

多优先级感兴趣区 H.264 计算资源分配

陈晓雷^{1,2}, 张爱华^{1,2}, 陈莉莉^{1,2}, 林冬梅^{1,2}, 杨欣翥³

(1. 兰州理工大学电气与信息工程学院, 兰州 730050; 2. 甘肃省工业过程先进控制重点实验室, 兰州 730050;
3. 兰州交通大学机电工程学院, 兰州 730070)

摘 要: 移动设备有限的电池能量是实时手语视频通信所面临的关键制约因素。为此, 提出多优先级感兴趣区的 H.264 计算资源分配方法。从原始手语视频中快速检测出面部、双手和背景区域, 依据视觉重要性的不同, 给面部和双手区域分配较多的计算资源, 给背景区域分配较少的计算资源。实验结果表明, 该方法在保证手语视频感兴趣区编码质量的同时, 可以大幅降低编码器的计算复杂度, 减少电池能量消耗, 延长能量受限设备的工作时间。

关键词: 人类视觉系统; H.264 标准; 感兴趣区; 计算资源分配; 手语视频; 电池能量

H.264 Computation Resource Allocation in Multiple Priority Region of Interest

CHEN Xiao-lei^{1,2}, ZHANG Ai-hua^{1,2}, CHEN Li-li^{1,2}, LIN Dong-mei^{1,2}, YANG Xin-zhu³

(1. School of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;
2. Key Laboratory of Gansu Advanced Control for Industrial Process, Lanzhou 730050, China;
3. School of Mechatronic Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

【Abstract】 The limitation of energy in battery powered devices is very challenging to apply H.264 to design a mobile sign language communication system. This paper therefore proposes a multiple priority Region of Interest(ROI) H.264 computation resource allocation scheme. In the system, the face, hands and background of sign language video is first detected with a fast algorithm. Then several coding parameters which allow the encoder to allocate computational resources differently are adaptively adjusted at the Macroblock(MB) level according to the relative priority of each MB. As a result, it keeps fine quality in face and hands, greatly reduces the computational complexity and prolongs the battery life.

【Key words】 human visual system; H.264 standard; Region of Interest(ROI); computation resource allocation; sign language video; battery power

DOI: 10.3969/j.issn.1000-3428.2013.04.065

1 概述

H.264 是现在广泛使用的最先进的视频编码标准之一, 和以前的标准相比, 在相同视觉质量下 H.264 能够节省大约 50% 的比特资源, 所以该标准非常适合带宽受限的移动应用。但是, H.264 编码算法的计算复杂度大幅增加, 在能量受限的移动设备上使用原始 H.264 编码器进行实时视频通话非常消耗计算资源及电池能量。

为此, 本文提出一种面向能量受限移动设备多优先级感兴趣区的 H.264 计算资源分配方法, 利用快速手语视频感兴趣区检测方法从手语视频中检测出面部、手部和背景区域, 给这些区域分配不同的 H.264 编码参数以保证编码

器给感兴趣区分配更多的计算资源, 给背景区分配较少的计算资源, 从而在保证手语视频感兴趣区编码质量的同时减少总计算资源消耗, 延长电池工作时间。

2 手语视频通话的制约因素

手语是聋哑人进行交流的一种视觉语言。过去, 聋哑人主要利用文字短信方式进行移动通信, 但是, 文字交流的速度受限于打字能力, 并且研究表明文字交流速度要比手语交流速度慢很多, 文字每分钟只能交流 5 个~60 个词, 手语每分钟可以交流 120 个~240 个词^[1]。另外, 聋哑人更习惯用手语交流, 而不是文字信息交流, 所以对于聋哑人而言文字交流方式并不是有效的移动通信方式。因此, 如

基金项目: 教育部博士点基金资助项目(1203ZBB0079); 甘肃省自然科学基金资助项目(1112RJZA033)

作者简介: 陈晓雷(1979—), 男, 讲师、博士研究生, 主研方向: 图像和视频处理; 张爱华(通讯作者), 教授、博士; 陈莉莉, 硕士研究生; 林冬梅, 助教、硕士; 杨欣翥, 讲师、硕士

