|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 重庆邮电大学  CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS | | | | |
| 博士学位论文  DOCTORAL DISSERTATION | | | | |
|  | | | | |
|  | | | | |
| **论文题目** | | **基于深度神经网络的HEVC帧内快速** | |  |
|  | | **深度决策算法研究** | |  |
|  | |  | |  |
|  | **学科专业** | | **计算机科学与技术** |  |
|  | **学 　号** | | **S210231282** |  |
|  | **作者姓名** | | **周广义** |  |
|  | **指导教师** | |  |  |
|  | **学　　院** | |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 学校代码 | 10617 | | | | | UDC |  | |
| 分类号 |  | | | | | 密级 |  | |
| 学　位　论　文 | | | | | | | | |
| **基于深度神经网络的HEVC帧内快速深度决策算法研究** | | | | | | | | |
|  | | | | | | | | |
|  | | | | **周广义** | | |  | |
|  | | | |  | | |  | |
|  | | | | | | | | |
| 指导教师 | | **某某某** | | | | | **教　授** |  |
|  | | **某 某** | | | | | **副教授** |  |
|  | |  | | | | |  |  |
|  | |  | | | | |  |  |
|  | |  | | | | | |  |
|  | |  | | | | | |  |
|  | |  | | | | | |  |
| 申请学位级别 | | **硕士** | | | 学科专业 | |  | |
| 专业学位领域 | |  | | | | | | |
| 答辩委员会主席 | | **某某某 教 授** | | | 论文答辩日期 | | **2021年5月20日** | |
| 学位授予单位和日期 | | | **重庆邮电大学** | | | | **2021年6月** | |

|  |  |
| --- | --- |
| **Intra-frame fast depthdecision algorithm for HEVC based on deep neural network** | |
| A Doctoral Dissertation Submitted to  Chongqing University of Posts and Telecommunications | |
|  | |
| Discipline | **Computer Science and Technology** |
| Student ID | **S210231282** |
| Author |  |
| Supervisor |  |
| School | **School of Computer Science and Technology** |

**重庆邮电大学**

**学位论文独创性声明**

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文中不包含其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在论文中以明确方式标明。本人完全知晓本声明的法律后果由本人承担。

学位论文作者签名：

日期： 年 月 日

**重庆邮电大学**

**学位论文使用授权书**

本人同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。

本学位论文属于 ：

□**公开论文**

□**涉密论文**，保密 年，过保密期后适用本授权书。

（请在以上方框内选择打“**√**”）

作者签名： 导师签名：

日期： 年 月 日

**摘　要**

随着大数据时代的到来和无时不在的移动互联网通信技术的发展，人们通信方式的内容形式已经变成了大量的图片和视频。为了方便视频数据的保存与传输最有效的方法是将视频进行压缩处理，现阶段最流行也是最普遍的处理视频压缩的方式是（高效视频编码HEVC），HEVC在上一代视频压缩标准（Advanced Video Coding：AVC）的基础上提高了50%的压缩率，同时带来了编码复杂度的提升，在当前的高效视频编码（HEVC）领域，帧内编码树单元（CTU）的深度划分决策对编码效率和视频质量具有显著影响。然而，传统的深度划分算法往往依赖于复杂的启发式规则或简单的机器学习方法，这限制了划分性能的提升并增加了编码复杂度。为了解决这些问题，本文旨在来研究通过利用深度神经网络（DNN）的强大特征提取和分类能力，实现更高效且更准确的CTU深度划分。从而降低HEVC的编码复杂度，本文的研究工作如下

（1）本研究的第一个贡献是设计了一种多尺度多输入的卷积神经网络（MMCNN），它能够同时处理多个不同量化参数（QP值）下的图像块输入，并且考虑不同尺度的CU特征。这种设计使得网络能够捕捉到跨尺度的特征关联，从而提升了特征的表达能力并增强了网络对不同QP条件下的适应性。

（2）通过分析研究工作（1）中提出方法的一些局限性，针对分类结果错误累积问题，本文提出了一种互补分类策略（CCS）。该策略由两个独立的(Multi-scale-multi-input Complementation Classification Network :MMCNN)组成，分别对CTU中的32×32 CU进行三分类深度预测，以及对16×16 CU进行二分类划分预测。通过将两个网络的预测结果进行融合，并采用投票机制来确定每个CU的最终深度值，实现了分类结果的互补，从而显著提高了深度划分的准确性。

为了验证所提出方法的有效性，我们在多个HEVC标准测试序列上进行了广泛的实验。实验结果表明，与现有的帧内深度划分方法相比，我们的MCCN框架在保持相似编码效率的同时，可以有效降低编码复杂度。此外，所提出的方法具有较好的泛化能力，即在不同类型的视频内容上都能达到稳定的性能提升。本文提出的模型在客观评价指标和主观视觉效果两个方面都获得了良好的效果

**关键词：**高效视频编码，编码树单元，多尺度多输入，深度神经网络，互补分类网络

**ABSTRACT**

Dissertation /Thesis is postgraduate’s main academic performance to display her/his works of scientific research, which shows the author’s new invention, new theory or new opinion in her/his research. It is the crucial document for the graduate students to apply for degree, and it is also the important scientific research literature and the valuable wealth of society.

In order to further standardize the format of dissertation/thesis writing and improve graduate dissertation/thesis quality, this temolate is formulated with reference to the national standard "Rules for Dissertation Writing" (GB/T 7713.1-2006) and the reality of CQUPT.

