

通用率失真模型及码率控制算法*

谢正光, 包志华, 徐 晨, 章国安, 张士兵, 杨永杰

(南通大学 电子信息学院, 江苏 南通 226019)

摘要: 为了提高码率控制的准确性,通过实验和理论推导,提出了一种通用率失真模型。在此基础上,定义了纹理复杂度,提出了纹理复杂度和图像块头用字节两自适应线性预测模型及宏块(组)比特分配模型;然后结合以上模型提出了一种新的码率控制算法。大量仿真实验结果表明,利用新提出的码率控制方法,测试序列的 PSNR 均有不同程度的提高,且码流曲线更平滑平稳,目标码率控制更准确。

关键词: 通用率失真模型; 纹理复杂度模型; 自适应线性预测模型; 码率控制

中图分类号: TN919.81 文献标志码: A 文章编号: 1001-3695(2010)01-0339-03

doi:10.3969/j.issn.1001-3695.2010.01.101

General rate distortion model and rate control scheme

XIE Zheng-guang, BAO Zhi-hua, XU Chen, ZHANG Guo-an, ZHANG Shi-bing, YANG Yong-jie

(School of Electronics & Information, Nantong University, Nantong Jiangsu 226019, China)

Abstract: To improve the accuracy of rate control, proposed a new general rate distortion model on the basis of theoretical analysis and extensive experiments. Meanwhile, this paper defined texture complexity, presented optimistic bit allocation model and two adaptive linear predictive models; texture complexity and overhead bit count predictive models. Based on above models, proposed a new rate control scheme for low bit-rate and real-time video coding applications. Simulation results show that the new rate control algorithm provides smoother bit stream, higher video quality and more accurate rate control than previous algorithms.

Key words: general rate distortion model; texture complexity model; adaptive linear predictive model; rate control

TMS^[1]、TMN8^[2]、VM18^[3]以及基于 ρ 域率失真模型^[4]的码率控制算法都无法解决 H. 264/AVC 蛋鸡悖论问题,为此许多机构对此展开了研究。目前主要有两种比较主流的算法: a)中国科学院计算所的马思伟等人提出的类似 TMS^[1]基于 Buffer 控制的比特分配和量化参数控制算法^[5](本文称之为 JVT-TM5 或 JVT-D030); b)吸取了 MPEG-4 的 VM18 码率控制经验,采用了基于 Buffer 控制和二项式 RQ(rate quantization)模型的控制算法^[6](本文称之为 JVT-VM18 或 JVT-G012)。此外,中国科学院计算所的 W. Yuan 等人在 JVT-G012 算法的基础上,通过充分挖掘视频序列的时空相关性,提出了一种 JVT-VM18 的改进算法 JVT-O016^[7]。M. Jiang^[8]和 Y. Liu 等人^[9]根据前后帧的 MAD(mean absolute difference)或 PSNR(peak signal noise ratio)的变化,在文献[6]的基础上提出了各自的修正方案。但理论仿真实验表明:这些方法的算法复杂度较高且对某些视频序列并没有多大改善。

尽管 JVT-D030、JVT-G012、JVT-O016 以及 Y. Liu 及 M. Jiang 等人的修正方案对其进行了一些有益的探索且做出了一些开拓性的工作,但不同程度地存在以下问题: MAD 的预测差错将导致错误的量化步长而影响 RDO 模式选择的性能;基于 Laplacian 模型的 RQ 模型对有些视频图像并不完全适合;仅根据 MAD 而不考虑宏块头的宏块(组)级比特分配策略很难最大化视频质量;用已编码块的平均宏块头字节作为待编码块的宏块头字节的预测值,对 H. 264 来说,将影响码率控制的准确