收稿日期: 2012-04-23 **修回日期:** 2012-06-25 **E-mail:** zhangah@lut.cn

何提高聋哑人之间的移动通信效率, 成为一个亟待解决的问题。随着移动通信技术的快速发展, 具备视频摄取及播放功能移动设备的日益普及和新一代视频编码标准 H.264 的诞生, 研究人员提出了移动手语视频通信技术^[2]。这种技术在不增加硬件模块的基础上可以用移动设备实时采集、编码和传输手语视频信息, 极大方便了聋哑人之间的交流。

但是实时的手语视频通话目前存在以下关键制约因素: (1)有限的移动网络带宽。(2)有限的移动设备电池能量。

为了克服第(1)个因素, 研究人员将感兴趣区比特资源优化方法引入到手语视频编码中。此类方法的目标是在保证手语视频可理解性的同时最大化压缩手语视频以满足带宽要求。目前, 用于手语视频编码的感兴趣区比特资源优化方法已取得了一些研究成果。文献[3]开发了一种基于肤色的手语视频分割方法, 该方法能够识别并分割出手语视频中对话者的面部和双手, 这些分割出的区域被视为感兴趣区并在编码时给予优先编码从而保证它们的视觉质量。文献[4]提出了一种视觉凹特性模型, 该模型可以给手语视频面部区域分配较小的量化系数, 给其他区域分配较大的量化系数, 从而可以保证面部区域的视觉质量。文献[5]利用基于动态肤色的感兴趣编码方法给手语视频中的肤色区域分配更多的比特资源以保证它们的编码质量。文献[6]利用重建块的运动矢量和帧间预测模式, 并结合人类视觉感知特征, 提出一种视频感兴趣区域快速提取算法及其优化编码方案。

但是针对第(2)个制约因素的研究工作却较少。文献[7]提出了可变帧率技术, 该技术可以检测出聋哑人是否在打手语。当聋哑人打手语时, 将视频编码帧率提高; 当聋哑人自己不打手语而只是在观看对方所打手语时, 将视频编码帧率降低。可变帧率技术可以增加 47% 的手机电池能量, 利用这些能量可以进行 68 min 的手语视频通话。本文则从移动手语视频通信技术中所使用的 H.264 视频编码标准入手减少移动设备电池能量消耗。

3 快速手语视频感兴趣区检测方法

精确的手语视频感兴趣区检测方法, 需要较高的计算要求, 在能量受限的移动设备上运行该类方法会过多地消耗视频编码所需的计算资源。为此, 本文提出了如图 1 所示的快速手语视频感兴趣区检测方法。

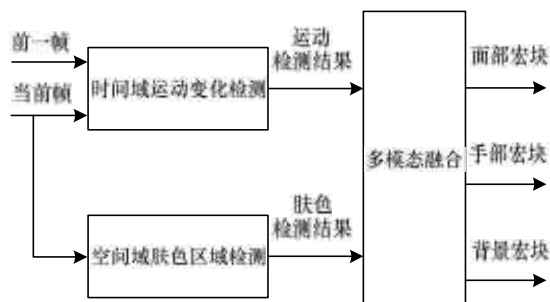


图 1 快速手语视频感兴趣区检测方法

3.1 时间域运动变化检测

设 $F(x, y, k)$ 为当前输入视频帧, $F(x, y, k-1)$ 为前一视频帧, 定义式(1)所示帧差:

$$DF_{k,k-1}(x, y) = \begin{cases} 1 & |F(x, y, k) - F(x, y, k-1)| > Thr_k \\ 0 & \text{else} \end{cases} \quad (1)$$

$DF_{k,k-1}(x, y)$ 中所有值为 1 的像素被视为运动像素, 所有值为 0 的像素被视为静止像素, 式中 Thr_k 是由式(2)定义的阈值^[8]:

$$Thr_k = \frac{1}{M \cdot N} \sum_{x=1}^M \sum_{y=1}^N |F(x, y, k) - F(x, y, k-1)| \quad (2)$$

