及相关信息的添加、修改、删除等功能; 检索病历及相关图像信息.

系统外观图如图所示:



图 1 基于裂隙灯的外眼图像采集与处理系统的外观