

高分辨率视频会议中基于模板的自适应编码算法

王菀栋 叶德建

(复旦大学软件学院网络多媒体实验室, 上海 201203)

摘要 如今,对高分辨率的便携计算机视频会议系统的需求日益高涨。虽然便携计算机使用方便、局域网带宽充足,能够为工作学习带来许多益处,然而计算资源的紧张却直接影响了高分辨率视频会议的效果。针对这个问题提出了自适应模板编码算法(template-based coding algorithm, TBCA),不仅能比传统自适应算法提高10%的计算资源利用率,还减少了近一倍的参数调整频率。TBCA已被实现在校园网会议系统(campus video meeting system, CVMS)中,成功地提高了用户体验。

关键词 自适应编码 实时视频系统 高分辨率 视频编码

中图分类号: TP301.6 **文献标识码**: A **文章编号**: 1006-8961(2007)10-1811-04

Applying Template-based Adaptive Video Coding Algorithm on High Resolution Live Video System

WANG Chun-dong, YE De-jian

(Networking and Multimedia Laboratory, Software School, Fudan University, Shanghai 201203)

Abstract There is a great demand to support high resolution local video conference in today's PCs. However, although PC is convenient and local access network bandwidth is sufficient, the computational capacity is a big bottleneck in high resolution laptop video conference system. Here, Template-based Coding Algorithm(TBCA) is proposed for such a situation which can improve the utilization of computing resources by 10% and reduce the adjustment frequency to half of the initial situation. A high resolution laptop video conference system based on the TBCA is implemented with satisfactory user experiences.

Keywords adaptive coding, live video system, high resolution, video coding

1 引言

视频捕获设备技术成熟成本下降,使得高分辨率协同工作时机成熟。据预测便携计算机的高分辨率视频系统(high resolution laptop video conference, HRLVC)具有巨大潜力,本文针对HRLVC系统中凸现的计算资源不足问题展开研究。当今常见的PC视频服务如MSN视频、G-Talk等仅能提供低分辨率视频,而HRLVC系统的目标则是在带宽充足的前提下提供高分辨率视频服务。

然而,虽然PC愈加便携易用而带宽愈加充足,计算资源不足却成为突出的问题。通常视频会议进

行 1280×1024 像素的实时编码将消耗至少90%的计算资源。不仅如此,视频会议系统作为辅助软件若独占计算资源,不仅影响其他软件,视频会议系统也会受此影响而产生诸如视频抖动、停顿等问题。

因此保持友好分享计算资源是视频会议系统的关键目标之一。而达成目标的关键便是自适应编码。在传统自适应编码过程中,运动检测和估计^[1-3]消耗计算资源最多,像素搜索范围便是其相关参数。而图像质量^[4-7]也是常见控制计算复杂度的参数。

本文提出了一种新的普适的自适应算法来解决CPU竞争的问题。与传统算法^[6]相比,它能更友好而有效地分享CPU运算资源,并能兼容各类编码器。它已经被应用在校园专用视频会议系统

收稿日期:2007-07-05;改回日期:2007-07-25

第一作者简介:王菀栋(1982-),男。现为复旦大学软件学院网络多媒体实验室计算机软件与理论专业硕士研究生。研究方向为流媒体传输、对等网络。E-mail: chundongwang@fudan.edu.cn

(campus video meeting system, CVMS)中了。

2 算法

2.1 当前自适应实时视频编码

视频编码算法的自适应性是指视频编码算法动态调整编码过程的特性,自适应编码算法能最大化资源利用率并适用于各种环境。

一般而言,最常见的自适应方法采用逐步调整参数来逼近最优。逐步逼近法在传统的自适应视频编码中被普遍使用,是因为编码器计算复杂度与参数之间没有稳定而普适的映射关系。固定步长自适应算法(fixed stepping adaptive algorithm, FSAA)及动态步长自适应算法(dynamic stepping adaptive algorithm, DSAA)是两种主要的逐步逼近算法。FSAA 进行线性的参数调整;DSAA 通过使用二次方程进行参数调整,有效地缩短了逐步逼近的过程。

2.2 可用的调整方法

对于视频编码算法而言,与算法复杂度攸成相关的可用调整方法主要有以下两个。

(1) 重采样原始视频图像 自适应控制器能通过抛弃原始图像信息来降低视频编码的复杂度,其中最有效的方式便是重采样。各类参数中尤以颜色、尺寸及帧率是重采样过程中效果显著的 3 个参数。

