# 摘 要

VP9是由Google主导的新一代开源视频编码标准，在保留了传统的视频编码框架的基础上，又采纳了一系列新技术，让它在同等画质下码率较上一代VP8可降低50％，和HEVC相差无几。但是编码效率的提升的同时，编码复杂度也大大增加，致使使用VP9编码器进行视频编码会耗费大量的时间，因此尽可能地减少编码时间，实现高清分辨率下的实时传输将是一个必然趋势。针对该问题，本文研究了VP9视频编码标准的帧间预测算法，期望降低其编码的复杂度，减少编码时间，使该标准更适应实时性要求较高的业务。

本文针对VP9视频编码的帧间预测过程，提出了基于划分模式映射的帧间预测快速算法。主要是根据当前块的时空相邻块的信息确定模式候选列表，然后通过代价估计减少划分模式或跳过一些不必要的划分模式。该算法不仅能保证编码性能不会下降太多，而且大大降低了编码复杂度。

关键词：VP9 帧间预测编码 模式选择 快速算法

# **ABSTRACT**

VP9 is the next generation video coding standard that developed by Google. It still utilizes the traditional hybrid video coding framework and also adopts a series of new video coding techniques in recent years. As a result, the coding performance of VP9 has a significant increment relative to the previous video coding standard VP8, and it achieves an average bit rate reduction of 50% compared with the VP8 coder while keeping the same perceptual video quality,which is almost the same as HEVC. However, the improvement of compression ratio sacrifices the coding speed at the same time,resulting in using VP9 encoder of video coding will spend a lot of time. Consequencely, it’s an inevitable trend to realize the real-time transmission of high resolution and reduce the encoding time. Based on the problem, this thesis mainly focuses on the fast algorithms for inter prediction with expecting to reduce the coding time and thus makes VP9 standard more suitable for real time services.

Based on the key technologies of inter mode prediction in VP9, a fast inter prediction algorithm based on partition mode mapping is proposed for VP9. The mode candidate list is determined according to temporal and spatial blocks of the current block.,and then reduce partition modes or skip some unnecessary partition modes by rdcost estimation. Not only can the algorithm guarantee the encoding performance will not drop too much, but also greatly reduce the encoding complexity.

**Key words: VP9 Inter prediction coding Mode selection Fast algorithm**

# 目 录

**摘 要** i

**ABSTRACT** iii

**目 录** i

**第一章 绪 论** 1

1.1论文研究背景及意义 1

1.2论文的主要内容 2

1.3论文的章节安排 3

**第二章 VP9视频编码技术** 5

2.1 VP9视频编码器特点 5

2.2 VP9视频编码器框架 6

2.3 VP9编码结构 8

2.3.1 编码单元CU 8

2.3.2 预测单元PU 9

2.3.3变换单元TU 9

2.4 VP9帧间预测编码技术 10

2.4.1 帧间预测模式选择 10

2.4.2 帧间划分模式选择 15

2.4.3 VP9帧间预测编码模式选择部分整体框架 17

2.5 已有帧间快速算法研究 18

**第三章 基于二维坐标系模式点映射的VP9帧间快速模式算法** 21

3.1 概述 21

3.2 帧间划分模式点映射 21

3.2.1 划分候选模式点映射 22

3.2.2 最佳预测划分模式点映射 28

3.3 基于二维坐标系模式点映射生成的划分模式候选集 30

3.4基于代价估计的早期停止优化方案 31

3.5 整体算法实现 33

3.6实验结果与分析 34

3.7 本章小结 35

**第四章 总结与展望** 37

**致 谢** 39

**参考文献** 41

# 第一章 绪 论

## 1.1论文研究背景及意义

数字视频技术在通信领域获得了日益广泛的应用，特别是20世纪90年代以来，随着Internet和移动通信的迅猛发展，视频信息在Internet和移动网络中的处理和传输技术成为了当前全球信息化的热点。在互联网中，与视频内容相关的流量已经占据了整个网络流量的80%以上。[1]高质量的视频内容势必需要更大的视频码流，占据更多的资源和网络流量。众所周知，视频信息具有一系列优点，如直观性、确切性、高效性、广泛性等等。但是视频信息的信息量太大，要使视频得到有效的应用，必须在保证质量的情况下解决视频压缩问题。为了实现这个目标，成立了两个视频编码标准的国际组织：ITU-T/VCEG和ISO-IEC/MPEG。目前，已有的视频编码标准包括MPEG提出的MPEG-1、MPEG-2、MPEG-4，VCEG提出的H.261、H.263以及有二者共同提出的H.264/AVC，此外还有H.265/HEVC、VP8、VP9等。

VP9是由谷歌推出的免费视频编码标准，其目标在同等画质下码率较上一代VP8可降低50%，编码效率超过H.265/HEVC，并支持更大的编码块、更多的色彩空间。[2]在2014年年初的CES上绚丽演示了由VP9编码技术的4K高清视频流媒体服务，让众多厂商看到了VP9技术的优势，所以除了谷歌之外，英特尔、英伟达、飞利浦、高通、三星、Realtek 、ARM、LG等行业巨头以及其他众多企业都已经正式加入VP9的支持行列。

随着数字视频的分辨率不断提高，目前人们所使用的平板电脑、智能手机等电子设备已经逐步将1080P的高清视频格式作为标配，在某些场合里，甚至有超高清的视频格式（如4Kx2K）的出现。[3]在VP9中，考虑到现在的视频越来越多是高清视频，所以VP9进行了技术层面的革新。技术层面的革新让VP9在压缩效率上超越了前代的VP8和H.264，自称可以降低30%~50%的码率即可表现相同的画质，能以更低的带宽传输更清晰的画面，和HEVC相差无几。根据谷歌的测试，用250MHz以下的频率便可以实现2160p数字电影画质的硬件解码，已具备高效编码的所有特点。

在软件支持方面，Chrome浏览器从2013年2月第25版开始加入了用户自开启VP9支持的功能，到2013年9月的第29版起默认开启VP9支持。内核相同的360、搜狗、腾讯、遨游等浏览器现在虽然部分因为没有更新内核而暂时未能支持VP9，但一旦完成升级，VP9将直接登陆浏览器市场份额近4成的用户，几十亿浏览器将可以直接播放VP9视频。下一版的Android操作系统也会内置对VP9编码的支持，影响整个智能手机和平板电脑市场。WebM平台的数字版权保护和发布传输已经形成了建议性的可操作协议方案，作为标准可以直接应用。同时，谷歌也发布了VP9的硬件编码和硬件解码器供硬件商整合成硬件加速产品，使VP9的应用不限于互联网行业。

谷歌对互联网持开放性(Open)战略，以开放的软件平台标准——HTML5、开放的网络客户端——Chrome浏览器、开放的操作系统——Android和Chrome OS、开放的多媒体格式—WebP、WebM、VP8、Vorbis、Opus，形成了一整套开放的互联网技术，而新加入的VP9同样是以开放赢得多媒体领域的核心技术之一。VP9使用自由的BSD版权协议，可以直接用于包括商用在内的各种用途，也可以更改代码或整合创新，无须缴纳任何版权许可费。对于可能出现但尚未发现的技术专利问题，谷歌也作出开放的承诺。对比其他要支出成百上千万真金白银许可费的专利编码，VP9的综合优势不言而喻。

实践表明，VP9的编码效率在H.264之上，但是在HEVC之下。VP9编码时间大约是x264的130倍，HEVC编码时间大约是VP9的7倍。而且，同等质量的前提下，x264相对于VP9节约了8.4%的码率。所以，在与H.264和H.265竞争的过程中，VP9的压力依然很大，在实际应用中还会有所限制，这就需要对其进行优化，VP9原程序中帧间预测时间约占总编码时间的90%，在原有基础上，改善其帧间编码模式预测部分的算法，旨在提高运算速度，加大精度。虽然现有的快速指令已经使其速度大大加快了，但对于实时性的要求还是有所欠缺。如果优化一旦完成，VP9的运算时间将会进一步减少，其实时性也将会更好，应用范围也会因此变得更加广泛。本文主要针对编码中耗时较多的帧间编码技术提出改进算法，做到在不影响压缩率和压缩质量的情况下提高编码性能，提高VP9的实用性。