**Keywords:**Dissertation/Thesis, Writing Specification,Thesis Template, Chongqing University of Posts and Telecommunications

目　录

[摘　要 I](#_Toc153361582)

[ABSTRACT II](#_Toc153361583)

[图目录 V](#_Toc153361584)

[表目录 VI](#_Toc153361585)

[主要符号表 VII](#_Toc153361586)

[缩略词表 VIII](#_Toc153361587)

[第1章 绪论 1](#_Toc153361588)

[1.1 研究背景及意义 1](#_Toc153361589)

[1.2 国内外研究现状 1](#_Toc153361590)

[1.3 论文研究主要内容 1](#_Toc153361591)

[1.4 论文组织结构 1](#_Toc153361592)

[第一章 绪论 2](#_Toc153361593)

[（一）研究背景及意义 2](#_Toc153361594)

[1. 研究背景 2](#_Toc153361595)

[2. 研究意义 2](#_Toc153361596)

[（二）国内外研究现状 2](#_Toc153361597)

[（三）论文研究主要内容 2](#_Toc153361598)

[（四）论文组织结构 2](#_Toc153361599)

[第2章 论文结构及文字格式 3](#_Toc153361600)

[2.1 本章引言 3](#_Toc153361602)

[2.2 论文结构 3](#_Toc153361603)

[2.3 字数要求 3](#_Toc153361604)

[2.3.1 硕士论文要求 3](#_Toc153361605)

[2.3.1 博士论文要求 4](#_Toc153361606)

[2.4 字体和段落 4](#_Toc153361607)

[2.5 本章小结 5](#_Toc153361608)

[第3章 图表、公式格式和印制要求 6](#_Toc153361609)

[3.1 本章引言 6](#_Toc153361611)

[3.2 图和表格式 6](#_Toc153361612)

[3.2.1 图 6](#_Toc153361613)

[3.2.2 表 8](#_Toc153361614)

[3.3 公式格式 9](#_Toc153361615)

[3.4 印制要求 10](#_Toc153361616)

[3.5 本章小结 10](#_Toc153361617)

[第4章 总结与展望 11](#_Toc153361618)

[4.1 主要结论 11](#_Toc153361620)

[4.2 研究展望 11](#_Toc153361621)

[参考文献 12](#_Toc153361622)

[附录A 各学院中英文名称对照表 15](#_Toc153361623)

[作者简介 16](#_Toc153361624)

[1. 基本情况 16](#_Toc153361625)

[2. 教育和工作经历 16](#_Toc153361626)

[3. 攻读学位期间的研究成果 16](#_Toc153361627)

[3.1 发表的学术论文和著作 16](#_Toc153361628)

[3.2 申请（授权）专利 16](#_Toc153361629)

[3.3 参与的科研项目及获奖 16](#_Toc153361630)

[致 谢 17](#_Toc153361631)

图目录

[图2‑1 学位论文基本结构 3](#_Toc152868291)

[图3‑1 不同缩放系数 ν 的缩放结果 6](#_Toc152868292)

[图3‑2 HARP模型中通过边坍塌进行结构粒度粗化的策略[64] 。 7](#_Toc152868293)

[图3‑3 HARP模型中通过边坍塌进行结构粒度粗化的策略[64] 7](#_Toc152868294)

表目录

[表2‑1 中、英文字号对应关系 4](#_Toc128917515)

[表2‑2 主要文字及段落格式要求 4](#_Toc128917516)

[表3‑1 电流类型对效率的影响 7](#_Toc128917517)

[表3‑2 球队的比赛结果统计表[5] 8](#_Toc128917518)

[表3‑3 学位论文页面设置 9](#_Toc128917519)

主要符号表

|  |  |  |
| --- | --- | --- |
| **符号** | **说明** | **页码** |
| 𝒫Ω(·) | 集合Ω上的投影算子 | 6 |
| *c* | 电磁波的相平面速度 | 10 |

缩略词表

|  |  |  |
| --- | --- | --- |
| **英文缩写** | **英文全称** | **中文全称** |
| CQUPT | Chongqing University of Posts and Telecommunications | 重庆邮电大学 |
| IEEE | Institute of Electrical and Electronics Engineers | 电气和电子工程师协会 |

第一章 绪论

## 1.1研究背景及意义

随着大数据时代的到来和无时不在的移动互联网通信技术的发展，以及互联网、物联网的全面覆盖，一个以大数据、大流通为特征的信息时代已经开启。人们对于具有生动、全面、信息量大为特征的视频信息的需求日益增大，使得大数据中的海量视频信息的占比达到一个新高度。随着4K、8K等高清视频格式的普及，视频数据量急剧增加，对存储和带宽提出了更高的要求，移动设备和社交媒体的普及导致视频内容的生产和消费量大幅上升。传统的视频编码标准（如H.264）在高清视频传输和存储方面面临带宽和存储限制。 网络服务提供商和内容分发网络寻求降低数据传输成本。因此视频压缩变得越来越重要。

新一代的视频压缩标准（High Efficiency Video Coding，HEVC）的出现，旨在提供比H.264更高的数据压缩率，理论上能够在保持相同视频质量的情况下将码率降低约50%。这意味着相同质量的视频，使用HEVC编码后，文件大小和所需的传输带宽大大减少。HEVC的出现使得更复杂的编码算法成为可能，它提出了编码树单元（Coding Tree Unit：CTU）的概念，每一个CTU可进行不同深度的划分，从64×64到最小的8×8大小不等，同时还将预测模式提高到了35种之多，以提高预测块的预测精度。通过递归地计算所有深度上所有图像块的率失真代价（Rate-Distortion cost：RD cost）来决定每个编码单元（Coding Unit ：CU)的深度。从而达到最优的编码效果，由于这个过程，计算复杂度和编码时间显着增加，HEVC 的编码时间比（H.264/AVC）平均高253%。这导致HEVC的编码复杂度变得极大，尤其是针对CTU深度的划分[1]-[4]。因此，如何能够在保证较低BD-BR（Bjøntegaard delta bit rate）的同时尽可能提升CTU深度划分的效率降低HEVC的编码复杂度已经成为了近年来人们研究的热点。