(2) 更改编码器参数 控制器还能通过改变编码器参数来调整运算资源的耗费。因调整编码器参数不涉及双方协商而代价较小。图像质量以及移动检测范围则是其中最为有影响力的两个参数。

虽然图像重采样占用极少计算资源,却需重新配置编码器并协商,从而引入延时。因此,本文将重采样参数称为“重量级参数”。相对地,将编码器内部参数称为“轻量级参数”。

2.3 基于模板的自适应编码算法

虽然逐步逼近的方式似乎是唯一可用的自适应方法,然而响应速度欠缺是其致命弱点。因实时视频会议本身对实时性的严格要求,逐步逼近法不仅响应速度难以满足要求,而且反馈的滞后性也将降低准确性并造成逼近过程的性能损失。

本文提出自适应模板编码算法(template based codec algorithm, TBCA),通过动态调整的参数模板,在编码过程中自动适应本地硬件环境以及应用场景,直接计算并应用适当的参数。由于模板调整速度也将影响收敛过程的时长,这里采用 DSAA 快速

收敛至最优组合。

输入:

$C_{free}(t)$: CPU 空闲值,代表剩余计算能力的标量。

算法过程:

(1) 使用默认值初始化模板序列。

(2) 寻找编码器参数 G_i ,使得 $C_{free}(t+1)$ 能够降到最低。

(3) 如果存在,那么将该 G_i 应用到编码器。

(4) 否则,使用 DSAA 来逐步逼近最佳参数。

(5) 如果 $C_{free}(t-1) < \delta$,将上轮的 G_i 对应的模板用 $C_{free}(t-1) < \delta$ 更新,带遗忘效应。

(6) 回到第 1 步直到停止。

输出:

$\{G_i\}$: 参数组。

图 1 中的“反馈控制器”就是应用了 TBCA 的控制器模块,负责控制编码器参数。它以 CPU 占用率作为反馈值,从经过本次调整的模板集中选取最佳模板。目前,CPU 占用率是关于有效的系统计算资源的主要指标,因为它能够独立于系统环境的变化而普遍适用。

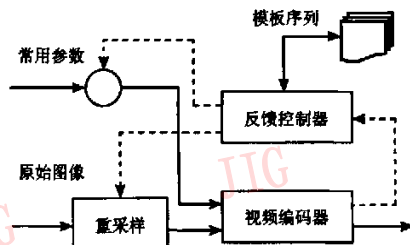


图 1 自适应视频编码器结构

Fig. 1 Structure of adaptive video codec

图中的模板序列则维护了从 CPU 利用率到编码器参数的动态映射关系。序列中每个模板代表着一组编码器参数 G_i 与一个相应的利用率值,而所有模板则按照编码器参数递增排列。

3 数学分析

定义 1 可能的参数向量集合定义如下:

$$P = \{p_i | p_i = (p_i^0, p_i^1, \dots, p_i^n)^T, 10^i \leq p_i^i \leq 10^{i+2}\}$$

函数 $C(p_i)$ 从参数向量映射到资源消耗标量 α ,

$$C: p_i \rightarrow \alpha$$

显然地, $C(p_i)$ 是一个 $\|p_i\|$ 上的单调函数。

定义 2 步进函数是指通过逐步调整参数从而逼近解的函数,即

$$g: (r, p_0) \rightarrow (p, (r \in S, p_0, p, \in P))$$

$$g(r, p_0) = \begin{cases} g_u(p_0) & r + C(p_0) < \bar{r} - E \\ p_0 & \bar{r} > r + C(p_0) > \bar{r} - E \\ g_d(p_0) & \text{其他} \end{cases}$$

其中, $g_u(p_0)$ 及 $g_d(p_0)$ 表示单步向上调整函数及单步向下调整函数。 S 表示可用资源值的集合。 E 是逼近目标的阈值, $\bar{r}$ 是最大可用资源。

优选函数 $h(r)$ 是基于有效 CPU 资源从记录的模板中选出合适的参数向量的映射函数。定义如下:

$$h: r \rightarrow p (r \in S, p \in P)$$

$$\forall r \in S, \forall \varepsilon > E, \exists p \in P, p = h(r)$$

$$|C(p) - r| < \varepsilon$$

结论 TBCA 所获得的参数向量解的平均利用率高于逼近算法。

TBCA 以 E 为长度将 S 划分为 $E/\bar{r}$ 部分。由于初始阶段持续时间很短,分析整体性能时可忽略它。

根据定义 2, 因 TBCA 具有近似计算算法特性,应用基于模板算法后时刻 t 的资源利用率。 $U_t(t)$ 能将参数解的平均误差控制在 $E/2$ 。

对于逼近算法而言,它们的平均误差为

$$\frac{E}{\bar{r}} \cdot \frac{E}{2} + \frac{\bar{r} - E}{\bar{r}} \cdot (g(r, p_0) - p_0)$$

