arXiv:2304.12384v1 [cs.MM] 24 Apr 2023

面向自适应视频流高效编码的绿色视频复杂度分析

Vignesh V Menon[1], Christian Feldmann[2], KlausSchoeffmann[1], MohammadGhanbari[3], 以及 Christian Timmerer[1]Christian Doppler 实验室 ATHENA，信息技术学院（ ITEC ），阿尔卑斯 - 亚得里亚大学克拉根福校区，奥地利[2]Bitmovin，奥地利[3]计算机科学与电子工程学院，埃塞克斯大学，科尔切斯特，英国

摘要

对于自适应流应用而言，需要低复杂度且准确的视频复杂度特征来实时分析视频内容，从而确保快速、压缩高效且不中断的视频流传输。当前最先进的视频复杂度特征为空间信息（SI）和时间信息（TI）特征，但它们与自适应流应用中的编码参数相关性不佳。为此，视频复杂度分析器（VCA）应运而生，其基于离散余弦变换 （DCT）能量确定相关特征。本文针对 VCA 提出了优化方案，以实现更快速且更节能的视频复杂度分析。实验结果表明，VCA v2.0 使用八个 CPU 线程、单指令多数据（SIMD）和低通 DCT 优化，以更高的准确度确定超高清 8 位视频的七项复杂度特征，速度最高可达 292.68 fps，能耗比参考 SITI 实现低 97.06%。

CCS 概念

• **信息系统** → **多媒体流**； • **软件及其工程** → 软件性能。

关键词

视频复杂度分析；离散余弦变换；多线程；低通优化。

ACM 参考格式：

Vignesh V Menon,[1] ChristianFeldmann[2],Klaus Schoeffmann[1],Mohammad Ghanbari[3], 和 Christian Timmerer[1]。2023。面向自适应视频流高效编码的绿色视频复杂度分析。在第一届 *ACM* 绿色多媒体系统研讨会（*GMSys '23*），*2023*年*6*月*7*日至*10*日，加拿大不列颠哥伦比亚省温哥华。，共 3 页。
https://doi.org/10.1145/3593908.3593942

1 引言

视频复杂度分析是众多视频流应用的关键步骤。Haseeb 等人 [6] 在率失真建模中使用了空间和时间复杂度信息。视频复杂度评估对于 QoE 评估指标也至关重要 [2, 5, 11, 13, 15, 17]。Pinson 等人 [15] 通过利用视频序列的空间内容对视频质量进行了客观测量

允许免费复制本作品的部分或全部内容用于个人或课堂用途，但不得为营利或商业利益而复制或分发，且复制件须附上本声明及首页的完整引文。必须尊重本作品第三方组件的版权。其他用途请联系所有者/作者。*GMSys '23*，*2023*年*6*月*7*日至*10*日，加拿大不列颠哥伦比亚省温哥华 © 2023 版权所有：所有者/作者。ACM ISBN 979-8-4007-0196-2/23/06。
https://doi.org/10.1145/3593908.3593942

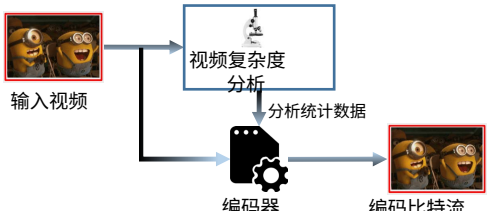

图1：采用视频内容复杂度分析的内容自适应编码框架。

。Bar- man 等人 [17], Goring 等人 [5], 以及 Zadtootaghaj 等人 [2] 提出了基于机器学习的QoE模型，其中空间复杂度和时间复杂度值与其他影响因素结合使用，以对游戏视频进行质量估计。快速视频复杂度分析用于在线逐标题编码方案 [4], ，此类方案可确定优化分辨率 [9], 、比特率阶梯 [10], 、帧率 [8],、可压缩性 [7] 以及其他适用于实时*HTTP*自适应流媒体（HAS）应用的相关编码参数 [14]。