为了降低HEVC的编码复杂度，减少编码时间，人们提出了两大类提升CTU深度划分效率的方法：基于启发式的方法（Heuristic-based Methods）和基于深度学习的方法（Deep-learning-based Methods）。对于启发式的方法而言，几乎都是通过人为提取图像中的纹理信息特征或者人为设置某些规则来提前终止CTU的深度划分，以避免暴力递归的RDO（Rate Distortion Optimization）搜索，从而达到降低深度划分的复杂度的目的，例如：利用CTU中每个CU的纹理特征以及其与相邻CU的空间相关性来决定CTU的划分结果[5]-[18]。

近年来，随着人工智能的发展。深度学习神经网络，尤其是卷积神经网络（Convolutional Neural Network：CNN）已经被广泛用于优化HEVC帧内模式的CTU深度划分中。用神经网络代替RDO的计算过程，通过给网络输入当前编码块的像素值，使用神经网络去做出深度的划分，这将大大节省HEVC的编码时间。

但是，无论是基于启发式的方法还是基于学习的方法，它们的性能仍具有较大的提升空间，网络模型需要进一步提升，划分策略也需进一步完善。因此，综上所述，本文的研究目标是充分利用神经网络在图像特征提取上的优势与HEVC编码系统相结合，进一步降低HEVC系统的编码复杂度。

## 1.2国内外研究现状

由HECV高压缩率所带来的复杂度的提升，一直是近年来的研究热点问题，在帧内编码过程中，发现CU划分的递归过程贡献了最大比例的编码时间，大约占据整个编码时间的80%。由于CU的深度划分是影响编码时间的重要部分。因此，在过去的五年中，出现了大量降低HEVC编码复杂性的方法。在这些HEVC复杂性降低方法中，大多数都试图简化CU分区的过程。这些方法的基本思想是提前预测CU分区，而不是蛮力递归RD优化 (RDO) 搜索。

预测CU分区的早期工作是启发式方法，它们探索一些中间特征，以便在检查所有可能的之前尽早确定CU划分。在启发式的方法中，也出现了一些基于机器学习的方法，但它仍需要人工设置一些CTU纹理和与相邻CTU之间的空间相关特征来描述CTU的分区。

在近些年，旨在加速HEVC帧内编码的研究一直在从基于启发式的策略转变为基于学习的策略，从而出现了大量基于深度学习的方法（Deep-learning-based Methods）。在这些方法中，大量使用了CNN（Convolutional Neural Network）网络结构，来预测CTU的划分。

### 基于启发式的方法

为了有效降低HEVC帧内编码的复杂度，人们提出了大量的基于启发式的方法来提升CTU深度划分效率。基于启发式的方法通常是通过人为提取CTU中CU的特征或者人为设置某些规则来提前终止CTU的深度划分过程，以避免暴力递归的RDO搜索，从而达到降低深度划分复杂度的目的。空间/时间/深度相关性、统计率失真 (RD) 和纹理信息是预测CU深度决策的三个常用特征。

在Chen[5]等人的研究中就利用了时间相关性来预测CU的深度范围，具体来说基于对前一帧获得的模式信息的分析来衡量视频内容的复杂度。然后，基于该特征与编码树单元（CTU）深度范围的关系构建模型。根据模型，跳过了编码单元 （CU） 划分的不必要操作。Cen[6]等人提出一种利用序列帧空间相关性的快速编码单元（CU）深度决策机制。所提出的机制包括一种自适应CU深度范围确定和CU深度比较算法。CU深度范围根据CU深度在同一序列中的分布确定。因此，可以跳过范围之外的 CU深度的速率失真（RD）成本计算。此外，通过将相邻 CU 的深度与当前CU深度进行比较，可以进一步跳过当前CU深度的RD成本计算。

Xiong[7]等人通过研究运动发散与率失真（RD）成本之间的隐含关系，提出了一种基于金字塔运动发散的模型来快速决定CU是否需要划分。Kim[8]等人]提出了一种根据每个CU中高频关键点的数量来决定CU是否分裂的方法。Shen[9]等人提出通过使用重要且计算友好的特征来做出CU大小决策，并定义了贝叶斯决策规则，以通过同时最小化贝叶斯风险来帮助准确快速地选择CU大小。

此外，基于编码图像的时间和空间特性，Wang[10]等人介绍了一种基于纹理复杂度和运动特性的剪枝策略。然后，进一步使用相邻编码块的运动信息来快速确定参考图像，从而减少编码时间。Liu[11]等人提出了一种基于纹理复杂度和图像方向的预测模式自适应决策方法。

Shen[12]等人提出了一种用于HEVC内编码的快速CU大小决策算法，通过减少每个CTU中需要检查的候选CU数量来加快该过程。该算法的新颖性在于（1）基于纹理均匀性，提出了一种基于自适应阈值的早期CU大小决策确定方法；（2）基于纹理属性和相邻编码CU的编码信息的组合，提出了一种新的大CU内部预测旁路策略。在文献[13]中，Gweon等人通过在亮度和色度的所有编码块标志Coded Block Flags（CBF）为零的情况下跳过对当前CU深度的RDcost计算来减少计算复杂度。