这里,

$$g(r, p_0) - p_0 = \max(g_u(p_0) - p_0, g_d(p_0) - p_0)$$

由于反馈系统的调整步数需满足以下不等式,

$$s_{\text{step}} \ll \bar{r}$$

而且,

$$\bar{r} > \begin{cases} \frac{s_{\text{step}}}{2} & \text{FSAA} \\ \frac{k_{\text{max}} \cdot s_{\text{step}}}{2} & \text{DSAA} \end{cases}$$

这里,

$$k_{\text{max}} \leq \left\lfloor \frac{\sqrt{1 + i \cdot 8\bar{r}/s_{\text{step}}} - 1}{2} \right\rfloor$$

这使得,

$$\frac{E}{\bar{r}} \cdot \frac{E}{2} + \frac{\bar{r} - E}{\bar{r}} \cdot (g(r, p_0) - p_0) > \frac{E}{2}$$

因此有应用基于步进算法后时刻 t 的资源利用率 $U_g(t)$ 的平均误差将大于 $U_t(t)$, 所以逼近算法 FSAA 和 DSAA 得到的参数向量解的平均利用率低于 TBCA。

4 性能分析

4.1 实验模型

与实时视频会议同步运行的背景应用程序的处理器请求在不同场景下虽然千差万别,但服务请求到达总体上仍满足 Poisson 分布。实验环境如表 1 所示。

表 1 TBCA 实验环境配置

Tab. 1 Experimental environment setup

	最小	最大
视频编码	H. 264	
视频分辨率	192 × 144	1 280 × 1 024
帧率	5	30
编码搜索范围	0	100
输出图像质量	0	100

在实验中,所有程序的计算与用户操作都被抽象为对 CPU 资源的请求。请求密度 (Density of Request, DoR) 是用来表示 CPU 请求到达密度的指标。它代表了与视频编码器共享计算资源的应用程序的计算复杂程度。

4.2 性能指标

实验中所使用的性能指标如下:

(1) CPU 利用率 CPU 利用率被定义为所有应用程序,包括视频编码器和当时的其他程序,所需要的所有 CPU 计算资源占总 CPU 计算资源的比例。CPU 利用率越高,意味着资源浪费越少,分配给编码器的资源越多,视频质量和用户体验也就越好。

(2) 自适应调整间隔 调整频率被定义为调整编码器参数 (包括原始图像重采样参数) 单位时间内的期望调整次数。调整频率越小,编码器重新配置越少,其所带来的短暂黑屏几率越低,同时跨越网络协商几率也越低,从而用户体验越好。

4.3 实验结果

图 2 与图 3 展示了 3 种算法的性能比较,包括 CPU 利用率以及参数调整频率。从图 2 中可以看出, TBCA 的平均 CPU 利用率高于 DSAA 及 FSAA。FSAA 的平均利用率为 87.63%, DSAA 为 89.16%, 而 TBCA 则拥有 95.34% 的平均利用率,相比逼近算法而言,拥有了 7% 的性能提升。TBCA 在初始阶段之后,利用模板快速获得适当参数,因此系统整体 CPU 利用率也就高了许多,因此随着 DoR 的提升,