其中, M 和 N 表示视频帧的行数和列数。

3.2 空间域肤色区域检测

肤色是人体表面最为显著的特征之一, 是人体面部及手部区别于其他部分的重要特征。

研究表明当视频的颜色模式是 YUV 时, 肤色区域像素的色度分量值 U 和 V 均属于一个较小的数值范围且对于不同的光照条件也较鲁棒^[9]。

利用这一特性, 对视频帧中的每一个像素进行肤色区域检测, 当像素的 U 分量值介于 77~127 之间同时 V 分量值介于 133~173 之间时, 认为该像素属于肤色区域。定义一个二进制参数 p 来指明一个像素是否属于肤色区域。 p 为 1 表明该像素属于肤色区域, p 为 0 表明该像素不属于肤色区域。

$$p = \begin{cases} 1 & \text{if } U \in [77, 127] \& V \in [133, 173] \\ 0 & \text{else} \end{cases} \quad (3)$$

3.3 时空域检测结果融合

将以上像素级的运动变化检测结果和肤色区域检测结果按式(4)融合产生宏块级的感兴趣区检测掩模, 将手语视频帧分为面部、手部和背景宏块。

$$I_{ROI}[i] = \sum_{j=1}^{16} \sum_{k=1}^{16} [DF_{k,k-1}(x, y) \cdot p(i, x, y)] \quad (4)$$

其中, i 是宏块索引; x, y 是某一个宏块中的像素索引。

4 多优先级感兴趣区的 H.264 计算资源分配

标准 H.264 编码器不区分每个宏块对聋哑人视觉系统的相对重要性, 对所有宏块采用相同的参数进行编码。而事实上, 研究表明聋哑人在观看手语视频时最关注人的面部, 其次是运动的双手, 最后是背景^[10]。

在背景区域对应的宏块上消耗过多的计算资源并不会改善手语视频的总可理解性。因此, 本文方法根据面部、双手和背景宏块的相对重要性自适应地为这 3 种宏块分配包括预测模式、搜索范围、参考帧数、像素精度和量化系数在内的编码参数, 从而给面部宏块分配最多的计算资源以保证其具有最好的编码质量, 给双手宏块分配较多的计算资源以保证其具有较好的编码质量, 同时给背景宏块分配最少的计算资源以减少编码器所消耗的总计算资源, 具体方法如图 2 所示。

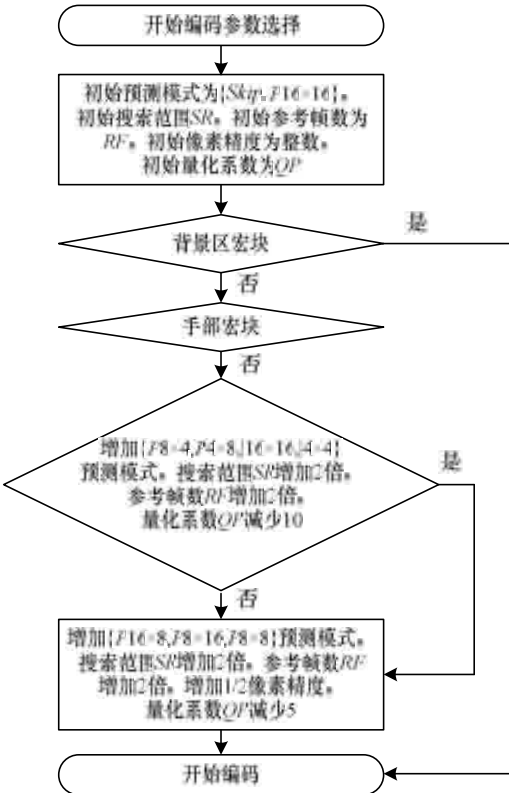


图 2 多优先级感兴趣区的 H.264 计算资源分配流程

5 仿真结果及分析

本文在 H.264 测试代码 JM16.2 基础上加入多优先级感兴趣区 H.264 计算资源分配方法后, 从总运动估计时间、亮度值峰值信噪比(PSNR_Y)和电池工作时间 3 个方面和原始测试代码 JM16.2 进行比较, 检验算法的有效性。

原始 JM16.2 编码器的参数配置如下: baseline Profile, 开启 RDO 和 CAVLC, 搜索方法为 UMHS Simple, 搜索范围为 8, 参考帧数为 5, 预测模式为 {Skip, P16×16, P16×8, P8×16, P8×8, P8×4, I16×16}。