SITI[1] 是用于评估空间和时间复杂度的先进开源软件。空间信息（SI）表示视频中的最大空间细节。对每个视频帧 $F(x, y, t)$应用Sobel滤波器 $[Sobel(F(x, y, t))]$进行滤波，然后计算每个经Sobel滤波的视频帧 $std[Sobel(F(x, y, t))]$的标准偏-差。选取视频序列中最大的标准差值作为SI，表示为 $SI = max\{std[Sobel(F(x, y, t))]\}$。时间信息（TI）是两个连续帧之间的最大时间变化。连续帧之间的差异表示为 $D(x, y, t) = F(x, y, t)) - F(x, y, t - 1))$。计算每个差值 $D(x, y, t)$的标准差 $std[D(x, y, t)]$。TI取视频序列中 $std[D(x, y, t)]$的最大值，表示为 $TI = max\{std[D(x, y, t)]\}$。

2 视频复杂度分析器

视频复杂度分析器[2] (VCA)确定七种基于分块DCT能量的特征，包括亮度纹理能量 E_Y、亮度纹理能量的梯度 h、发光度 L_Y、色度纹理能量 E_U 以及 E_V（分别针对 U 和 V 平面），以及色度 L_U 和 L_V（分别针对 U 和 V 平面） [12]。为了实现绿色且快速的视频复杂度分析，引入了以下优化：

（*1*）多线程优化：为提升多核 CPU 上的性能，VCA 利用多线程机制。多线程优化会在一个 VCA

[1]https://github.com/Telecommunication-Telemedia-Assessment/SITI，最后访问时间：2023年4月20日。[2]https://vca.itec.aau.at，最后访问时间：2023年4月20日。

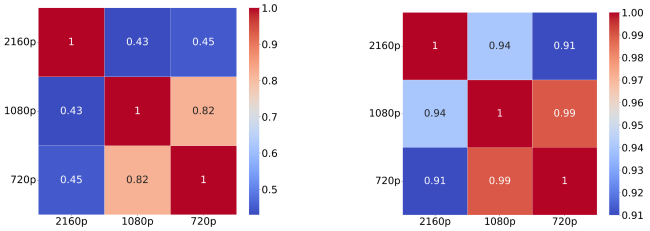

图2：VCD数据集 [1]在多个分辨率下空间复杂度特征（a）SI和（b） E_Y 之间的皮尔逊相关系数。

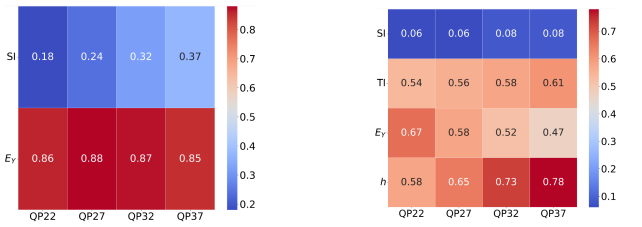

图3：在全帧内配置下使用x265编码器的*medium* 预设时，（a）SI与 E_Y分别和码率之间的皮尔逊相关系数；以及在低延迟P图像配置下使用x265编码器的*ultrafast* 预设时，（b）SI、 E_Y、TI和 h 与码率 之间的皮尔逊相关系数，VCD数据集为 [1]。

执行实例中创建多个线程。这些线程共享进程资源，彼此独立并发执行。各线程独立计算每个块的 DCT 能量。

（2）*x86 SIMD*优化： VCA利用以内在函数代码和汇编代码实现的DCT函数的SIMD优化 [16] ，适用于x86架构。DCT的内在函数代码针对Intel SSSE3架构执行。手写的Intel SSSE3和AVX2指令可加速每个块的DCT计算。

（3）低通 *DCT* 优化：与 SI 不同，如图 2 所示， E_Y特征在不同分辨率之间表现出更好的相关性。因此，VCA v2.0 实现了低通 DCT 优化，即在空间下采样倍数为二的视频上评估复杂度特征。预计该优化可提高处理速度并降低能耗。

3 性能评估

首先，将 E_Y 的准确性与最先进的 SI 特征进行比较。在此基础上，评估 SI 和 E_Y 特征与全帧内配置下码率 [3] 的相关性，因为该配置下的码率被视为空间复杂度的真实值。如图 3a 所示，SI 与码率之间的平均皮尔逊相关系数（PCC）为 0.28，而 E_Y 与码率之间的平均 PCC 分别为 0.86。因此， E_Y 比 SI 能更好地表示空间复杂度。