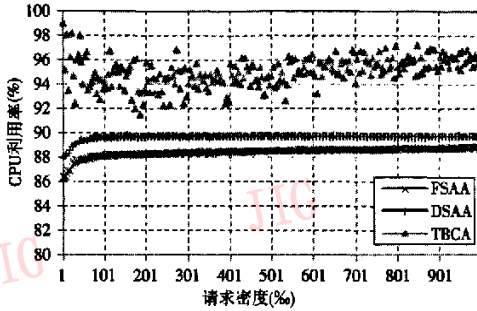


图 2 CPU 利用率, 这里 DoR 在 0% 与 100% 之间

Fig. 2 CPU utilization

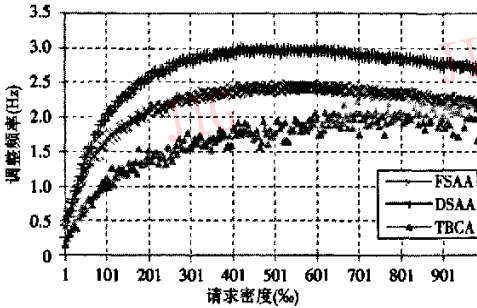


图 3 调整频率, 这里 DoR 在 0% 与 100% 之间

Fig. 3 Adjustment frequency

可快速计算请求愈多, 利用率优势也愈加明显。

TBCA 的性能优势不仅仅体现在高 CPU 利用率上, 还可从更少的调整次数中看出其在用户体验方面的优势。图 3 展示了这方面 TBCA 所获得的令人印象深刻的性能增益。当参数调整频率超过 2Hz 时, 将导致用户体验大幅下降。图 3 中, DSAA、FSAA 都在请求密度分别超过 100%、180% 时超过了 2Hz, 而 TBCA 则在请求密度超过 600% 后, 才逐步维持在 2Hz 附近。TBCA 一般仅需两步以内的调整即可, 因此在请求密度低于 800% 前均保持了较大优势。

4.4 系统实现

TBCA 算法已经被应用 CVMS 中了。CVMS 是一个基于 P2P 的高分辨率视频会议系统, 它专为校园网络环境设计, 能够提供多方参与的高分辨率实时会议服务, 具体应用环境参照表 2。高分辨率的视频服务以及自适应算法使复旦大学与 UMASS 大学的师生能利用其进行日常交流和定期会议, 而完全取代了传统的视频服务例如 MSN 视频、G-Talk 等。

CVMS 受到如此热烈的欢迎不仅因为良好的视频质量, 还因为它能够友好的与其他应用程序分享计算资源。因此不会因为使用它而影响了其他应用

表 2 CVMS 应用环境配置

Tab. 2 Environment setup of CVMS

	最小	最大
CPU	Pentium III 1.8GHz	Core Duo 3.4GHz
每会话用户数	3	7
视频分辨率	192 × 144	800 × 600

程序如开发环境或是文档处理软件等的执行。

5 结 论

基于反馈的自适应算法能够最大化资源利用率, 由于实时性的限制, TBCA 通过类直接计算的方法几乎一次得到最适参数, 大大加快逼近速度, 减少性能损失, 并兼顾不同的环境和场景。目前仅考虑了 CPU 资源, 扩大监控资源将是下一目标。

不仅如此, TBCA 这种通过记录历史来将资源占用映射至对应参数的算法, 不但能快速而准确地得到最适解, 亦能自适应于各种环境。因此对于受限资源环境对实时性和灵活性有严格要求的应用都很适合, 将它推广到其他的领域也是下一个目标。

参考文献 (References)

- 1 Wang Hui, Mao Zhi-gang. An adaptive motion estimation algorithm based on evolution strategies [A]. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing [C], Montreal, Quebec, Canada, 2004: 353 ~ 356.
- 2 Sorwar G, Murshed M, Dooley L. Fully Adaptive Performance Scalable Block-based Motion Estimation [A]. In: Proceedings of IEEE International Conference on Information, Communications and Signal Processing [C], Bangkok, Thailand, 2005: 1145 ~ 1149.
- 3 Sun H, Vetro A. Adaptive Video Coding Method [P]. United States Patent 5,969,764, 1999-10-19.
- 4 Chen W H, Smith C H. Adaptive coding of monochrome and color images [J]. IEEE Transactions on Communications, 1977, COM-25: 1285 ~ 1292.
- 5 Donald A, Lee M C. Scene-adaptive transform domain video partitioning [J]. IEEE Transactions on Multimedia, 2004, 6: 58 ~ 69.
- 6 Tang Ge, Zhu Li. An algorithm of adaptive post-processing to reduce the block artifacts in the block-based transform image coding techniques [J]. Geometrics and Information Science of Wuhan University, 2002, 27(1): 81 ~ 84. [汤戈, 朱立. 一种减少块效应的自适应后处理算法 [J]. 武汉大学学报信息科学版, 2002, 27(1): 81 ~ 84.]
- 7 Yin Hao, Zhang Jiang-shan. A robust adaptive video encoder based on Human visual model [J]. Journal of Electronics (China), 2003, 20(2): 142 ~ 149.
- 8 Gupta S, Williamson C L. A performance study of adaptive video coding algorithms for high speed networks [A]. In: Proceedings of IEEE International Conference on Local Computer Networks [C], Minneapolis, Minnesota, USA, 1995: 317 ~ 325.