性,导致编码性能下降。

1 通用率失真模型

现有码率控制策略均假设残差块的 DCT 系数基本不相关且服从方差为 σ^2 的 Laplacian 分布,编码后的码流为 $R(D)$ 。理论分析表明,DCT 系数分布模型的选择对码率控制非常重要,因为它决定着码率控制中比特分配的策略和量化步长的计算。许多学者和研究机构对视频残差信号的分布模型进行了研究,提出的模型主要有:高斯、通用高斯^[10]、混合高斯、拉普拉斯^[11]、通用拉普拉斯^[12]和柯西^[13]等模型。大量实验表明,TMN-8 和 VM-8 中采用的是基于拉普拉斯的简化模型,很难适合所有的视频信源。在大量实验的基础上,本文采用文献[14]的更通用的 RD(rate distortion)模型:

$$R(D) = (1/\alpha) \ln(\varepsilon^2 \sigma^2 / D) \quad (1)$$

其中: ε^2 是一个与分布模型有关的参数,如对拉普拉斯模型, $\varepsilon^2 = 1.2$,对高斯模型为 1.4; α 等于 $2 \ln 2$; $R(D)$ 纹理编码用比特; D 是失真测度,约为 $\rho \times Qstep^{2[15]}$ ($Qstep$ 表示量化步长),这里的 ρ 是与量化方式和 DCT(discrete cosine transform)系数分布模型有关的参数。

大量的仿真实验表明:高码流应用中,对于编码后的码流 $R(D)$ 与式(1)基本吻合;低码流应用中, $R(D)$ 与方差接近于线性关系;两条曲线在点 $\rho\varepsilon^2$ 处重合。所以可将式(1)在点

收稿日期: 2009-04-28; 修回日期: 2009-06-06 基金项目: 江苏省高校自然科学基金基础研究项目(08KJB510018);南通市应用研究项目(K2008028);南通大学博士启动基金资助项目(08R09);江苏省高校“青蓝工程”(2008)

作者简介: 谢正光(1967-),男,湖南洞口人,副教授,博士,主要研究方向为移动通信、无线多媒体、视频压缩、图像处理等(xie_zg@126.com);包志华(1955-),男,教授;徐晨(1960-),男,教授;章国安(1965-),男,教授;张士兵(1962-),男,教授;杨永杰(1969-),男,副教授,主要研究方向为信号处理、无线通信等。

$\rho e / \varepsilon^2$ 处进行泰勒展开:

$$R(Qstep) = R(D) = \frac{1}{\alpha} \ln \left[e \left(\frac{\varepsilon^2 \sigma^2}{\rho e \times Qstep^2} - 1 + 1 \right) \right] = \frac{1}{\alpha} + \frac{1}{\alpha} \ln \left[\left(\frac{\varepsilon^2 \sigma^2}{\rho e \times Qstep^2} - 1 \right) + 1 \right] = \frac{1}{\alpha} + \frac{1}{\alpha} \left[\left(\frac{\varepsilon^2 \sigma^2}{\rho e \times Qstep^2} - 1 \right) - \frac{1}{2} \left(\frac{\varepsilon^2 \sigma^2}{\rho e \times Qstep^2} - 1 \right)^2 + \dots + \frac{(-1)^{n-1}}{n} \left(\frac{\varepsilon^2 \sigma^2}{\rho e \times Qstep^2} - 1 \right)^n + \dots \right] \approx \frac{1}{\alpha} + \frac{1}{\alpha} \left(\frac{\varepsilon^2 \sigma^2}{\rho e \times Qstep^2} - 1 \right) = \gamma \frac{\sigma^2}{Qstep^2} \quad (2)$$

其中: $\gamma = \varepsilon^2 / (\alpha \rho e)$ 是与 DCT 系数分布模型及量化方案有关的系数,理论上它需在编码过程中随着特定帧的统计特性而自适应调整。所以可得到通用 RD 模型:

$$R(D) = \begin{cases} \frac{1}{\alpha} \ln \left(\frac{\varepsilon^2 \sigma^2}{D} \right), & \frac{\sigma^2}{Qstep^2} > \frac{1}{\alpha \gamma} \\ \gamma \frac{\sigma^2}{Qstep^2}, & \frac{\sigma^2}{Qstep^2} \leq \frac{1}{\alpha \gamma} \end{cases} \quad (3)$$

J. R. Corbera 等人^[15]通过大量仿真实验得到的并被成功应用在 TMN-8 码率控制中的 RD 模型公式只是上述公式的拉普拉斯特例,此时 $\gamma = e / \ln 2$ 。需要说明的是,对于低码率应用, $\sigma^2 / Qstep^2 < \rho e / \varepsilon^2$; 而高码率中,对应 PSNR > 40 dB 的情况^[15], $\sigma^2 / Qstep^2 > \rho e / \varepsilon^2$ 。本文后继部分主要考虑低码率应用。