其次，在低延迟 P 图像配置 [3]下，根据 E_Y 和 h 与编码比特率的相关性评估其准确性。如图 3b 所示， E_Y和 h 与编码比特率高度相关。

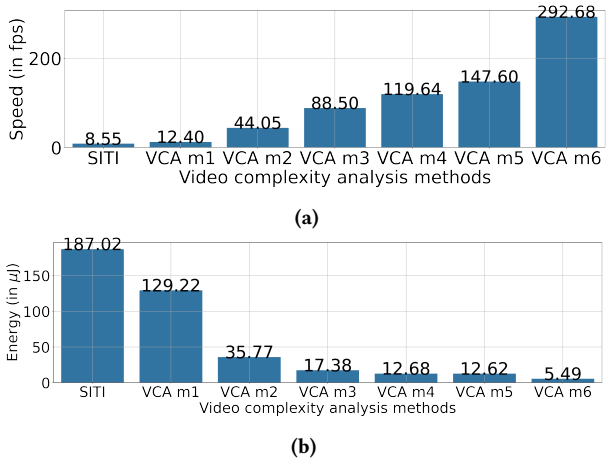

图 4：（a）视频内容复杂度分析方法的处理时间；（b）能耗。

第三，在配备英特尔 i7-11370H 处理器和 16GB 随机存取存储器的计算机上，将 VCA 与 SITI 在处理超高清（UHD）8 位视频时的分析速度和能耗进行比较。能耗使用 *codecarbon*[3] 测量。如图 4 所示，参考 SITI 实现的内容分析速率为 8.55fps，能耗为 187.02 μJ。参考 VCA 实现（记为 VCA m1）即未采用多线程、SIMD 和低通优化，分析速率为 12.40 fps，能耗比 SITI 实现低 30.90%。仅使用低通 DCT 优化（VCA m2）时，与 VCA m1 相比，能耗降低 72.01%。在 VCA m2 的基础上添加 SIMD 优化（记为 VCA m3）后，能耗进一步降低 51.41%。VCA m4、m5 和 m6 模式分别表示使用两个、四个和八个 CPU 线程。随着使用的线程数增加，分析速度加快，能耗降低。启用所有优化并使用八个 CPU 线程时，VCA 的速度是 SITI 的 33 倍，能耗降低 97.06%。

4 结论

本文提出了一种面向视频流应用中高效编码的节能型视频内容复杂度分析方法。利用 VCA v2.0 提取低复杂度的基于 DCT 的能量特征，编码器可据此制定比特率阶梯、帧类型、块划分等决策。多线程、x86 单指令多数据和低通 DCT 优化提升了 VCA 实现的能效。与最先进的视频复杂度分析 SITI 实现相比，VCA v2.0 能够更准确地估计视频复杂度，同时将能耗降低 97.06%。VCA 以 GNU GPLv3 许可证发布，网址为 https://github.com/cd-athena/VCA。