Choi[14]等人通过判断当前CU深度下的最佳预测模式是否为SKIP来提前结束对该CU深度的划分。Yang[15]等人通过采用差分运动矢量和编码块标志两个特征作为CU是否划分的标准来降低编码器的复杂度。Cho[16]等人使用了具有低复杂度和低RDcost的贝叶斯决策规则来对CU的划分做出判决。Zhang[17]等人采用时空相关性进行自适应的CU深度搜索，从而降低CTU划分的复杂度。Zhao[18]等人采用低成本的Hadamard cost来作为终止CU划分的标准。Khan[19]等人提出了一种基于视频帧内容自适应的快速CU决策方法：根据视频帧中不同局部区域的内容，自适应的将较小的CU组合成较大的CU。

这些传统的方法在一定程度上提高了编码单元分割的效率，降低了HEVC的编码复杂度。在之后出现了一些使用机器学习的方法来进一步挖掘图像的纹理特征。

在文献[20]中，Zhang等人提出了一种基于梯度的机器学习粗模式决策和RD优化算法。为了实现快速的CU划分决策，Zhang等人提出了两个以深度差和RD代价为特征的线性支持向量机来进行早期CU分裂和早期CU终止决策。

Hu[21]等人基于离散余弦变换系数透明复合模型，将帧内编码模式判决问题转换成贝叶斯判决问题，提出了一种基于离群点的快速帧内模式判决算法。Liu[22]等人提出了一种基于支持向量机的三分类器来判断CU是否需要划分。Jamali[23]等人提出了一种针对帧内模式的RDO成本预测方法和一种基于梯度的模式筛选方法。Jamali等人采用Prewitt算子来从候选模式列表中筛选出方向相关的模式。由以上分析可知，现有的基于启发式的方法几乎都是通过人为手动的对图像中某些成分的特征（如边缘、纹理）来设置一些规则，使得满足这些特征的CU可以提前终止划分，来达到降低HEVC帧内编码复杂度的目的。然而自然图像中边缘和纹理是千变万化的，不同CU之间的边缘/纹理特征以及它们之间的关系是十分复杂的，这些算法不能正确地考虑在各种序列中的CTU的分割趋势。也就是说，它们没有根据图像纹理和编码信息综合考虑CTU划分趋势。它们使用在编码过程中可以获得的各种因素来决定是否跳过对RD成本的操作。然而，存在太多的因素，并且难以具体地确定哪些与CTU划分相关。因此，仅参考元素中的一些来确定CTU分割，这可能导致较差的预测。这种人为设置相关规则的方式是几乎不可能满足所有的CU的。因此，虽然启发式的方法能够在一定程度上降低HEVC的编码复杂度，但其性能仍有较大提升空间。

### 基于学习的方法

近年来，随着深度学习神经网络的出现，尤其CNN，因为其强大的特征提取能力和非线性表达能力和自动特征学习，它们可以有效地用于各种视频编解码场景，从而增强了所提出的算法的通用性，因此，基于CNN的方法越来越受到关注，并在视频编码和解码领域中找到各种应用，目前已经被广泛用于优化HEVC帧内模式的CTU深度划分中。通过神经网络可以很好的代替人工方式来提取CTU中的图像特征和纹理信息，来提前终止CU的深度划分。

Yu[24]等人提出了一种基于CNN的二分类CTU深度划分方法：从初始的64×64图像块开始，判断其是否需要划分，如需要划分则继续对下一层的每个32×32图像块进行二分类判断，如不需要划分则直接终止对该层CU的划分，他们在非常有限的环境下应用CNN。编码仅在INTRA模式下执行，该模式具有相对较轻的计算负载。此外，他们为所有深度设计了相同的神经网络架构，无法有效地考虑每个CU深度划分的属性。

在文献[25]中，Li等人使用CNN来降低INTRA模式下的CTU决策复杂度，他们针对不同深度的CU设计了不同的神经网络以实现更具针对性的CU深度预测，但是Li等人却忽略了不同尺寸CU间的关联性。因此它不能很好地反映CTU分区趋势。Kim[26]设计了一种基于神经网络推理的编码单元深度判决算法，该算法不仅适用于帧内模式，而且适用于帧间模式，并根据神经网络推理的结果将编码单元划分为64×64、32×32、16×16三种深度。此外，通过为四个量化参数（QP）设计每个网络，可以适用于各种QP。Xu[27]等人提出了一种ETH-CNN（Early Terminated Hierarchical-CNN）来进行分等级的CU深度地图预测，以降低帧内模式HEVC的复杂度。Kuanar[28]等人提出了一种基于卷积神经网络的算法，该算法可以学习区域图像特征并执行分类工作将空间模式分为四类，这些分类结果用于编码器下游系统，以在每个CTU块中寻找最佳编码单元，从而减少预测模式的数量。

Li[29]等人提出了一种基于CNN的速率失真优化算法，通过结合混合注意力机制模块，以优化HEVC中CTU划分的RDO。首先，Li等人设计了一个与当前CU大小相兼容的CNN来预测CU的深度划分。其次，他们设计了一个卷积块，以增强CU块之间的信息交互，并将卷积块注意力模块（Convolution Block Attention Module:CBAM）引入到CNN中，构建了CNN-CBAM。该模块能够正确关注目标对象，提高CU划分的准确率。

针对CU深度划分的复杂递归计算，Qin[30]等人同样提出了一种基于CNN的CU划分方法：设计了一种改进的VGGNet来替代暴力搜索策略，有效降低了帧内编码的计算复杂度。Fan[31]等人提出了一种采用块划分神经网络来预测CTU的块划分结构的方法。Fan等人采用自适应阈值技术来准确管理CNN的预测误差。Wang[32]等人提出了一种基于密集连接卷积神经网络（Densely Connected Convolution Neural Network:D-CNN）来的CU深度划分方法。Wang等人首先设计了一个密集连接块来充分提取CTU的像素特征；然后，Wang等人采用有效的通道注意力（Efficient Channel Attention: ECA）和自适应卷积核大小应来捕获D-CNN卷积通道的特征信息；最后，Wang等人制定了一种阈值优化策略，通过为每个深度选择一个最佳阈值来平衡视频编码的计算复杂性和RD性能。Feng[33]等人提出了一种面向HEVC帧内编码的CNN快速CTU深度划分方法，但是该方法仅以单个尺度的CTU作为输入，导致网络难以学习到小尺寸CU的特征。Hari[34]等人采用不同大小的卷积核（5×5、4×4、2×2）来预测CTU的深度地图。