2 新提出的纹理复杂度模型

2.1 纹理复杂度 TC 模型

2.1.1 纹理复杂度 TC 的定义

现有的码率控制,定义 MAD 或方差 σ^2 为图像块复杂度。本文定义的纹理复杂度为

$$TC = \gamma \times \sigma^2 \quad (4)$$

该式(称为纹理复杂度定义公式)定义的纹理复杂度不但与方差有关,而且还与 DCT 系数分布模型及量化方案有关。从编码角度看,纹理复杂度的高低意味编码后占用比特数的多少,所以本文定义的纹理复杂度更科学,物理意义也更明确。方差可由下式导出^[16]:

$$N^2 \sigma^2 = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \left[f(x,y) - \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \right]^2 = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f^2(x,y) - \frac{1}{N^2} \times \left[\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \right]^2 = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} F^2(u,v) - F^2(0,0) \quad (5)$$

理论上方差可在运动估计或 DCT 后可由上式导出,但对于估计匹配准则为 SAD(sum of absolute difference)而不是 SSD(sum of squared difference),或无需运动估计的转码应用中,需增加额外的计算。由式(3)得,TC 的计算也可在该宏块(组)编码完成后,通过下式估算:

$$TC = R(Qstep) \times Qstep^2 \quad (6)$$

上述公式称为 TC 的计算公式。

2.1.2 纹理复杂度 TC 的预测模型

由纹理复杂度 TC 的定义式(4)可得当前待编码图像块的复杂度 TC_{cb} 和参考帧中对应位置图像块复杂度为

$$TC_{cb} = \gamma_{cb} * \sigma_{cb}^2 \text{ and } TC_{pb} = \gamma_{pb} * \sigma_{pb}^2 \quad (7)$$

其中: γ_{cb} 和 σ_{cb}^2 分别为当前待编码块的 γ 和 σ^2 ; γ_{pb} 和 σ_{pb}^2 则分别是参考帧中对应位置图像块的 γ 和 σ^2 。因视频信源的时空强相关性, σ_{cb}^2 可用下面的线性自适应预测模型通过 σ_{pb}^2 估计:

$$\sigma_{cb}^2 = \xi_1 * \sigma_{cb}^2 + \xi_2 \quad (8)$$

其中 ξ_1 和 ξ_2 是该模型的两个参数。将式(8)代入式(7),可得基于 TC 定义公式的 TC 线性自适应预测模型:

$$TC_{cb} = \gamma_{cb} * \sigma_{cb}^2 = \gamma_{cb} * (\xi_1 * \sigma_{pb}^2 + \xi_2) = \xi_1 * \gamma_{cb} * \sigma_{pb}^2 + \xi_2 * \gamma_{cb} = \frac{\xi_1 * \gamma_{cb}}{\gamma_{pb}} * (\gamma_{pb} * \sigma_{pb}^2) + \xi_2 * \gamma_{cb} = c_1 * TC_{pb} + c_2 \quad (9)$$

其中: c_1 和 c_2 是该模型的两个参数,初值分别置为 1 和 0,且每编码完一个图像块后利用线性回归的方法自动更新。基于 TC 计算公式的 TC 线性自适应预测模型为

$$TC_{cb} = c_1 * TC_{pb} + c_2 = c_1 * [R(Qstep) * Qstep^2]_{pb} + c_2 \quad (10)$$

考虑视频信源的时空强相关性, TC_{cb} 的预测值被限制在一定的范围内以防止模型失效时产生奇异值。

$$TC_{cb} = \max((1 - 0.8 * \omega) * TC_{avg}, TC_{cb}), TC_{cb} = \min((1 + 0.8 * \omega) * TC_{avg}, TC_{cb}), TC_{cb} = \theta * TC_{cb} + (1 - \theta) * TC_{pb} \quad (11)$$

其中: TC_{avg} 为平均复杂度,它等于:

$$TC_{avg} = \left(\frac{u_{bw}}{F_r N_{bu}} - H_{avg} \right) \times Qstep_{avg}^2 \quad (12)$$