5 致谢

谨此致谢奥地利联邦数字与经济事务部、国家研究、技术与发展基金会以及克里斯蒂安·多普勒研究

[3]https://codecarbon.io/，最后访问日期：2023年4月20日

协会提供的资助。克里斯蒂安·多普勒实验室 ATHENA: https://athena.itec.aau.at/。

参考文献

[1] Hadi Amirpour、Vignesh V Menon、Samira Afzal、Mohammad Ghanbari 和 Christian Timmerer。2022。VCD：视频复杂度数据集（*MMSys '22*）。计算机协会，纽约州纽约，美国，234–239。https: //doi.org/ 10 1145 3524273 3532892 。/。N. Barman、E. Jammeh、S. Ghorashi 和 M. Martini。2019。基于机器学习的无参考视频质量估计，用于被动游戏视频流媒体应用。*IEEE Access* 7（2019），74511–74527。https: //doi.org/ 10 1109 2019 2920477 。/ACCESS。Jill Boyce、Karsten Suehring、Xiang Li 和 Vadim Seregin。2018。*JVET-J1010*：*JVET* 通用测试条件和软件参考配置。[4] Jan De Cock、Zhi Li、Megha Manohara 和 Anne Aaron。2016。云端基于复杂度的一致质量编码。载于*2016*年*IEEE*国际图像处理会议（*ICIP*）。1484–1488。https://doi.org/10.1109/ICIP.2016.7532605 ISSN：2381-8549。[5] S. Göring、R. Rao 和 A. Raake。2019。nofu——用于游戏内容的轻量级无参考像素级视频质量模型。载于*2019*年第十一届多媒体体验质量国际会议（*QoMEX*）。1–6。https://doi.org/ 10 1109 2019 8743262 。/QoMEX。A. Haseeb、M. Martini、S. Cicalò 和 V. Tralli。2012。实时 MGS 编码的码率与失真建模及自适应。载于*2012*年无线先进技术会议（*WiAd*）。 85 89 10 1109 2012 6296574 –。https://doi.org/。/WiAd。Maria Martini 和 Shaymaa Al-Juboori。。直播视频压缩优化的内容特征分析。（2022年12月）。https://doi.org/10.36227/techrxiv。 21743312 2。v Vignesh V Menon、Hadi Amirpour、Mohammad Ghanbari 和 Christian Timmerer。2022。CODA：用于高帧率视频流传输的内容感知帧丢弃算法。载于*2022*年数据压缩会议（*DCC*）。475–475。https://doi.org/ 10 1109 52660 2022 00086 。/DCC。Vignesh V Menon、Hadi Amirpour、Mohammad Ghanbari 和 Christian Timmerer。2022。OPTE：用于直播视频流传输的在线逐标题编码。载于*ICASSP 2022 - 2022*年*IEEE*国际声学、语音与信号处理会议（*ICASSP*）。 1865 1869 10 1109 43922 2022 9746745 –。https://doi.org/。/ICASSP。Vignesh V Menon、Hadi Amirpour、Mohammad Ghanbari 和 Christian Timmerer。2022。面向感知的自适应视频流逐标题编码。载于*2022*年*IEEE*国际多媒体与博览会会议（*ICME*）。IEEE计算机学会，洛斯阿拉米托斯，加利福尼亚州，美国，1–6。https://doi.org/10.1109/ ICME52920.2022.9859744[11] Vignesh V Menon、Reza Farahani、Prajit T Rajendran、Mohammad Ghanbari、Hermann Hellwagner 和 Christian Timmerer。2023。自适应视频流的转码质量预测。载于第二届*Mile-High*视频会议论文集。计算机协会，纽约州纽约，美国。https://doi.org/10.1145/3588444.3591012[12] Vignesh V Menon、Christian Feldmann、Hadi Amirpour、Mohammad Ghanbari 和 Christian Timmerer。2022。VCA：视频复杂度分析器（*MMSys '22*）。计算机协会，纽约州纽约，美国，259–264。https: //doi.org/10.1145/3524273.3532896[13] Vignesh V Menon、Prajit T Rajendran、Reza Farahani、Klaus Schoeffmann 和 Christian Timmerer。2023。融合纹理信息的流媒体应用视频质量评估。arXiv：cs.MM/2302.14465 https://arxiv.org/ abs/2302.14465[14] Vignesh V Menon、Prajit T Rajendran、Christian Feldmann、Klaus Schoeffmann、Mohammad Ghanbari 和 Christian Timmerer。2023。面向自适应直播流传输的基于JND感知的两遍逐标题编码方案。（2023）。https://doi.org/ 10 36227 22256704 1 。/techrxiv。v M.H. Pinson 和 S. Wolf。2004。客观测量视频质量的一种新的标准化方法。*IEEE*广播学报50，3（2004），312–322。https://doi.org/ 10 1109 2004 834028 。/TBC。Praveen Kumar Tiwari、Vignesh V Menon、Jayashri Murugan、Jayashree Chandrasekaran、Gopi Satykrishna Akisetty、Pradeep Ramachandran、Sravanthi Kota Venkata、Christopher A Bird 和 Kevin Cone。2018。使用英特尔®高级矢量扩展512加速x265。英特尔开发者页面上的白皮书（2018）。[17] S. Zadtootaghaj、N. Barman、S. Schmidt、M. Martini 和 S. Möller。2018。NR-GVQM：无参考游戏视频质量指标。载于*2018*年*IEEE*国际多媒体研讨会（*ISM*）。IEEE，131–134。https://doi.org/10.1109/ISM。2018.00031