## 1.2论文的主要内容

人们日益增长的对高清视频的需求促使高压缩比的视频压缩标准的发展，VP9就是在这个背景之下发展起来的，但是编码效率提升的同时，编码复杂度也大大增加，致使使用VP9编码器进行视频编码会耗费大量的时间，因此尽可能地减少编码时间，实现高清分辨率下的实时传输将是一个必然趋势。VP9原程序中帧间预测时间约占总编码时间的90%，减少帧间预测模式选择的时间将是一个关键点。

视频序列中的两帧相邻的图像，后一帧图像与前一帧图像之间有较大的相关性，这反映为时间冗余。同一帧同一景物表面上采样点的颜色之间通常存在着空间相关性，相邻各点的取值往往相近或者相同，这就是空间冗余。因为视频中存在极强的相关性，也就是存在着极大的冗余度。冗余信息造成比特浪费，消除这些冗余信息可以节约码字。视频压缩的基本过程是先通过帧间编码去除时域和空域冗余冗余信息，再使用熵编码对残差进行压缩，得到视频的编码比特流进行传输。

在标准的VP9中，帧间编码单元采用四叉树循环划分，每个CU都要尝试4种划分模式，复杂度非常高，是实时编码的最大障碍之一。本文先对VP9的关键性技术进行研究，通过深层学习和统计分析来寻求降低编码复杂度的优化算法。本文主要研究了帧间预测过程中两个比较关键的过程：划分模式和帧间预测模式选择。通过对已有的帧间快速预测模式算法的研究，针对由于四叉树循环划分的候选候选划分模式众多导致模式选择过程复杂度极高的问题，本文提出了一种基于二维坐标系模式点映射的VP9帧间划分模式快速算法，以加速模式选择过程。

首先，本文主要介绍了帧间预测中块划分和帧间预测模式选择的详细过程，然后分析了原本算法的时间复杂度，研究了一些已有的块划分优化算法，并利用分析结果来提出优化算法。通过大量测试和实验，得到影响帧间预测模式选择的各个因素。本文先分析了影响块划分的主要因素：量化参数、CU大小以及视频自身的时空域特征。经过详细分析，提出了一种基于二维坐标系模式点映射的VP9帧间划分模式快速算法。算法的总体思路是利用视频自身时空域特征构建模式候选列表。接着，通过研究一些早期停止算法，利用代价估计等算法进一步简化模式候选列表 。从实验结果可以看出，该算法可以对不同大小的不同纹理的序列，在各个量化参数下都能大幅度减少编码时间，并且视频质量几乎没有损失。

## 1.3论文的章节安排

全文总共分为四章，各章的内容安排如下。

第一章：绪论，第一节主要介绍视频编码标准的历史、VP9产生的背景、VP9的综合优势和现有的不足之处，以及分析研究帧间快速算法的必要性。第二节介绍论文的研究内容，第三节介绍论文的章节安排。

第二章：先是介绍了VP9系列视频编码的发展历程和VP9视频编码器的特点，重点介绍了VP9帧间预测编码的过程，接着介绍了对已有VP9帧间快速算法研究。

第三章：是针对块划分模式的帧间快速算法。第一节简要分析了研究帧间块划分优化算法的必要性以及研究过程。第二、三节针对预测块的划分过程提出了基于二维坐标系模式点映射的快速算法，第四节针对该算法提出了早期停止的优化算法。第五节是该算法的具体实现以及对算法性能做的测试和分析。

# 第二章 VP9视频编码技术

VP9是一个由Google开发的开放格式、无使用授权费的视频压缩标准，在同等画质下码率比H.264降低50%，并支持更大的编码块、更多的色彩空间。VP9与视频编码标准中的其他标准，如H.265组成的功能块基本一致，主要不同在于各个功能块的具体实现细节。VP9视频编码器由六个基本模块组成，分别为块划分、帧内预测、帧间预测、变换量化、熵编码以及环路滤波。

## 2.1 VP9视频编码器特点

随着数字视频的分辨率不断提高，目前人们所使用的平板电脑、智能手机等电子设备已经逐步将1080P的高清视频格式作为标配，在某些场合里，甚至有超高清的视频格式（如4Kx2K）的出现。考虑到现在的视频越来越多是高清视频，VP9对每帧画面采用新一代的宏区块划分方法，在VP8最大为16×16(单位为像素或色彩分量)宏区块的基础上，发展出了超级区块(Superblock)的设计。这种设计使画面元素能被更优化地复用，也使得编码压缩效率直接提升约一成以上。

VP9在帧内预测部分继承了VP8对4x4块的10个预测模式。在帧间预测部分对于每个16x16及更大的区块，VP9允许有1或2个运动向量。VP8和H.264同代编码的运动向量可以表现1/4个像素的位置变化，而VP9在支持1/4像素精度的同时，可适时提升一倍精度到1/8像素，能更有效地表现慢动作。宏区块的亮度、色度、透明度的变化在VP8中用最高6阶插值滤波器来表现。VP9提升精度到8阶插值滤波器，增加了两个分别应对锐利变化和平滑变化的值滤波器。

对于识别宏区块和预测运动向量后，仅用已知信息仍不能完全重建一致的色彩信息，用DCT(离散余弦变换)和DST(离散正弦变换)等矩阵变换和量化记录其残留的变化情况。VP9支持不同大小变换块包括32x32、16x16、8x8、4x4四种变换尺寸。并且VP9去掉了VP8中的WHT变换，加入了一种新的变换类型ADST(Asymmetric Discrete Sine Transform)。在帧内预测中，ADST可以更好地应对信号边缘的情况。VP9量化参数的取值范围比VP8大了一倍，VP8量化参数的取值范围为0~127，VP9量化参数的取值范围为0~255。

VP9允许每一帧使用大小不同的分辨率，在编码时会用8阶1/16像素精度的缩放滤镜对运动向量编码，在解码播放时还原成统一的大小。在用于直播时，这样的功能允许传输中不间断调整码率，一定程度上改善了可适性。VP9对编码后的16×16及以上的宏区块增加了分段(Segmentation)功能，宏区块带有一个段标识(Segment\_ID)来判断是否属于同一个段。段数据分散在各帧，上面记录对该段允许使用哪些参考帧、省略系数、量化因子、回环滤波强度等参数，能更好地避免冗余和应对不同的应用需要。

VP9对无损压缩和有损压缩都支持，并且支持多核多线程的并行，利用并行技术可以显著地提高编码速率。VP9中每帧的头部都有并行模式的标识。VP9也和HEVC一样引入了tile的概念，通过把一帧图像分为不同的tile，可以让每个tile独立编解码，这样可以更好地实现并行。

## 2.2 VP9视频编码器框架

如图2.1和图2.2所示，VP9和H.264/AVC、H.265/HEVC等其它传统的视频编码标准的编解码器的功能模块的基本一致，主要体现在各功能模块的实现细节上有所不同。为了满足现在视频日趋增加的实时性要求，VP9视频编码器采用了一系列相对应的自适应技术。让VP9适合互联网应用是Google的设计初衷。对于常见的互联网应用场景，VP9都有简易、实用的解决方案设计。但互联网稳定性差，容易阻塞或信号衰落，因此需要编码器需要有相应的自适应技术。为了提高编码效率VP9视频编码器做了很多创新，大大提高了编码器的复杂度，不过编码效率得到了极大的提高使得VP9视频编码器的编码效率足以和H.265/HEVC视频编码器相媲美。