Lorkiewicz等人[35]提出了一种预训练的神经网络（Artificial Neural Network：ANN）来预测CTU的深度地图，但Lorkiewicz等人同样只考虑了单一尺度的CTU输入。

Ren[36]等人提出了一种基于CNN的帧内块划分（intra block partition CNN：IBP-CNN）网络，通过学习图像块的纹理特征来实现块的深度划分。Feng等人[37]提出了一种基于CNN的HEVC帧内深度范围（HEVC intra depth range CNN：HIDR-CNN）预测网络，通过预测CTU可能的深度范围，并仅在预测的深度范围内执行RDcost计算来减少RDcost的计算次数，从而达到降低编码复杂度的目的。

Li[38]等人提出了一种端到端的快速算法，以辅助帧内编码中CTU的结构划分决策。Imen[39]等人基于LeNet-5和AlexNet，有效降低了检查所有块决策候选者的计算时间，从而深度优化了HEVC编码单元划分模块的全帧内配置。Yao等人[40]构建了一种基于CNN的双网络模型：预测网络和目标网络。通过基于RD函数的强化学习方法，在确定最优CU划分模式的同时获得CU分区的最佳RD估计。

由以上分析可知，虽然现有的基于CNN的深度预测方法能够通过自主学习CTU的结构来做出对应的深度划分决策，但是，一方面，这些方法几乎都是仅采用了一种串联的二分类策略来决定CTU的划分结果，存在着明显的分类错误累积影响。另一方面，现有的基于CNN的CTU深度划分方法几乎都只考虑了单一尺度的输入，而忽略了不同尺度CU之间的特征关联。

## 论文研究的主要内容

在众多的针对加快HEVC编码速度、降低其编码复杂度的方法中本论文关注和研究的重点基于深度神经网络（Deep Neural Network，DNN）的方法，并对目前已有方法的缺点和不足 进行了简要总结与概括,主要包括:

1. 部分方法只使用到了局部特征，而忽略了局部特征与全局特征的结合。
2. 大部分方法未能考虑CTU中不同尺度的特征融合，，无法充分利用原始图像中 的多尺度信息。
3. 神经网络设计过于简单，不能够很好的提取CTU图像特征用于深度的划分。
4. 未能充分考虑量化参数QP对CU深度划分的影响。
5. 对于深度的划分采用串联二级分类方式，容易出现预测的错误累计效应。

综上所述，针对现有方法的缺陷，本论文提出了一种新颖的多尺度多输入卷积神经网络（Multi-scale Multi-input Convolutional Neural Network：MMCNN）。提出的MMCNN同时考虑了将多个QP值（本论文考虑了22、27、32和37四个不同的QP值）和多个不同尺度的CU图像块作为输入，通过对不同尺度CU之间特征信息的有效融合，进一步提升了网络的特征提取能力和表达能力，极大提升了HEVC帧内CTU深度划分的效率

除此之外，不同于现有方法的串联二分类策略，本论文还将多分类融入到CU的深度预测中，提出了一种互补的分类策略（Complementary Classification Strategy ：CCS）：首先，创建两个MMCNN来分别预测CTU中每个的CU的深度（三分类）和判断每个的CU是否需要划分（二分类）；然后，综合考虑三分类和二分类的结果，并采用投票机制来决定每个CU的最终深度结果。通过结合多分类和投票机制，CCS中的三分类和二分类很好的实现了分类结果的相互互补，很好地避免了现有方法中存在的错误累积效应，较大提升了CTU中CU深度预测的准确性。

## 论文组织结构

# 视频编码相关理论



## 本章引言

学位论文……

## 论文结构

## 字数要求

学位论文……

### 硕士论文要求

各学科和学部自定。

### 博士论文要求

各学科和学部自定。

## 字体和段落

## 本章小结

本章介绍了……

# 图表、公式格式和印制要求



## 本章引言

学位论文……

## 图和表格式

### 图

### 表

算法表的示例：XXXXXXXX，如算法3-1所示。

## 公式格式

(3-4)