其中: u_{bw} 是预定义的信道带宽; F_r 表示帧率; N_{bu} 为每帧图像块数; θ 的值一般设为 0.5 ~ 0.8; H_{avg} 和 $Qstep_{avg}$ 分别代表参考帧中宏块头用平均字节和平均量化步长; ω 等于:

$$\omega = H_{avg} / (u_{bw} / F_r * N_{bu}), \omega = \max(\omega, h_ratio_1), \omega = \min(\omega, h_ratio_2) \quad (13)$$

h_ratio_1 和 h_ratio_2 分别设为 0.2 ~ 0.4 和 0.6 ~ 0.8。

2.2 头用字节预测模型

宏块(组)头用比特主要包括编码模式、运动矢量和量化步长等占用的二进制位数。因 H. 264 的可变块大小和众多的编码模式导致其相邻宏块(组)头用比特的变化较大,致使像 JVT-G012 中的宏块(组)头用比特预测模型: $H_{cb} = H_{avg}$ 的误差较大。综合考虑视频信源的时空强相关性,提出了下列宏块(组)头用比特线性自适应预测模型:

$$H_{cb} = h_1 H_{pb} + h_2 \quad (14)$$

其中: H_{cb} 为当前待编码宏块(组)的宏块头用比特; H_{pb} 是参考帧中对应位置宏块(组)的宏块头用比特; h_1 和 h_2 是该模型的两个系数,初值置为 1 和 0,且每编码完一个图像块后用线性回归法自适应更新。考虑到空域的相关性,预测后的宏块(组)头用比特再按下式进行修正:

$$H_{cb} = \tau H_{cb} + (1 - \tau) H_{avg} \quad (15)$$

其中: τ 一般为 0.5 ~ 0.8。

3 新提出的码流控制方案

码率控制是通过图像组层、图像层和宏块(组)层等三个层次来进行比特率控制的。GOP 层根据目标比特率和目标帧率为一个 GOP 分配所需码流;图像层根据该 GOP 层剩余的码率和缓冲区情况为该帧分配比特数;宏块(组)层根据该图像块的重要程度、纹理复杂度和该帧中分配给该块的码流情况来确定编码一个宏块(组)所用的量化参数。

3.1 基本单元级比特率控制

3.1.1 像块层比特率分配

像块层比特率控制中,分配给本图像块的比特数不仅与图像复杂度 TC 有关,还应该与像块头用字节有关。且像块头用字节与图像块复杂度间并没有存在近似的线性关系,所以像 JVT-G012 仅根据纹理复杂度来分配图像块比特数显然不合理,从而难以最大化视频质量。

最优化比特分配就是在给码流下,分配给该编码帧中不同图像块不同的比特数,以最大化视频质量。步骤如下:

a) 估计该编码帧中未编码图像块的复杂度和宏块(组)头

用字节。未编码图像块的复杂度的计算利用式(10)~(13),然后按下式求其和。

$$\sum TC = \sum_{r=\text{current}}^{\text{last}} TC_r \quad (16)$$

宏块(组)头用字节的计算用式(14)和(15),然后按下式求其和。

$$\sum H = \sum_{r=\text{current}}^{\text{last}} H_r \quad (17)$$

其中: TC_r 和 H_r 分别表示本编码帧中未编码图像块的纹理复杂度 TC 和头用字节; $\sum TC$ 是未编码图像块的纹理复杂度 TC 之和; $\sum H$ 为未编码图像块头用字节之和。

b)按下式估算待编码图像块的纹理用比特数 R_c 。

$$R_c = (T_r - \sum H) * TC_{cb} / \sum TC \quad (18)$$

其中: T_r 是本编码帧中未编码图像块可用的比特数,初值为按式(27)分配给本帧的预算比特数,每编码完一图像块后自动更新。为防止 R_c 过小,设置了两个下限值:

$$\begin{aligned} \text{if } R_c < \min i_bit/2 \text{ then } R_c &= \min i_bit/2 \\ \text{else } R_c &= \max \{R_c, \min i_bit\} \end{aligned} \quad (19)$$

其中: $\min i_bit$ 为

$$\min i_bit = (u_{bu}/F_r * N_{bu} - H_{avg})/2 \quad (20)$$

3.1.2 量化步长的计算

获得本帧待编码图像块的纹理复杂度估算值和纹理用字节预算比特 R_c 后,根据 TC 的计算式(6)可得本图像块的量化步长 $Q_{step_{cb}}$ 为

$$Q_{step_{cb}} = \sqrt{TC_{cb}/R_c} \quad (21)$$

然后,按照 H.264 的量化参数 QP 与量化步长的计算关系,将 $Q_{step_{cb}}$ 转换成 H.264 的量化参数 QP_{cb} 。为了保证时空相邻图像块的视频质量波动不至太大, QP_{cb} 需按 H.264 的规定再次修正。本图像块中所有的宏块将使用该修正后的 QP_{cb} 进行 RDO 模式选择和量化编码。