如图2.1所示，VP9编码器采用的是混合编码框架，框架包括了帧内预测编码、帧间预测编码、变换、量化、熵编码和环路滤波等模块。输入的原始视频帧首先被划分为互不重叠的64x64的超级块，然后以超级块为基本单位被VP9编码器处理。每个超级块采用帧内预测或帧间预测进行编码，并且它们的编码过程都是参考重建后的像素来进行预测，得到预测像素值（图2.1中的P）。对于帧内预测编码，其预测值是由当前块周围的重建像素预测得到的。而对于帧间预测编码，其预测值P是由前面重建帧（如图2.1中的）中经过运动搜索和运动补偿得到的。为了提高预测精度，VP9采用了多参考帧技术，相对于H.265/HEVC中的32个参考帧，VP9中参考帧只有3个，但每个参考帧的性质都不相同。这样不仅减小了处理器的压力，也降低了编码复杂度，提升了编码速度。最重要的是多个参考帧可以保证预测过程的预测精度。



图2.1 VP9视频编码器框图



图2.2 VP9视频解码器框图

预测编码结束后将原始块的参考像素和预测块对应像素相减得到残差块，再对残差块进行变换编码，去除频域上的一些冗余信息。变换块经过变换量化后得到DCT系数，对DCT系数熵编码，得到编码比特。然后将解码所需的一些头信息（如帧头信息、预测模式、量化参数、跳过标志位等）和熵编码压缩后的信息组成要传输视频码流。

为了得到进一步预测所需要的参考帧，编码器需要具有重建帧的模块。重构图像的模块与解码器的功能几乎一致。重构过程即为残差信号经过反量化和反变换后与预测值进行求和得到未滤波的重建帧（如图2.2中的）。为了提高重建帧的质量，需要进行环路滤波以去除方块效应，滤波后的图像即为重建帧（如图2.2中的），重建帧将在后续预测编码中作为参考帧。

## 2.3 VP9编码结构

为了更好地应对待编码视频图像中可能出现的各种运动情况、各种变化情况的视频图像，VP9将帧处理单元定义了三部分：编码单元、预测单元和变换单元。三个单元相互独立，从而使得编码、预测和变化各个模块也变得更加灵活，因此VP9的编码结构能够更好地处理具有不同纹理特征和运动变化的视频序列。

### 2.3.1 编码单元CU

视频编码通常把一张画面细分成小矩形图块作为编码的基础单位，这就是宏区块(MacroBlocks)。近年来视频的分辨率越来越高，大尺寸的视频越来越成为主流视频源，大尺寸图像的一个比较突出的特点就是存在大面积的平坦区域。据此，VP9对每帧画面采用新一代的宏区块划分方法，在VP8最大为16×16(单位为像素或色彩分量)宏区块的基础上，发展出了超级宏块(Superblock)的设计。超级宏块最大可以为64x64，超级宏块内可以再细分。为了更灵活、更有效使地表现视频内容，VP9采用了树形结构单元，对亮度和色度分别进行树形编码。用四叉树的结构安排这些宏块，把每个超级宏块细分成4块32x32的小编码块，每个小编码块也可以再分成4块更小的16x16编码块，尺寸可以继续划分到最小4x4编码块的基础大小。这种设计使画面元素能被更优化地复用，也使得编码压缩效率直接提升约一成以上。

VP9的这种表示方法比VP8中的宏块划分方法更灵活，也有更多的优势：

1、编码块的大小可以大于传统的宏块大小(16×16)。对于平坦区域，使用较大的编码单元进行编码能减少使用的比特数，从而可以提升视频编码的压缩比从而提高编码效率。

2、通过这种循环四叉树分层结构更加有利于编码器选择出合适的编码块大小以及合适的最大划分深度，编码器的编码结构能够根据不同的应用需求、图片大小、以及图片内容获得最大程度的优化。



图2.3 编码块的循环四叉树分层结构

### 2.3.2 预测单元PU

预测单元规定了编码块的所有预测模式，所有预测处理和信息都在预测块中获得和保存，包括帧内预测模式、帧间预测参考帧索引号、帧间预测模式以及运动向量在内的所有与预测相关的信息。预测块有4种划分形式： 2N×2N、2N×N、N×2N、N×N。如图2.4所示。其中2N×N、N×2N不能继续往下划分，对称划分即NxN划分模式只适用于不小于8x8的编码块之中。



图2.4 VP9预测块的划分方式

### 2.3.3变换单元TU

变换单元是完成变换和量化的基本单元，VP9的变换块可支持大小为32×32至4×4的编码变换，其大小也是根据实际情况变化的，实现了编码结构的灵活独立。图2.5给出了当前编码块为64x64下各个预测块中变换块的分割结构的划分情况。根据预测残差的局部变化特性，变换块可以自适应地选择最优尺寸和模式。尺寸较大的变换块时能够更好地集中能量，尺寸较小的变换块时能够更好地保存视频图像中的细节。变换块的这种独立的划分选择方式，让变换后的残差信息得到充分地压缩，进一步提高了压缩比。



图2.5 变换块的划分示意图

## 2.4 VP9帧间预测编码技术

视频序列中存在极强的相关性，也就是存在着极大的冗余度。冗余信息造成比特浪费，消除这些冗余信息可以节约码字。预测编码是利用信号之间存在的相关性，利用前面一个或多个信号预测下一个信号进行，然后对原始值和预测值的差进行编码。如果预测比较准确，误差就会接近于0。在一定精度条件的要求下，仅利用较少的比特表示和传输，就能达到压缩数据的目的。视频序列中两帧相邻的图像有较大的时间相关性；同一帧同一景物表面上采样点的颜色之间通常存在着空间相关性，相邻各点的取值往往相近或者相同。如今大部分视频编码采用帧内预测和帧间结合的预测方式，去除更多的视频空域和时域的冗余度，提高编码效率。

在本章前面部分主要介绍了VP9编码器的框架，编码结构，包括编码单元、预测单元、变换单元。本节将主要介绍VP9视频编码器中的帧间预测编码技术。

### 2.4.1 帧间预测模式选择

帧间预测编码技术主要是利用视频图像中已经编码的图像像素来预测当前帧的图像像素从而去除视频时空域上的冗余，进而实现压缩编码。

帧间预测中有两个关键技术：参考帧的选取以及运动向量的确定。参考帧主要用于复制像素到预测帧，为当前编码帧提供像素基础，然后根据运动向量构建出重建帧。

1、参考帧

VP9的帧间预测最多支持3个参考帧：Last Frame、Golden Frame和Altref Frame。

Last Frame 为当前帧在时间轴上的前一帧。

Golden Frame 是指时间上任意远的一帧，用来存储和重建编码视频背景内容的更新信息。Golden Frame的特点是质量较高，量化参数较小。

Altref Frame同Golden Frame相似，量化参数较小，质量较高，但该类型帧只能作为参考帧不能被显示。Altref Frame参考帧有两种，第一种Altref Frame是通过普通编码方式得到的，只是其QP较小，但质量高于普通帧。第二种Altref Frame较为特殊，它是由多个未编码的帧拼凑而成的，它的宏块可能来自不同的帧。

由于为了避免使用MPEG专利，VP9没有采用可以在前后双向参照的B帧，而使用单向参考帧，好像显得VP9不足；但VP9新引入了复合预测的方式，允许Altref Frame的存在，它不会被显示。换句话说，复合预测即变相地实现了双向B帧，同时又规避了B帧的专利。

Last Frame、Golden Frame和Altref Frame三种参考帧的位置关系如图2.6所示：



图2.6 VP9中参考帧之间的位置关系

2、帧间预测编码运动向量

帧间的PMV有五种，分别为NEARESTMV、NEARMV、NEWMV、NEARESTMV和SPLITMV，其中SPLITMV只适用于小于8x8的块，其他 MV适用于大于等于8x8的块。