本文……

## 印制要求

## 本章小结

本章介绍了……

# 总结与展望



## 主要结论

## 研究展望

参考文献

1. Sullivan G J, Ohm J R, Han W J, et al. Overview of the high efficiency video coding (HEVC) standard[J]. IEEE Transactions on circuits and systems for video technology, 2012, 22(12): 1649-1668..
2. Wiegand T, Sullivan G J, Bjontegaard G, et al. Overview of the H. 264/AVC video coding standard[J]. IEEE Transactions on circuits and systems for video technology, 2003, 13(7): 560-576.
3. Lainema J, Bossen F, Han W J, et al. Intra coding of the HEVC standard[J]. IEEE transactions on circuits and systems for video technology, 2012, 22(12): 1792-1801.
4. Pourazad M T, Doutre C, Azimi M, et al. HEVC: The new gold standard for video compression: How does HEVC compare with H. 264/AVC?[J]. IEEE consumer electronics magazine, 2012, 1(3): 36-46.
5. Chen F, Jin D, Peng Z, et al. Fast intra coding algorithm for HEVC based on depth range prediction and mode reduction[J]. Multimedia Tools and Applications, 2018, 77(21): 28375-28394.
6. Cen Y F, Wang W L, Yao X W. A fast CU depth decision mechanism for HEVC[J]. Information Processing Letters, 2015, 115(9): 719-724
7. Xiong J, Li H, Wu Q, et al. A fast HEVC inter CU selection method based on pyramid motion divergence[J]. IEEE transactions on multimedia, 2013, 16(2): 559-564.
8. N. Kim, S. Jeon, H. Shim, B. Jeon, S.-C. Lim, H. Ko, Adaptive keypoint-based CU depth decision for HEVC intra coding,in: Proc. 2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, 2016.
9. X.L. Shen, L. Yu, J. Chen, Fast coding unit size selection for HEVC based on Bayesian decision rule, in: Proc. 2012Picture Coding Symposium, 2012
10. X.J. Wang, Y.L. Xue, Fast HEVC inter prediction algorithm based on spatio-temporal block information, in: Proc. 2017IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, 2017.
11. X.G. Liu, Y.B. Liu, P.C. Wang, C.-F. Lai, H.-C. Chao, An Adaptive Mode Decision Algorithm Based on Video Texture Characteristics for HEVC Intra Prediction, IEEE Transactions on Circuits and Systems for Video Technology 27(8)(2017)
12. Shen L, Zhang Z, Liu Z. Effective CU size decision for HEVC intracoding[J]. IEEE Transactions on Image Processing, 2014, 23(10): 4232-4241.
13. Gweon R, Lee Y L. Early termination of CU encoding to reduce HEVC complexity[J]. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2012, 95(7): 1215-1218.
14. Choi K, Park H M, Jang E S. JCTVC-F092 Coding tree pruning based CU early termination[J]. Joint Collaborative Team on Video Coding (JCTVC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 Torino July, 2011: 1-11.
15. Yang J, Kim J, Won K, et al. Early SKIP detection for HEVC. JCT-VC of ITU-T SG16 WP3 and ISO[R]. IEC JTC1/SC29/WG11, Document JCTVC-G543, 2011.
16. Cho S, Kim M. Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(9): 1555-1564.
17. Zhang Y, Wang H, Li Z. Fast coding unit depth decision algorithm for interframe coding in HEVC[C]//2013 Data Compression Conference. IEEE, 2013: 53-62.
18. Zhao L, Fan X, Ma S, et al. Fast intra-encoding algorithm for high efficiency video coding[J]. Signal Processing: Image Communication, 2014, 29(9): 935-944
19. Khan M U K, Shafique M, Henkel J. An adaptive complexity reduction scheme with fast prediction unit decision for HEVC intra encoding[C]//2013 IEEE International Conference on Image Processing. IEEE, 2013: 1578-1582.
20. Zhang T, Sun M T, Zhao D, et al. Fast intra-mode and CU size decision for HEVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 27(8): 1714-1726.
21. Hu N, Yang E H. Fast mode selection for HEVC intra-frame coding with entropy coding refinement based on a transparent composite model[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25(9): 1521-1532.
22. Liu D, Liu X, Li Y. Fast CU size decisions for HEVC intra frame coding based on support vector machines[C]//2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech). IEEE, 2016: 594-597.
23. Jamali M, Coulombe S. Fast HEVC intra mode decision based on RDO cost prediction[J]. IEEE transactions on broadcasting, 2018, 65(1): 109-122.
24. Yu X, Liu Z, Liu J, et al. VLSI friendly fast CU/PU mode decision for HEVC intra encoding: Leveraging convolution neural network[C]//2015 IEEE International Conference on Image Processing (ICIP). IEEE, 2015: 1285-1289.
25. Li T, Xu M, Deng X. A deep convolutional neural network approach for complexity reduction on intra-mode HEVC[C]//2017 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2017: 1255-1260.
26. Kim K, Ro W W. Fast CU depth decision for HEVC using neural networks[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 29(5): 1462-1473.
27. Xu M, Li T, Wang Z, et al. Reducing complexity of HEVC: A deep learning approach[J]. IEEE Transactions on Image Processing, 2018, 27(10): 5044-5059.
28. Kuanar S, Rao K R, Bilas M, et al. Adaptive CU mode selection in HEVC intra prediction: A deep learning approach[J]. Circuits, systems, and signal processing, 2019, 38: 5081-5102.
29. Li H, Wei G, Wang T, et al. Reducing Video Coding Complexity Based on CNN-CBAM in HEVC[J]. Applied Sciences, 2023, 13(18): 10135.
30. Qin L M, Zhu Z J, Bai Y Q, et al. A Complexity-Reducing HEVC Intra-Mode Method Based on VGGNet[J]. Journal of Computers, 2022, 33(4): 57-67
31. Fan J, Song L. Fast Intra-frame Prediction Algorithm for HEVC Based on Neural Networks and Adaptive Threshold[C]//Proceedings of the 2022 6th International Conference on Video and Image Processing. 2022: 127-134.
32. Wang T, Wei G, Li H, et al. A Method to Reduce the Intra-Frame Prediction Complexity of HEVC Based on D-CNN[J]. Electronics, 2023, 12(9): 2091.
33. Feng A, Gao C, Li L, et al. Cnn-based depth map prediction for fast block partitioning in hevc intra coding[C]//2021 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2021: 1-6.
34. Hari P, Jadhav V, Rao B K N S. CTU Partition for Intra-Mode HEVC using Convolutional Neural Network[C]//2022 IEEE International Symposium on Smart Electronic Systems (iSES). IEEE, 2022: 548-551.
35. Lorkiewicz M, Stankiewicz O, Domanski M, et al. Fast Selection of INTRA CTU Partitioning in HEVC Encoders using Artificial Neural Networks[C]//2021 Signal Processing Symposium (SPSympo). IEEE, 2021: 177-182.
36. Ren W, Su J, Sun C, et al. An IBP-CNN based fast block partition for intra prediction[C]//2019 Picture Coding Symposium (PCS). IEEE, 2019: 1-5.
37. Feng Z, Liu P, Jia K, et al. HEVC fast intra coding based CTU depth range prediction[C]//2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC). IEEE, 2018: 551-555.
38. Li Y, Li L, Fang Y, et al. Bagged tree and ResNet-based joint end-to-end fast CTU partition decision algorithm for video intra coding[J]. Electronics, 2022, 11(8): 1264.
39. Imen W, Amna M, Fatma B, et al. Fast HEVC intra-CU decision partition algorithm with modified LeNet-5 and AlexNet[J]. Signal, Image and Video Processing, 2022, 16(7): 1811-1819.
40. Yao C, Xu C, Liu M. RDNet: Rate–Distortion-Based Coding Unit Partition Network for Intra-Prediction[J]. Electronics, 2022, 11(6): 916.
41. 王晓琰, 殷建芳, 王晓峰, 等. 关于连续出版会议论文著录格式的探讨[J]. 学报编辑论丛, 2019: 162-165.