3.2 图像(组)层比特率控制

基于线性跟踪理论的流量模型被应用到 GOP 层和图像层的比特率分配上。设 N_{gop} 为每 GOP 的帧数, $n_{i,j}$ ($i=1,2,\dots,j=1,2,\dots,N_{gop}$)是第 i 个 GOP 的第 j 帧编号,编码完第 i 个 GOP 的第 j 帧后,缓冲区的充盈度 $B_c(n_{i,j})$ 可用下式计算:

$$\begin{aligned} B_c(n_{i,j+1}) &= \min \{ \max \{ 0, B_c(n_{i,j}) + A(n_{i,j}) - \frac{u(n_{i,j})}{F_r} \}, B_s \} \\ B_c(n_{1,1}) &= B_s/8 \quad B_c(n_{i+1,1}) = B_c(n_i, N_{gop+1}) \end{aligned} \quad (22)$$

$A(n_{i,j})$ 是第 $n_{i,j}$ 帧实际编码所花费的比特数; $u(n_{i,j})$ 表示实际可用带宽; B_s 为缓冲区初始容量设置值。

每 GOP 的第二个 P 帧的目标缓冲级别 $Tbl(n_{i,2})$ 定义为

$$Tbl(n_{i,2}) = B_c(n_{i,2}) \quad (23)$$

其中: $B_c(n_{i,2})$ 为编码完第 i 个 GOP 的第一个 P 帧后的缓冲区充盈度。其他 P 帧的目标缓冲级别 $Tbl(n_{i,j+1})$ 按下式计算:

$$Tbl(n_{i,j+1}) = Tbl(n_{i,j}) - \frac{Tbl(n_{i,2}) - B_s/8}{N_{gop} - 1} \quad (24)$$

按基于目标缓冲级别、缓冲区充盈度、帧率、可用带宽和帧级码流分配策略,分配给第 $n_{i,j}$ 帧的预算 $\hat{N}(n_{i,j})$ 比特数为

$$\hat{N}(n_{i,j}) = \frac{u(n_{i,j})}{F_r} + \eta * (Tbl(n_{i,j}) - B_c(n_{i,j})) \quad (25)$$

其中: η 是等于 0.5 的常量。按本 GOP 剩余比特数分配策略,分配给第 $n_{i,j}$ 帧的预算比特数 $\hat{f}(n_{i,j})$ 为

$$\hat{f}(n_{i,j}) = R_r/N_r \quad (26)$$

其中: R_r 为用来编码本 GOP 还未编码的 N_r 个 P 帧的比特数。最终分本给第 $n_{i,j}$ 帧的预算比特数为

$$T = \beta * \hat{f}(n_{i,j}) + (1 - \beta) * \hat{N}(n_{i,j}) \quad (27)$$

其中: β 常设为等于 0.5 的常数。得到本帧的预算码流后,再按基本单元级码率控制来完成本帧的预算码流并使图像质量最大化。

4 仿真实验

在 H.264 测试模型 JM98 Baseline 的基础上,对基于本文提出的基于新的 R-D 模型的新比特率控制方法进行了实验仿真。测试序列有 Moblie&Calendar、Cheers 和 Sailormen 格式为 CIF(common intermediate format)、D1 或 720P,帧率为 30 fps,测试 100 帧。此外,还对其他格式、帧率和标准测试序列进行了测试。表 1 列出了实验环境参数。测试时,先将各标准测试序列分别用固定量化参数 24、28、32、36 和 40 编码一次。其产生的码率作为码率控制方案的目标控制码率,对应固定的 QP 设为码率控制的初始 QP ,分别用 JVT-G012 和本文提出的码率控制算法进行控制。

表 1 测试环境

运动矢量精度	参考帧数	Hadarnard 变换
1/4 像素	1	开
熵编码	基本单元宏块数	
CABAC	11(CIF)/45(D1)/80(720P)	
GOP 结构	搜索范围	RDO
IPPP	± 16	开