每一种PMV，根据参考帧类型及参考帧数目的不同，可以继续分为五种情况：

当只选取一个参考帧时可以得到的组合类型有三种情况，分别是：{ Last Frame ，None}、{ Golden Frame ，None}或{ Altref Frame ，None}。当选取两个参考帧时，由于当存在第二参考帧时，参考帧类型一定是Altref Frame，并且由于两个参考帧的类型不能一样，所以可以得到的组合类型是{ Last Frame ，Altref Frame }或{ Golden Frame ，Altref Frame }。

下面主要介绍帧间PMV的获取方法：

1. NEARESTMV和NEARMV

确定参考帧后，当前编码块有8种空域候选块和1种时域候选块。编码块空域候选块的位置和顺序随当前块尺寸的不同而有所差异。图2.7给出了不同预测块的空域候选位置，图中每一个候选块的坐标均为与当前编码块的相对坐标。其中ABCDEFGH为当前块的空域候选块，再加上一个前一帧中对应位置的时域候选块，一共9个候选块，尺寸均为8x8。

①在选取当前编码块的PMV时，首先遍历候选块的参考帧，若候选块的参考帧类型至少有一个与当前块的参考帧类型相同，则候选块的MV会被放入参考帧列表，当有两个满足条件的MV，就早期停止不再遍历。如果遍历所有的候选块仍然没有找到两个满足条件的MV，则直接将前两个候选块的MV放入参考帧列表。

②在获取了候选列表之后，列表中的第一个运动向量就是NEARESTMV，第二个运动向量就是NEARMV。因为一共有五种参考帧组合方式，并且每一种参考帧组合都需要找到1个NEARESTMV和1个NEARMV，所以最终可能会得到5个NEARESTMV和5个NEARMV。

1. NEWMV

NEWMV是以NEARESTMV和NEARMV为基础进行运动搜索获取的运动向量，NEWMV并非预测运动向量，而是真实运动向量，其具体计算过程如下：

①从NEARESTMV和NEARMV中选取一个失真即SAD较小的MV，当作 

图2.7 空域候选块位置

NEWMV运动搜索的起点。

②找出搜索代价最小的MV即为NEWMV。

③NEWMV模式的PMV为NEARESTMV，其MVD不为0而是NEWMV与NEARESTMV的差值。

1. NEARESTMV

NEARESTMV计算比较简单，即其直接将预测运动向量PMV赋值为0，绝对误差均值也为0。

(4)SPLITMV



图2.8 8x8块的划分示意图

SPLITMV只适用于小于8x8的块， 8x8块划分方式如图2.8所示。

3、 帧间预测模式

VP9帧间预测模式主要由参考帧和运动向量确定，总共30种。

1. NEARESTMV:

{NEARESTMV, LAST FRAME, NONE}，

{NEARESTMV, ALTREF FRAME, NONE}，

{NEARESTMV, GOLDEN FRAME, NONE}，

{NEARESTMV, LAST FRAME, ALTREF FRAME}，

{NEARESTMV, GOLDEN FRAME, ALTREF FRAME}，

2) NEARMV

{NEARMV, LAST FRAME, NONE}，

{NEARMV, ALTREF FRAME, NONE}，

{NEARMV, GOLDEN FRAME, NONE}，

{NEARMV,LAST FRAME,ALTREF FRAME}，

{NEARMV, GOLDEN FRAME, ALTREF FRAME}，

3) NEWMV

{NEWMV, LAST FRAME, NONE}，

{NEWMV, ALTREF FRAME, NONE}，

{NEWMV, GOLDEN FRAME, NONE}，

{NEWMV, LAST FRAME, ALTREF FRAME}，

{NEWMV, GOLDEN FRAME, ALTREF FRAME}，

4) NEARESTMV

{NEARESTMV, LAST FRAME, NONE}，

{NEARESTMV, GOLDEN FRAME, NONE}，

{NEARESTMV, ALTREF FRAME, NONE}，

{NEARESTMV, LAST FRAME, ALTREF FRAME}，

{NEARESTMV, GOLDEN FRAME, ALTREF FRAME}.

5) SPLITMV

{SPLITMV,LAST FRAME,NONE}，

{SPLITMV,GOLDEN FRAME,NONE}，

{SPLITMV,ALTREF FRAME,NONE}，

{SPLITMV,LAST FRAME,ALTREF FRAME}，

{SPLITMV,GOLDEN FRAME,ALTREF FRAME}，

### 2.4.2 帧间划分模式选择

对于64x64的SuperBlock，VP9中提供了3种块划分策略，具体如下：

①四叉树循环划分：每个编码块遍历4种划分模式，即NONE、HORIZON、VERTICAL、SPLIT。通过树形遍历四种模式，根据率失真准则选出最佳的划分模式。CU块大小的上下限可以参考相邻SuperBlock的划分自适应设置；

②使用固定尺寸划分：根据速率参数确定的一个固定的CU大小，从而完成对当前的SuperBlock进行划分；

③使用上一帧的划分：使用上一帧对应位置SuperBlock的划分模式，对当前SuperBlock进行划分。

这三种划分中四叉树划分的复杂度最高，性能也是最好，在实现过程中根据所分配的复杂度通过一系列的速率参数控制其它两种快速划分模式的开关。下面重点介绍一下四叉树编码块的划分。

编码块一共有4种划分形式：2Nx2N（NONE）、2NxN（HORIZON）、Nx2N（VERTICAL）、NxN（SPLIT）。其中SPLIT中的每个不小于8x8的子块可以进一步进行四叉树划分，如图2.9所示。

对当前CU遍历4种划分模式，首先进行NONE，直接以当前CU为单位进行预测模式选择。接下来如果当前块的尺寸不小于8x8就进行SPLIT，每个子块再分别遍历四种划分模式，这是一个逐层递归的过程。然后做HORIZON和VERTICAL，每个长方形子块分别进行预测模式选择。最后比较4种划分模式下的RDCost，选择出RDCost最小的划分模式作为最佳划分模式，然后将相应的模式信息和残差系数进行打包，传给熵编码模块。



图2.9 四叉树编码划分

帧间部分变换块的尺寸选择主要的流程图如图2.10所示。



图2.10 变换块的尺寸选择流程图

### 2.4.3 VP9帧间预测编码模式选择部分整体框架

整个帧间预测编码模式选择部分以率失真函数作为目标函数，通过计算每种模式下的RDCost，选取出RDCost最小的编码模式作为编码块最终的编码模式。整个模式选择部分的整体代码框架如图2.11所示：

图2.11 模式选择部分代码框架

①首先判断当前编码块的尺寸是否大于8x8，如果当前编码块的尺寸小于等于8x8，则遍历候选块从而获得NEARESTMV、NEARMV，通过运动搜索获取NEWMV，并将NEARESTMV置为0，最后再计算SPLITMV。然后通过运动补偿，获得预测像素值，再进行变换、量化、反量化、反变换，选择出变换块的大小，进而计算出RDCost，进而跳转到③；

②如果当前编码块的尺寸大于8x8则进入该分支，对编码块遍历候选块从而获得NEARESTMV、NEARMV。再利用NEARESTMV、NEARMV找到运动搜索的起点，然后进行运动搜索得到NEWMV。将NEARESTMV置为0。在各个运动向量模式下通过运动补偿，获得预测像素值，再进行变换、量化、反量化、反变换，选择出变换块的大小，进而计算出RDCost；

③通过对比NEARESTMV、NEARMV、SPLITMV、NEWMV四种运动向量下的RDCost，从中选出RDCost最小的模式，此模式即为当前编码块选取的最终编码模式。