**顺序编码制参考文献表的著录格式和示例**

**期刊论文示例:** [序号] 作者. 文题[J]. 期刊名, 年, 卷(期): 起-止页码.

1. WU D, YAN J, WANG H, et al. Social attribute aware incentive mechanism for device-to-device video distribution[J]. IEEE Transactions on Multimedia, 2017, 19(8): 1908-1920.
2. HEWITT J A. Technical services in 1983[J]. Library Resource Services, 1984, 28(3): 205-218.
3. 李炳穆. 理想的图书馆员和信息专家的素质与形象[J]. 图书情报工作, 2000(2): 5-8.

**会议论文示例:** [序号] 作者. 文题[C]. 会议名, 会议地, 会议年: 起-止页码.

1. BERGAMASCO F, ALBARELLI A, COSMO L, et al. Adopting an unconstrained ray model in light-field cameras for 3D shape reconstruction[C]. IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 3003-3012.
2. 中国力学学会. 第3届全国实验流体力学学术会议论文集[C]. 天津: 南开大学出版社, 1990: 10-12.

**普通图书示例:**  [序号] 作者. 书名[M]. 译者. 版本. 出版地: 出版者, 出版年: 起-止页码.

1. 竺可桢. 物理学[M]. 北京: 科学出版社, 1973: 56-60.
2. 汪继祥. 作者编辑手册[M]. 北京: 科学出版社, 2004.
3. 汪昂. (增补)本草备要[M]. 石印本. 上海: 同文书局, 1912: 31-32 .
4. ROOD H J. Logic and structured design for computer programmers[M]. 3rd ed. Watertown: Brooks/Cole Thomson Learning, 2001: 105-116.
5. CRAWFPRD W, GORMAN M. Future libraries: dreams, madness, & reality[M]. Chicago: American Library Association, 1995: 20-21.
6. 罗杰斯. 西方文明史：问题与源头[M]. 潘惠霞, 魏婧, 杨艳, 等译. 2版. 大连: 东北财经大学出版社, 2011: 15-16.

**学位论文示例:** [序号] 作者. 文题[D]. 授位单位所在地: 授位单位, 授位年: 起-止页码.

1. 李娜芬. 障碍环境中Swarm突现计算模型研究及行为控制[D]. 重庆: 重庆邮电大学, 2013: 11-47.

**报纸文章示例:** [序号] 作者. 文题[N]. 报纸名, 出版日期(版面数).

1. 顾春. 牢牢把握稳中求进的总基调[N]. 人民日报, 2012-03-31(3).
2. 丁文祥. 数字革命与竞争国际化[N]. 中国青年报, 2000-11-20(15).
3. 张田勤. 罪犯DNA库与生命伦理学计划[N]. 大众科技报, 2000-11-12(7).

**报告示例:** [序号] 作者. 文题[R]. 出版地: 出版者, 出版年.

1. 冯西桥. 核反应堆压力容器的LBB分析[R]. 北京: 清华大学核能技术设计研究院, 1997.
2. U. S. Department of Transportation Federal Highway Administration. Guidelines for handling excavated acid-producing material: PB 91-194001[R]. Springfield: U.S. Department of Commerce National Information Service, 1990.
3. World Health Organization. Factors regulating the immune response: report of WHO Scientific Group[R]. Geneva: WHO, 1970.

**专利示例:** [序号] 发明人. 专利名: 专利号[P]. 授权日期.

1. 张凯军. 轨道火车及高速轨道火车紧急安全制动辅助装置: 201220158825. 2[P]. 2012-04-05.

**标准示例:** [序号] 发布单位. 标准名: 标准号[S]. 出版地: 出版者, 出版年: 起-止页码.

1. 国家技术监督局. 国际单位制及其应用: GB 3100-1993[S]. 北京: 中国标准出版社, 1994: 3-6.
2. 国家技术监督局. 有关量、单位和符号的一般规则: GB/T 3101-1993[S]. 北京: 中国标准出版社, 1994: 13-20.
3. 中国国家标准化管理委员会. 信息与文献 参考文献著录规则: GB/T 7714.1-2015[S]. 北京: 中国标准出版社, 2015: 1-18.