采用本文方法时手语视频各区域的预测模式、搜索范围、参考帧数和像素精度根据图 2 进行分配。选取 Silent 和 Irene 2 个 QCIF 分辨率(176×144 像素)的手语视频序列进行实验, 这些序列各取 100 帧, 第 1 帧为 I 帧(帧内编码), 其余 99 帧全部为 P 帧(帧间编码)。实验平台是由锂电池供电的笔记本电脑, 该电脑配置如下: Intel Core2 Duo CPU, 主频 2.13 GHz, 内存 2 GB DDR, 锂电池容量 4 400 mAh。

表 1 给出了本文方法和 JM16.2 的客观性能比较, 其中, QP=35。

表 1 原始 JM16.2 和本文方法的性能比较

视频序列	编码方法	总运动 估计时间/s	PSNR_Y/dB			
			面部	双手	背景	平均
Silent	JM	197.3	33.5	33.4	33.7	33.6
	本文方法	92.2	37.7	35.1	33.3	33.2
Irene	JM	195.9	35.3	35.2	35.7	35.5
	本文方法	91.5	39.3	37.0	34.2	35.4

从表中实验结果可见: 和 JM16.2 相比, 本文方法可以大幅减少运动估计时间。由于本文方法给面部和双手宏块分配了更多的计算资源, 因此面部、双手的 PSNR 要高于 JM16.2 面部、双手的 PSNR。同时由于本文方法给背景区宏块分配了较少的计算资源, 因此背景区的 PSNR 要低于 JM16.2 背景区的 PSNR。此外, 本文方法的平均 PSNR 也略低于 JM16.2 的平均 PSNR。

总体而言, 本文方法以牺牲背景区编码质量为代价, 在保证手语视频感兴趣区编码质量的同时, 大幅减少了计算复杂度。

为了验证本文方法对电池能量消耗的影响, 利用笔记本电脑电池监视工具检测不同编码方法下的电池工作时间(指从电池充满电到电池电量耗尽所花的时间)。

从表 2 实验结果可见, 对于 Silent 视频序列, 本文方法可以延长电池工作时间 18 min, 对于 Irene 视频序列, 本文方法可以延长电池工作时间 22 min。

表 2 不同编码方法下的电池工作时间

视频序列	编码方法	电池工作时间/min
Silent	JM	108
	本文方法	126
Irene	JM	123
	本文方法	145

图 3 给出了本文方法和 JM16.2 的主观视觉质量比较。从实验结果可见, 本文方法以牺牲背景区信息为代价, 保留了更多的面部和双手信息, 更有利于聋哑人识别手语视频中的面部表情和手部动作, 提高了手语视频的可理解性。



(a)原始未编码视频 (b)JM16.2 (c)本文方法
Silent 第 218 帧

图 3 不同编码方法下的主观视觉质量

6 结束语

针对能量受限移动设备上的 H.264 手语视频编码, 本文提出了多优先级感兴趣区 H.264 计算资源分配方法。该方法依据聋哑人视觉系统选择特性进行计算资源分配, 在保证手语视频面部和双手编码质量的同时, 可以大幅降低编码器的计算复杂度, 减少电池能量消耗, 延长能量受限移动设备的工作时间。