## 2.5 已有帧间快速算法研究

为了减少VP9编码器的整体计算复杂度，维持VP9的编码性能，目前已经有很多帧间预测快速算法，大多数都是基于块划分的快速算法。

随着视频分辨率的不断提高，大分辨率的视频在生活中越来越常见。视频序列中图像的内容大多数是平坦的，对于平坦的块采用大块划分可以减少传输比特，而细节复杂的区域采用小块更能保留图像特征。因此块划分模式对整体的编码性能有很大的影响。

已有的基于块划分的快速算法主要分为以下五类：

1. 基于简化率失真代价函数的快速算法

简化率失真代价函数的快速算法主要是通过分析率失真代价函数的特点，采用新的算法简化率失真代价函数的计算或者代替率失真代价函数，达到减小降低帧间编码过程复杂度的目的。文献[15]中，作者通过预测比特开销和失真来进行代价估计，利用预测RDCost来提前终止块划分算法，得到了比较理想的效果。

1. 早期终止的快速算法

文献[16]中，作者通过对具有相同块划分模式的宏块进行分析，找到一些共同的特点，根据分析结果设置一些门限，如果在块划分的过程中早期终止的条件成立，则提前终止块划分算法。实验结果表明该算法可以明显的减少编码时间而且维持编码性能仅有微小的变化。

1. 基于当前块的纹理分析的快速算法

图像中纹理简单均匀的部分更倾向于较大的CU块划分，而纹理比较复杂的部分则倾向于较小的CU块的划分。纹理是一个编码块比较直观而且重要的特征，文献[17]通过分析当前块像素值的平均平方误差MSD（Mean Square Deviation）和时空域相邻块的的信息来估计当前块的纹理复杂度，最后根据纹理复杂度减少候选划分模式。

1. 基于变换域的快速算法

这类算法的主要思想是利用变换系数的一些特征减少候选划分模式，从而减少计算复杂度。文献[18]中主要利用DCT变换系数以及残差数据通过一个透明复合模型来分别估算当前块的空间复杂度和时间复杂度，然后进行早期终止决策完成快速模式划分。

1. 基于复杂度分析的快速算法

复杂度是衡量一个编码器性能的重要指标，为了使编码器满足复杂度的要求，一些基于复杂度受限的快速算法应运而生。文献[19]中，利用每个GOP所分配的复杂度，为每一时间层的的图像分配不同的复杂度因子，利用复杂度因子来调整候选划分模式集的大小。文献[20]中利用服务器的硬件条件以及时空域相邻块的划分信息为当前编码块分配块级复杂度，取得了良好的效果。

# 第三章 基于二维坐标系模式点映射的VP9帧间快速模式算法

## 3.1 概述

由第二章得知VP9技术革新之后拥有更多的块划分方式和预测模式，这让帧间编码的压缩效率急剧提高的同时也造成编码复杂度的大幅度增加。本章从块划分角度进行VP9帧间算法的优化，提出基于二维坐标系模式点映射的VP9帧间快速模式算法，旨在维持编码器整体性能几乎不变的情况下降低编码复杂度。

在自然的视频序列中有很多的同质或固定的区域，所以在VP9视频编码中，时间或空间相邻编码块之间往往具有相同或相似的编码划分模式。因此，通过优先检查这些类似的编码划分模式，如果满足一些早期终止条件，模式决策过程就可以停止，这样可以大大减少编码复杂度。基于这一观察，本文提出了一个以二维坐标系模式点映射为基础的算法，利用该算法构建一个模式候选列表用于VP9的帧间快速划分模式决策。

在该算法中首先根据最小编码块4x4的归一化运动向量（NMV）将不同大小的编码块所有可能的编码块划分模式映射到同一个二维坐标系下，得到各个编码块划分模式的映射点。然后对于一个当前的编码块CU，首先将其时域相邻块的划分模式与空域相邻块的划分模式映射为二维坐标系中的模式点，根据时空域相邻块的划分情况推导出一个最佳中心点。最后根据最佳中心点与所有编码划分模式点间的距离，建立一个基于优先级的模式候选列表。最后的模式决策过程是按照模式候选列表中顺序依次遍历模式候选列表中的模式

基于最优停止理论，如果遍历有限数量的候选划分模式而且要保持编码器性能，这些被遍历的候选划分模式应具有最高的预测概率。因此在模式决策过程中，提出了基于最优搜索半径的早期停止策略、基于代价估计的早期停止策略用来优化模式候选列表。实验证明结合早期终止算法，帧间划分模式决策过程消耗的复杂度大大减少。

## 3.2 帧间划分模式点映射

本文通过研究不同编码划分模式的运动向量之间的一致性（MVS），提出了一种基于二维坐标系的模式点映射方法。在二维坐标系中利用模式点（MPS）代表不同的划分模式，用最佳预测点（SP）代表所预测的最优划分模式，然后用它们之间的欧氏距离代表不同划分模式之间的关联。

在基于二维坐标系的模式点映射方法中，提出了以下四个假设：

1、每个划分模式都能映射为二维坐标系中的一个点MP

2、在模式空间中，两个模式点MP之间的距离越小意味着对应两个划分模式之间具有更高的相关性。

3、对于每个CU，在模式空间存在有一个对应的真实最佳模式点SP，这个模式点不一定能与划分模式的映射点MP重合。

4、最优划分模式是所有候选模式中与真实最佳模式点SP距离最小的模式MP。

### 3.2.1 划分候选模式点映射

目前主流的视频编码标准帧间预测部分都采用了基于块的运动补偿技术，如图3.1所示。[21]其主要原理是为当前图像的每个像素块在之前已编码的图像中寻找一个最佳匹配块，这个过程就是运动估计（Motion Estimation，ME）。其中用于预测的图像称为参考图像，参考块到当前像素块的位移称为运动向量。

图3.1 运动补偿的原理图

通过测试不同块大小、不同划分模式对应的4x4子块的运动向量MV之间的一致性，测试结果如下。

图3.2 划分为NONE的64x64 CU块的4x4 sub\_CU的MV分布

图3.3 划分为SPLIT的64x64 CU块的4x4 sub\_CU的MV分布

图3.4 划分为NONE的32x32 CU块的4x4 sub\_CU的MV分布

图3.5 划分为SPLIT的32x32 CU块的4x4 sub\_CU的MV分布

图3.2和图3.4表明，对于一个划分模式为NONE的CU块，它的子块sub\_cu 4x4的运动向量MV之间具有更高的一致性。相反对于一个划分为SPLIT的CU块，它的子块sub\_cu 4x4的运动向量MV之间的一致性相对比较差，如图3.3和图3.5所示。由此我们可以根据MV的一致性建立二维坐标系，然后求出各个划分模式映射在坐标系中的模式点（MPs）。

为了计算划分模式映射在二维坐标系中的位置，[22]我们引入4x4块归一化运动向量（NMV）的概念，计算公式如3-1所示：



其中x，y分别是4x4子块在编码块CU中水平方向和垂直方向上的位置；是在编码块CU中（x，y）位置的4x4子块的运动向量；pflag是预测方式的一个标志，pflag = 0表示预测方式为单向预测（前向预测或后向预测），pflag = 1表示预测方式为复合预测；c、r分别指当前帧的编号和参考帧的编号（、为复合预测时两个参考帧的编号）。

如何表示NMV的一致性是一个要解决的问题。衡量一个样本波动大小可以使用方差，方差越大，样本数据的波动就越大，样本一致性也就越差。因此可以用方差表示一致性。我们定义NMV的方差为var，通过计算当前编码块每一行与每一列的方差来求得各个4x4子块之间NMV的一致性，公式如3-2所示：