**电子文献示例：**[序号] 作者. 文题[文献类型标识/文献载体标识]. 出版地: 出版者, 出版年: 起-止页码 (更新或修改日期) [引用日期]. 获取或访问路径. 数字对象唯一标识符.

1. CLERC M. Discrete particle swarm optimization: a fuzzy combinatorial box[EB/OL]. 2010-07-16, http://clere.maurice.free.fr/pso/Fuzzy\_Discrere\_PSO/Fuzzy\_DPSO.htm.
2. Dublin core metadata elsment set:version 1.1[EB/OL].(2012-06-14)[2014-06-11].http://dublincore.org/docu-ments/dces/.
3. 河北绿洲生态环境科技有限公司. 一种荒漠化地区生态植被综合培育种植方法: 01129210.5[P/OL]. 2001-10-24[2002-05-28]. http://211.152.9.47/sipoasp/zlijs/hyjs-yx- new.asp? recid= 01129210.5&leixin.
4. 陈建军. 从数字地球到智慧地球[J/OL]. 国图资源导刊, 2010, 7(10): 93[2013-03-20].http://d.g.wanfangdata.com.cn/Periodical\_hunandz201010038.aspx.DOI:10.3969/j.issn.16725603.2010.10.038.
5. WALLS S C, BARICHIVICH W J, BROWN M E. Deluge and declines:the impact of precipitation extremes on amphibians in a changing climate[J/OL]. Biology, 2013, 2(1): 399-418[2013-11-04]. http://www.md-pi.com/2079-7737/2/1/399. DOI:10.3390/biology2010399.
6. 吴玉芳. 面向中文信息处理的现代汉语并列结构研究[D/OL]北京: 北京大学, 2003[2013-10-14].http://thesis.lib.pku.edu.cn/dlib/List.asp?lang= gb&type= Reader&DocGroupID= 4& DocID= 6328.

**著者-出版年制参考文献表的著录格式和示例**

顾春, 2012. 牢牢把握稳中求进的总基调[N]. 人民日报, 03-31 (3).

河北绿洲生态环境科技有限公司, 2001. 一种荒漠化地区生态植被综合培育种植方法: 01129210.5[P/OL]. http:// 211.152.9.47/sipoasp/zlijs/hyjs-yx-new.asp? recid= 01129210.5&leixin.

李炳穆, 2000. 理想的图书馆员和信息专家的素质与形象[J]. 图书情报工作, (2): 5-8.

尼葛洛庞帝, 1996.数字化生存[M]. 胡泳, 范海燕, 译. 海口: 海南出版社.

全国信息与文献标准化技术委员会, 2007. 学位论文编写规则: GB/T 7713.1-2006[S]. 北京: 中国标准出版社, 17-20.

全国信息与文献标准化技术委员会, 2015. 信息与文献 参考文献著录规则: GB/T 7714-2015[S]. 北京: 中国标准出版社, 1-18.

汪冰, 1997. 电子图书馆理论与实践研究[M]. 北京: 北京图书馆出版社: 16.

BAERGAMASCO F, ALBARELLI A, COSMO L, et al. 2015. Adopting an unconstrained ray model in light-field cameras for 3D shape reconstruction[C]. IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 3003-3012.

DOWLER L, 1995. The research university’s dilemma:resource sharing and research in a transinstitutional environment[J]. Journal of Library Administration, 21(1/2): 5-26.

KENNEDY W J, GARRISON R E, 1975a. Morphology and genesis of nodular chalks and hardgrounds in the Upper Cretaceous of southern England[J]. Sedimentology, 22: 311-386.

KENNEDY W J, GARRISON R E, 1975b. Morphology and genesis of nodular phosphates in the cenomanian of South-east England[J]. Lathaia, 8: 339-360.

附录A 各学院中英文名称对照表

|  |  |  |
| --- | --- | --- |
| **序号** | **中文名称** | **英文名称** |
| 01 | 通信与信息工程学院 | School of Communications and Information Engineering |
| 02 | 计算机科学与技术学院/人工智能学院 | School of Computer Science and Technology / School of Artificial Intelligence |
| 03 | 自动化学院/工业物联网学院 | School of Automation / School of Industrial Internet of Things |
| 04 | 先进制造工程学院 | School of Advanced Manufacturing Engineering |
| 05 | 光电工程学院/重庆国际半导体学院 | School of Optoelectronic Engineering / Chongqing International Semiconductor College |
| … | …… | …… |

作者简介

## 基本情况

张某某，男，重庆人，1993年8月出生，重庆邮电大学XX学院XX专业2018级博士研究生。

## 教育和工作经历

2010.08～2014.06 重庆邮电大学光电工程学院，本科，专业：电子科学与技术

2014.08～2015.06 华为，技术研究工程师

2015.08～2018.06 重庆邮电大学光电工程学院，硕士研究生，专业：电子科学与技术

2018.08～2022.06 重庆邮电大学通信与信息工程学院，博士研究生，专业：信息与通信工程

## 攻读学位期间的研究成果

### 发表的学术论文和著作

1. **ZHANG M** , XX, XX. XXXX[J]. International Journal of Machine Learning and Cybernetics, 2021, 12(9): 2543–2557. (SCI期刊)
2. XX, **张某某**, XX. XXXX [J]. 计算机学报, 2022. (已录用)
3. XX, XX, XX,**张某某**等. XXXX[M]. 科学出版社, 2021. (专著)

### 申请（授权）专利

1. **张某某**, XXX, XXX等. 专利名称: 专利号[P]. 授权日期.

### 参与的科研项目及获奖

格式：XXX项目, 项目名称, 起止时间, 完成情况, 作者贡献.

1. 国家自然科学基金重点项目, XXXX (No.000000), 2017.01-2020.12, 参与.
2. 重庆邮电大学博士研究生人才培养项目, XXXX (No.000000), 主持.
3. XXX, **张某某**, XXX等. 科研项目名称. 重庆市科技进步三等奖, 获奖日期.

致 谢

感谢老师、同学们的关心、支持和帮助！