由于能量受限移动设备的电池能量是不断变化的, 因此下一步工作将研究如何把能量因素引入手语视频编码系统, 以进一步减少电池能量消耗。

参考文献

- [1] James C L, Reischel K M. Text Input for Mobile Devices: Comparing Model Prediction to Actual Performance[C]//Proc. of Conference on Human Factors in Computing Systems. New York, USA: ACM Press, 2001: 365-371.
- [2] Cavender A, Ladner R, Riskin E. MobileASL: Intelligibility of Sign Language Video as Constrained by Mobile Phone Technology[C]//Proc. of the 8th International ACM SIGACCESS Conference on Computers and Accessibility. New York, USA: ACM Press, 2006: 71-78.
- [3] Saxe D M, Foulds R A. Robust Region of Interest Coding for Improved Sign Language Telecommunication[J]. IEEE Trans. on Information Technology in Biomedicine, 2002, 6(4): 310-316.
- [4] Agrafiotis D, Canagarajah N, Bull D R, et al. A Perceptually Optimised Video Coding System for Sign Language Communication at Low Bit Rates[J]. Signal Processing: Image Communication, 2006, 21(7): 531-549.
- [5] Chon J, Cherniavsky N, Riskin E, et al. Enabling Access Through Real-time Sign Language Communication over Cell Phones[C]//Proc. of the 43rd Annual Asilomar Conference on Signals, Systems, and Computers. Piscataway, USA: IEEE Press, 2009: 588-592.
- [6] 刘鹏宇, 贾克斌. 基于重建编码信息的视频 ROI 提取算法[J]. 计算机工程, 2011, 37(24): 278-280.
- [7] Cherniavsky N, Cavender A C, Ladner R E, et al. Variable Frame Rate for Low Power Mobile Sign Language Communication[C]//Proc. of the 9th International ACM SIGACCESS Conference on Computers and Accessibility. New York, USA: ACM Press, 2007: 163-170.
- [8] Sun Yu, Ahmad I, Li Dongdong, et al. Region Based Rate Control and Bit Allocation for Wireless Video Transmission[J]. IEEE Trans. on Multimedia, 2006, 8(1): 1-10.
- [9] Chai D, Ngan K N. Face Segmentation Using Skin-color Map in Videophone Applications[J]. IEEE Trans. on Circuits and Systems for Video Technology, 1999, 9(4): 551-564.
- [10] Muir L J, Richardson I E G. Perceptions of Sign Language and Its Application to Visual Communications for Deaf People[J]. Journal of Deaf Studies and Deaf Education, 2005, 10(4): 390-401.

编辑 顾逸斐

(上接第 282 页)

6 结束语

本文处理共因失效的方法为进行含共因失效的复杂系统可靠性分析提供了一种新的思路。下一阶段将在更加全面地考虑共因失效对系统影响的同时,着手研究如何用 GO 法解决多阶段系统考虑共因失效时的可靠性分析问题,拓展 GO 法在可靠性分析领域的适用范围。

参考文献

- [1] 李翠玲, 李剑锋, 谢里阳. 冗余系统共因失效概率预测模型[J]. 东北大学学报: 自然科学版, 2006, 27(2): 213-216.
- [2] Levitin G. Incorporating Common-cause Failures into Non-repairable Multistate Series-parallel System Analysis[J]. IEEE Trans. on Reliability, 2001, 50(4): 380-388.
- [3] Xing Liudong, Meshkat L, Susan K D. Reliability Analysis of Hierarchical Computer-based Systems Subject to Common-cause Failures[J]. Reliability Engineering and System Safety, 2007, 92(3): 351-359.
- [4] Tang Zhihua, Dugan J B. An Integrated Method for Incorporating Common Cause Failures in System Analysis[C]//Proc. of Annual Reliability and Maintainability Symposium. Los Angeles, USA: IEEE Press, 2004: 610-614.
- [5] 沈祖培, 黄祥瑞. GO 法原理及应用[M]. 北京: 清华大学出版社, 2004.
- [6] Shen Zupei, Gao Jie, Huang Xiangrui. A New Quantification Algorithm for the GO Methodology[J]. Reliability Engineering and System Safety, 2000, 67(3): 241-247.
- [7] Shen Zupei, Gao Jia, Huang Xiangrui. An Exact Algorithm Dealing with Shared Signals in the GO Methodology[J]. Reliability Engineering and System Safety, 2001, 73(2): 177-181.
- [8] 沈祖培, 唐 辉. 有共因失效的系统可靠性的 GO 法分析[J]. 清华大学学报: 自然科学版, 2006, 46(6): 829-832.
- [9] 沈祖培, 黄立刚, 李晓东, 等. 大亚湾核电站外电源系统可靠性分析中 GO 法的应用[J]. 核动力工程, 2003, 24(1): 68-72.
- [10] 宋东昱, 沈祖培. 天然气管道系统可靠性分析中 GO 法的应用[J]. 石油工业技术监督, 2000, 16(11): 1-7.
- [11] 林 洁, 颜兆林, 龚时雨, 等. GO-FLOW 方法在共因故障中的应用[J]. 工业安全与环保, 2004, 30(2): 11-13.
- [12] 林 洁, 颜兆林, 龚时雨, 等. GO-FLOW 方法及其改进的精确算法[J]. 系统工程与电子技术, 2005, 27(1): 193-196.

编辑 顾逸斐