式3-2中（x0，y0）表示当前编码块左上角4x4子块的位置，N表示当前编码块一行或一列4x4子块的个数，如果当前编码块的大小为64x64，则N = 16，如果当前编码块大小为32x32，则N = 8；var表示当前编码块以4x4子块为单位第k行N个4x4子块的方差，其中表示两个向量之间的欧氏距离。

最终划分模式点在二维坐标系中的位置为，为水平分量，为垂直分量，其定义如下：



通过分析不同序列不同划分模式点的位置，发现在CU为64x64时，NONE、HORIZON、VERTICAL、SPLIT四种划分模式可以形成一个近似方形的平面，类似的情况在sub\_CU中同样适用，不同的只是方形的大小不一样。为了验证此现象，以视频序列Kimono\_1920x1080为例，用式3-3求出不同深度下不同划分模式点的位置坐标，经过统计分析将2Nx2N划分模式点平移到坐标原点，各划分模式的相对坐标如表3.1所示。

表3.1视频序列Kimono\_1920x1080中各划分模式的相对坐标

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| CU depth | | 0 | 1 | 2 | 3 |
| CU size | | 64x64 | 32x32 | 16x16 | 8x8 |
| MPs | 2Nx2N | （0,0） | （55,58） | （98,95） | （113,114） |
| 2NxN | （0,56） | （56,96） | （98,115） | （118,131） |
| Nx2N | （60,0） | （93,55） | （109,99） | （129,112） |
| NxN | （55,58） | （98,95） | （113,114） | （128,133） |

在VP9的四叉树编码中，CU的深度（depth）可以代表当前编码块（CB）的大小，depth = 0表示CB = 64x64，depth = 1表示CB = 32x32。将各个模式点在二维坐标系上表示出来，发现不仅每一层四个划分模式形成一个类似方形的平面而且相邻层之间方形的对角线之间呈现出一种相似的比例关系，如图3.6所示，即R1/R2 = a，这里称a为比例因子。



图3.6 相邻层之间各划分模式的位置分布

为了更精准的确定每个划分模式点的坐标以及相邻层之间的比例因子a，本文对不同分辨率不同内容的视频（BasketballDrive\_1920x1080、BQTerrace\_1920x1080、ParkScene\_1920x1080、legend\_1280x720、wallstreet\_1280x720、LineWalk\_1280x720），用不同的量化参数Qp（38~50）在相同的测试环境下进行统计，最终得到6组不同的模式点映射的比例因子a。

经过测试发现大多数视频中，当depth = 3时，CU块划分为NxN对应的SPLIT模式点的坐标大多集中在（126,126）附近。为了从6组不同的模式点映射的比例因子中选出一个最有效的比例因子a，本文固定8x8 CU块的SPLIT模式点的坐标为（126,126），运用上述6种不同的比例因子进行性能测试。用峰值信噪比PSNR（Peak Signal to Noise Ratio）的下降值（dB）以及码率BR(BitRate)的增加值（%）来进行比较，表达式如式3-4、3-5所示：





上式中和是利用该算法测得的预测PSNR和预测Bitrate；和指的是利用VP9原代码测得的原始PSNR和Bitrate。因为所有的比例因子a保持几乎相同编码时间，所以在整个编码时间节省方面的性能没有给出。测试用了BasketballDrive\_1920x1080、BQTerrace\_1920x1080、ParkScene\_1920x1080、legend\_1280x720、wallstreet\_1280x720、LineWalk\_1280x720不同分辨率不同内容的序列在不同的QP（38~50）下进行。最终结果如表3.2所示。

表3.2 不同比例因子下编码器的性能

|  |  |  |  |
| --- | --- | --- | --- |
| 比例因子 a | | ΔPSNR | Δbitrate |
| a1 | 2/1 | -1.256 | 7.89 |
| a2 | 3/2 | -1.473 | 5.4 |
| a3 | 4/3 | -1.432 | 6.32 |
| a4 | 5/3 | -1.357 | 6.53 |
| a5 | 5/4 | -1.396 | 6.88 |
| a6 | 7/5 | -1.348 | 6.62 |



图3.7 6种不同比例因子下对应的编码器的性能

从图3.7中可以观察到，比例因子a1、a6、a4和a2在帕累托边界，也就是说就比特率和PSNR而言，它们是最佳的选择。在本文中，选择a2为最佳模式映射的比例因子，因为比例因子a2的比特率性能明显优于a1、a6和a4，而它们的PSNR性能差别不大。因此，最终的模式映射的比例因子为R2/R1=a=3/2。

比例因子a和8x8CU块的SPLIT模式点的坐标确定之后，不同depth对应的不同划分模式的映射点MP就可以求出来，如表3.3所示

表3.3 不同层不同划分模式在同一坐标系下对应的坐标

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| CU depth | | 0 | 1 | 2 | 3 |
| CU size | | 64x64 | 32x32 | 16x16 | 8x8 |
| MPs | 2Nx2N | (0,0) | (53,53) | (88,88) | (111,111) |
| 2NxN | (0,53) | (53,88) | (88,111) | (111,126) |
| Nx2N | (53,0) | (88,53) | (111,88) | (126,111) |
| NxN | (53,53) | (88,88) | (111,111) | (126,126) |

### 3.2.2 最佳预测划分模式点映射

如果一个深度为d的块已经被编码，那么根据它的划分模式MPs（如果它继续向下划分，那么也包括他所有子块的划分模式MPs）可以求得最终划分模式映射到二维坐标系中的位置SP。具体计算方式如式3-6：



式中d和分别代表当前编码块的深度以及它的最大深度；m代表当前编码块中CBs的个数。

因为最优划分模式点的对应位置SP只能在已编码的块中求得，而未编码的块就不存在SP，因此本文引入最佳预测划分模式点的概念。[23]由于在自然视频序列中相邻块间存在较强的相关性，因此本文根据已经编码的相邻块来预测当前块的最佳预测划分模式点的位置。本文将相邻编码块分为两大类：时域相邻块和空域相邻块。其中时域相邻块指的是上一帧中对应位置的编码块；空域相邻块是指同一帧中相邻的左方、上方和左上方对应位置的编码块。如图3.8所示

 

（a）空域相邻块 （b）时域相邻块

图3.8 时域相邻块与空域相邻块的位置

对于时域相邻块的CU，它的最佳划分模式点SP的坐标记作；对于空域相邻块的CU，它们最佳划分模式点SPs的坐标算术平均值记作。本文通过当前编码块最佳划分模式点的位置分别与它对应的和之间的归一化距离Dist来表示它们之间的关系， Dist的定义如式3-7：



其中为当前编码块最佳划分模式点的位置，为相邻块最佳划分模式点的位置，为坐标系中的最大距离。经过对不同分辨率不同内容的视频序列（RaceHorses\_832x480、PartyScene\_832x480、FourPeople\_1280x720、Johnny\_1280x720、BasketballDrive\_1920x1080、BQTerrace\_1920x1080）的不同大小的编码块在不同的QP下进行测试，最终结果如表3.4所示

表3.4 最佳划分模式点的分别与它对应的和之间的归一化距离Dist

|  |  |  |
| --- | --- | --- |
| CU size |  |  |
| 64x64 | 0.3016 | 0.1864 |
| 32x32 | 0.3268 | 0.2031 |
| 16x16 | 0.3577 | 0.2264 |
| 8x8 | 0.3713 | 0.2538 |

从表3.4中可以看出，当前编码块的空间相关性比时间相关性更强。为了更精确的减小预测误差，本文设置一个自适应权重因子动态地改变时域块与空域块所占的比重，如式3-8所示：





式3-9中为经过自适应权重因子改善后的最佳划分模式预测点的位置坐标。当空域相邻块不存在时，为0；t是当前帧的帧号，n为一个预定义的帧数，它代表可用来调整自适应权重因子的帧的个数。将式3-9展开可得：