图 1~6 是本文方法与 JVT-G012 的部分码流曲线及 PSNR 曲线的比较。对于低码率实时 CBR(constant bit rate)视频应用,码流曲线要求越平稳越好。因本算法采用了最优化比特分配策略,从而保证了一定码率约束条件下的图像质量最大化。从图 1~6 的 PSNR 曲线可以看出,所有测试序列的 PSNR 值均有不同程度的提高,码流曲线比 JVT-G012 更平稳,对目标码率的控制更准确。

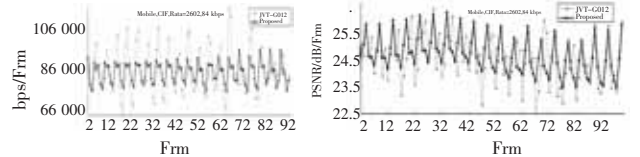


图 1 Moblie&Calendar 的码流曲线比较

图 2 Moblie&Calendar 的 PSNR 曲线比较

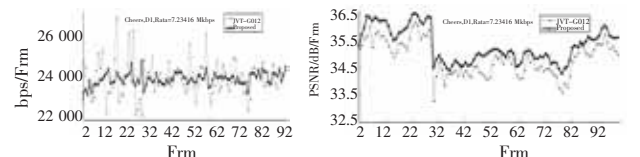


图 3 Cheers 的码流曲线比较

图 4 Cheers 的 PSNR 曲线比较

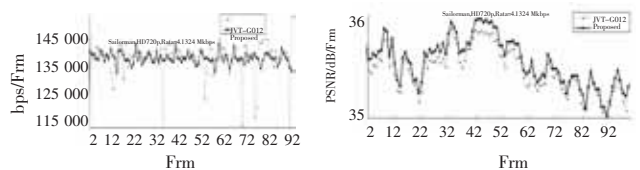


图 5 Sailorman 的码流曲线比较

图 6 Sailorman 的 PSNR 曲线比较

测试结果表明,本算法能够提供更稳健的码率控制方案,相同带宽下能获得更高图像质量,更平滑平稳的码流曲线,更准确的目标码率控制。这些优越性得益于本码率控制策略的更通用的 RD 模型,更准确的复杂度定义,更优化的比特分配策略和两个线性自适应预测模型等。

5 结束语

在对现有码率控制算法的优缺点和 H.264 编码的特点进行系统研究后,提出了一种更通用的工程率失真模型;然后根据此率失真模型,定义了纹理复杂度,提出了 (下转第 344 页)

用新算法的图像分割结果更自然,更多的颅骨细节被保持;图 2 中新算法的分割结果与原始图像更接近;在图 3 中,当使用新算法时,远处景物细节被保留,而用传统的均值漂移算法不能清晰地分割远处的景物细节;图 4 使用新算法时,帽子和头发被清楚地分割显示。

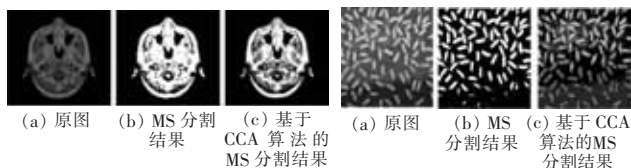


图 1 Mri 图实验结果对比

图 2 Mri 图实验结果对比



图 3 Cameraman 图实验结果对比

图 4 Lena 图实验结果对比

4 结束语

带宽是均值漂移算法的重要参数,它不但决定了参与迭代的采样点数量,而且还会影响算法的收敛速度和准确性,因此带宽的选择非常重要。本文利用相关性比较算法计算带宽,然后再用均值漂移算法进行图像分割,避免了经典的均值漂移算法人工确定带宽的主观性。实验结果表明:基于相关性比较算法的均值漂移图像分割算法的分割结果更接近于原始图像,更自然,能够保留更多的图像细节。但是基于相关性比较算法的

均值漂移图像分割算法运行时间较长,如何加快该算法的运行速度是需要进一步研究的内容。

参考文献:

- [1] FUKUNAGA K, HOSTETLER L D. The estimation of the gradient of a density function with applications in pattern recognition [J]. *IEEE Trans on Information Theory*, 1975, 21(1): 32-40.
- [2] CHENG Y Z. Mean shift, mode seeking, and clustering [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1995, 17(8): 790-799.
- [3] COMANICIU D, MEER P. Mean shift: a robust approach toward feature analysis [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 603-619.
- [4] COMANICIU D, RAMESH V, MEER P. The variable bandwidth mean shift and data-driven scale selection [C]//Proc of IEEE International Conference on Computer Vision. 2001: 438-445.
- [5] 王兆虎, 刘芳, 焦李成. 一种基于视觉特性的遥感图像分割 [J]. *计算机学报*, 2005, 28(10): 1687-1691.
- [6] COMANICIU D. An algorithm for data-driven bandwidth selection [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2003, 25(2): 281-288.
- [7] COMANICIU D, RAMESH V, MEER P. Real-time tracking of non-rigid objects using mean shift [C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2000: 142-149.
- [8] WU K L, YANG M S. Mean shift-based clustering [J]. *Pattern Recognition*, 2007, 40(11): 3035-3052.
- [9] YANG M S, WU K L. A similarity-based robust clustering method [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2004, 26(4): 434-448.
- [10] PSNR-based frame complexity estimation [J]. *IEEE Trans on Consumer Electronics*, 2005, 5(1): 281-286.
- [9] LIU Y, LI Z G, SOH Y C. A novel rate control scheme for low delay video communication of H. 264/AVC standard [J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2007, 17(1): 68-78.
- [10] MULLER F. Distribution shape of two-dimensional DCT coefficients of natural images [J]. *Electronics Letters*, 1993, 29(22): 1935-1936.
- [11] REINIGEK R C, GIBSON J D. Distributions of the two-dimensional DCT coefficients for images [J]. *IEEE Trans on Communications*, 1983, 31(6): 835-839.
- [12] ZHOU F, LU Z, LING H. Statistical model of quantized DCT coefficients [C]//Proc of Presented at International Conference on Signal Processing (ICSP). 2004: 2572-2575.
- [13] ALTUNBASAK Y, KAMACI N. An analysis of the DCT coefficient distribution with the H. 264 video coder [C]//Proc of Presented at IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP). 2004: 177-180.
- [14] HANG H M, CHEN J J. Source model for transform video coder and its application-Part I: fundamental theory [J]. *IEEE Trans on Circuits and Systems for Video Technology*, 1997, 7(2): 287-298.
- [15] JORDI R C, LEI S. Rate control in DCT video coding for low-delay communications [J]. *IEEE Trans on Circuits and Systems for Video Technology*, 1999, 9(1): 172-185.
- [16] XIE Z, LIU Y, LIU J, et al. A general method for detecting all-zero blocks prior To DCT and quantization [J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2007, 17(2): 237-241.
- [17] THOMAS W, MICHAEL L, CAMPBELL T G, et al. A rate-constrained encoding strategy for H. 263 Video Compression [R]. 1995.

(上接第 341 页)纹理复杂度和图像块头用字节两自适应线性预测模型及宏块(组)比特分配模型。在以上模型的基础上提出了一种新的码率控制算法。大量仿真实验结果表明,利用新提出的码率控制方法,所有测试序列的 PSNR 均有不同程度的提高,其码流曲线更平滑平稳,目标码率控制更准确。这些优越性得益于本码率控制方案更通用的率失真模型,更准确的复杂度定义,更优化的比特分配策略和两个线性自适应预测模型等。

参考文献:

- [1] Video test model, version 5 (TM5), ISO/IEC JTC1/SC29/WG11 [R]. 1993.
- [2] Video codec test model, version 8 (TMN8), ITU-T/SG15 [R]. 1997.
- [3] Video verification model, version 18 (VM18), ISO/IEC JTC1/SC29/WG11 [R]. 2001.
- [4] HE Z, KIM Y K, MITRA S K. Low-delay rate control for DCT video coding via p-domain source modeling [J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2001, 11(8): 928-940.
- [5] MA S, GAO W, LU Y. Rate-distortion analysis for H. 264/AVC video coding and its application to rate control [J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2005, 15(12): 1533-1544.
- [6] LI Z, PAN F, LIM K P, et al. Adaptive basic unit layer rate control for JVT, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-G012 Document [R]. 2003.
- [7] WU Y, LIN S, ZHANG Y, et al. Optimum bit allocation and rate control for H. 264/AVC [J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2006, 16(6): 705-715.
- [8] JIANG M, LING N. On enhancing H. 264/AVC video rate control by