将式3-10继续分解可以得到两部分：第一部分为从第t-n到第t-2帧所作的贡献如式3-11；第二部分为第t-1帧所做的贡献如式3-12，为第t-1帧的权重因子如3-13：







当前块的自适应权重因子定义如式3-14：



为了确保的值在[0,1]，有



通常为了简化的计算，通常将设置为一个定值，一般取0.1，将的初始值。

## 3.3 基于二维坐标系模式点映射生成的划分模式候选集

利用3.2.2求得的并不是最终编码后的SP，所以不能仅仅把距离SP最近的MP作为最终的划分模式，根据最佳中心点与所有编码划分模式点间的距离，建立一个基于优先级（距离越小优先级越高）的模式候选列表。最后的模式决策过程是按照模式候选列表中的顺序依次遍历模式候选列表中的模式（从最重要的模式到最不重要的模式）。对于每一层，模式候选列表中会有NONE、SPLIT、HORIZON、VERTICAL四种模式按照优先级排列。[24]如果遍历所有模式那么该算法就毫无意义，所以基于最优停止理论，为了用有限数量的划分模式保持研发性能，这些候选模式应该有最高预测概率。因此本文提出了终止距离的概念，其表达式如式3-16所示：



如图3.9所示，为与所有MP的距离中最长的距离，为与所有MP的距离中最小的距离；为终止距离，Θ为自适应因子，。如果MP所代表的模式与的距离R>，则跳过该模式。



图3.9 Rc与和的关系

Θ越大，遍历的划分模式就越多。Θ = 0则，只遍历一个划分模式；Θ = 1则，遍历所有的划分模式，本文中取Θ=0.65。

## 3.4基于代价估计的早期停止优化方案

假设有一种情况：模式候选列表中有三个候选模式，而候选模式中的第一个划分模式就是最优的划分模式，如果算法中没有早期停止的条件，那么编码器会继续遍历剩下的两种划分模式导致了整体复杂度的浪费。因此，本节提出一种基于代价估计的早期停止优化方案。

为了确保最终选定的划分模式具有较低的RDCost，最常见的方式就是设置一个RDCost阈值，在VP9编码器中，RDCost的表达式如式3-17所示：



式3-17中，RM和DM是与量化参数相关的常量。R为编码当前预测模式下所有信息（如变换系数、模式信息等）所需的比特数，D为当前预测模式下的失真。

为了减小计算复杂度，最有效的方法就是从已经编码过的块中获取RDCost信息。在VP9视频编码中，当前的编码块CU的划分模式与其时域相邻块和空域相邻块有很大的相关性。通过对不同分辨率不同内容的视频序列在不同的QP下进行测试，表3.5说明了当前块的最优RDCost分别与时域对应块的最优RDCost和空域相邻块的最优RDCost（平均值）之间的关系。表3.5中表示当前块的最优RDCost与空域相邻块的最优RDCost平均值的相关系数；表示当前块的最优RDCost与时域对应块的最优RDCost之间的相关系数；表示在当前块与时域对应块的最优划分模式相同的条件下当前块的最优RDCost与时域对应块的最优RDCost之间的相关系数。

表3.5 当前块的RDCost分别与时域和空域的相关性

|  |  |  |  |
| --- | --- | --- | --- |
| 视频序列 |  |  |  |
| Cactus\_1920x1080 | 0.802 | 0.941 | 0.968 |
| Kimono\_1920x1080 | 0.814 | 0.938 | 0.957 |
| FourPeople\_1280x720 | 0.738 | 0.902 | 0.934 |
| Johnny\_1280x720 | 0.756 | 0.911 | 0.926 |
| BasketballDrill\_832x480 | 0.729 | 0.889 | 0.907 |
| BQMall\_832x480 | 0.711 | 0.892 | 0.916 |
| 平均值 | 0.758 | 0.912 | 0.935 |

表3.5说明当前块的最优RDCost与时域相邻块的最优RDCost之间具有比较高的相关性，而且虽然在当前块与时域对应块的最优划分模式相同的条件下当前块的最优RDCost与时域对应块的最优RDCost之间具有更高的相关性，但是当前块与时域对应块的最优划分模式不相同当前块的最优RDCost与时域对应块的最优RDCost之间仍具有较高的相关性。因此本文为早期停止条件设置一个RDCost阈值：



式3-18中代表时域对应块的最优RDCost；γ为一个修正因子，一般取γ=1。在遍历候选列表中的划分模式时，如果当前划分模式的RDCost小于阈值，则早期停止不再遍历。

## 3.5 整体算法实现

基于二维坐标系模式点映射的VP9帧间划分模式快速算法的整个流程如图3.10所示。



图3.10 帧间划分模式快速算法的流程图

## 3.6实验结果与分析

将上述快速算法添加到VP9帧间预测中，对视频序列进行编码性能测试并与VP9原编码器的编码性能进行比较。本测试中采用的量化参数QP分别为152、168、184、200。测试结果共包括5项，分别为该算法相对原算法编码时间的减少量所占的百分比，比特率的增量和PSNR的变化量、和。在该算法下对应的编码性能如表3.6所示

表3.6 编码性能测试结果

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 视频序列 | QP | (%) | (%) | (dB) | (dB) | (dB) |
| Basketball | 152 | -48.2 | 2.26 | -0.037 | -0.081 | -0.021 |
| 168 | -49.9 | 3.52 | -0.146 | -0.044 | -0.007 |
| 184 | -50.3 | 5.08 | -0.247 | -0.071 | -0.035 |
| 200 | -51.5 | 7.06 | -0.388 | -0.011 | -0.068 |
| ParkScene | 152 | -49.6 | 1.98 | -0.013 | -0.06 | -0.09 |
| 168 | -50.3 | 2.79 | -0.009 | -0.282 | -0.317 |
| 184 | -51.9 | 5.69 | -0.015 | -0.25 | -0.218 |
| 200 | -52.3 | 9.61 | -0.028 | -0.215 | -0.257 |
| BQTerrace | 152 | -47.3 | 2.04 | -0.081 | -0.142 | -0.094 |
| 168 | -48.1 | 4.38 | -0.026 | -0.098 | -0.073 |
| 184 | -50.3 | 5.55 | -0.019 | -0.076 | -0.154 |
| 200 | -51.4 | 6.68 | -0.007 | -0.06 | -0.213 |
| Kimono | 152 | -48.8 | 2.52 | -0.102 | -0.112 | -0.136 |
| 168 | -49.6 | 3.98 | -0.064 | -0.042 | -0.09 |
| 184 | -51.2 | 4.63 | -0.113 | -0.011 | -0.059 |
| 200 | -52.7 | 6.78 | -0.236 | -0.007 | -0.026 |
| Cactus | 152 | -46.9 | 1.88 | -0.073 | -0.064 | -0.057 |
| 168 | -48.1 | 2.73 | -0.041 | -0.051 | -0.031 |
| 184 | -49.6 | 5.21 | -0.032 | -0.028 | -0.027 |
| 200 | -50.7 | 10.36 | -0.011 | -0.007 | -0.002 |
| 平均 | ---- | -50.02 | 4.49 | -0.095 | -0.096 | -0.113 |

实验结果表明，本文的帧间快速算法可以在一定程度上降低帧间编码复杂度，使编码时间与VP9原版相比平均降低50%左右。与此同时，PSNR下降在0.01左右，但是比特率平均上升4.5%。本算法基本达到了预期效果。

在本算法中，涉及到坐标、门限值的设置，这些值是根据大量实验数据来设置为定值，不具有自适应性，所以该算法还有一定的提升空间。

## 3.7 本章小结

为了降低VP9编码器的整体复杂度，减少编码时间，本章基于CU的划分模式提出了一种基于二维坐标模式点映射的帧间快速算法。该算法利用当前块的时空域信息建立基于模式点映射的候选模式列表，然后根据最优搜索半径与代价估计对算法进行优化，最终使得候选划分列表得到缩减。

本章的最后进行了大量实验对该算法进行了测试，结果表明该算法在PSNR下降很少的情况下降低了编码复杂度，减少了编码时间。不过该算法使得比特率上升了许多，所以还有一定的提升空间。

# 第四章 总结与展望

随着数字视频的分辨率不断提高，目前人们所使用的平板电脑、智能手机等电子设备已经逐步将1080P的高清视频格式作为标配，在某些场合里，甚至有超高清的视频格式（如4Kx2K）的出现。人们日益增长的对高清视频的需求促使高压缩比的视频压缩标准的发展，VP9是由谷歌推出的免费视频编码标准，其目标在同等画质下码率较上一代VP8可降低50%，编码效率超过H.265/HEVC，并支持更大的编码块、更多的色彩空间。但是编码效率提升的同时，编码复杂度也大大增加，致使使用VP9编码器进行视频编码会耗费大量的时间，因此尽可能地减少编码时间，实现高清分辨率下的实时传输将是一个必然趋势，减少帧间模式选择的时间将是一个关键点。

本文以VP9视频编码器为研究对象，主要研究VP9中帧间快速编码模式预测部分，对编码中耗时较多的帧间编码技术提出快速算法，做到在不影响压缩率和压缩质量的情况下提高编码性能，提高VP9的实用性。由此本文提出了一个以二维坐标系模式点映射为基础的算法，利用该算法构建一个模式候选列表用于VP9的帧间快速划分模式决策。首先利用当前块的时空域信息建立基于模式点映射的候选模式列表，然后在模式决策过程中，提出了基于最优搜索半径的早期停止策略、基于代价估计的早期停止策略用来优化模式候选列表。实验证明结合早期终止算法，帧间划分模式决策过程消耗的复杂度大大减少。

在本人所提出的算法中涉及到坐标、门限值的设置，这些值是根据大量实验数据来设置的一般设为定值，不具有自适应性，所以如何让它们自适应也要进一步的研究。未来视频编码必然会引入数量更多更复杂的块侦测机制，随着硬件运算速度的提高，编码将会以提高复杂度换取空间的节省、以加大缓冲内存占用空间提升侦测的准确性、以更多线程的设计更优化进行并行运算、以可适性技术增强在线传输的稳定性和容错性，从而推动硬件更新换代和新应用方式的产生。或许将来会出现能经过简单定制适应各种行业应用需求的通用视频编码，而现在，至少VP9已经前进了一步。而随着Android系统的普及，以其为基础的智能手机、平板、智能电视等产品，也许同样会成为VP9的重要推动力。

# 致 谢

论文写到这里已经快要完成，这也意味着我的大学本科四年也即将结束。此刻，回想这四年大学时光，我的心里充满了不舍与感激。

首先，我要特别感谢我的研究生导师杨付正教授。感谢杨老师在这半年的毕设阶段对我在科研学术上的谆谆教导与在生活上的指点迷津。杨老师在对生活的态度、人生观和世界观方面也深深感染了我。在生活上，杨老师关心着实验室的每个成员，时刻提醒着我们锻炼身体的重要性，让我们养成了良好的生活习惯。在人际交往方面，杨老师教我们待人真诚、友善，常怀感恩之心。他的严格要求和悉心指导，让我对待问题有了更深的体会与认识，也让我顺利地完成了本科毕业设计。

在毕业设计期间遇到了很多问题与困难，非常感谢马彦卓老师在生活上对我的关怀与帮助和在学习上对我的关心与指导。感谢冯俊凯师兄、李文荣师兄在视频编码方向对我的指导，感谢李璞、邱孟品、俞宏捷、闵楠师兄师姐在学习上对我的帮助，感谢同窗好友张赛萍、陈志杰、汪泽伟、周凡、饶童、王学昭对我在学习和生活上的帮助。

感谢我的大学室友陈元魁、孟勇、荣磊、王磊、范佳乐、宋建伟、王存发在我大学四年里对我无微不至的关怀与照料，是你们让我的大学生活里充满了欢声笑语，你们的陪伴让我笑对每一次困难与挫折，感谢你们让我拥有一个美好的大学时光。

最后我要深深地感谢我的家人，感谢爸爸妈妈一直在背后默默无闻的付出，感谢你们一直以来对我的支持与理解。感谢我的姐姐对我每一次的关怀与鼓励，让我更加有信心面对挫折与失败，让我更加勇敢地面对未来！

最后向参与论文审阅和答辩的专家和老师们表示感谢。感谢你们能在百忙之中抽出宝贵时间来参加我的答辩。

# 参考文献

1. Video coding standards: AVS China, H.264/MPEG-4 part 10, HEVC, VP9, DIRAC and VC-1 Rao, K.R. 2013.
2. A Technical Overview of VP9 – The Latest Open-Source Video Codec Mukherjee, Debargha 2013.
3. 刘小鹤. 开放的VP9谷歌新一代互联网高清视频编码. 微型计算机. 2014年3月上
4. 何海东, 董全武, 陈宏. HEVC & VP9在网络直播点播中的应用探索(上). 广播与电视技术. 2014, 41(8). 171-172.
5. 毕厚杰. 新一代视频压缩编码标准—H.264/AVC [J]. 人民邮电出版社. 2005.
6. Wei Jiang, Hanjie Ma, and Yaowu Chen. “Gradient Based Fast Mode Decision Algorithm for Intra Prediction in HEVC ,“Signal Image Technology and Internet Based Systems（SITIS）,2012 Eighth International Conference on. 2012, 225-229.
7. Jiang W, Ma H, and Chen Y. “Gradient based fast mode decision algorithm for intra prediction in HEVC ,” Processdings of the 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet). 2012: 1836-1840.
8. Da S, Thaisa L, Agostini L V, et al.“ Fast HEVC intra prediction mode decision based on EDGE direction information ,” Proceedings of the 20th European Signal Processing Conference (EUSIPCO). 2012: 1214-1248.
9. Feng Pan, Xiao Lin, and Susanto Rahardja.“ Fast Mode Decision Algorithm for Intra Prediction in H.264/AVC Video Coding ，”. IEEE Transaction on Circuits and Systems for Video Technology. 2005.
10. Tiesong Zhao ,SamKwong , and HanliWang,“H.264/SVC Mode Decision Based on Optimal Stopping Theory,” IEEE Transactions On Image Processing, VOL. 21, NO. 5, MAY 2012
11. 杨金凤. VP8实时视频编码器软件实现及优化[D]. 西安电子科技大学, 2013. DOI:doi:10.7666/d.Y2397186.
12. 崔子冠. H.264率失真优化和码率控制算法的研究与改进[D]. 南京航空航天大学, 2008. DOI:doi:10.7666/d.d052159.
13. En-Hui Yang, Xiang Yu, Jin Meng, and Chang Sun,“Transparent Composite Model for DCT Coefficients: Design and Analysis，” IEEE Transactions On Image Processing, VOL. 23, NO. 3, MARCH 2014
14. En-hui Yang and Xiang Yu,“Transparent Composite Model for Large Scale Image/Video Processing,” IEEE International Conference on Big Data, 2013.
15. Soon-heung Jung and HyunWook Park, “A fast mode decision method in HEVC using adaptive ordering of modes,” IEEE Transactions on Circuits and Systems for Video Technology. 2015.2473303
16. Jaehwan Kim, Jungyoup Yang, Kwanghyun Won, and Byeungwoo Jeon,“Early Determination of Mode Decision for HEVC，” Picture Coding Symposium，May 7-9, 2012.
17. X.Yang, G.Teng, H.Zhao, G.Li, P.An, and G.Wang,“Fast PU Decision Algorithm Based on Texture Complexity in HEVC，”in Proc.IEEE Int. Conf. Signal Processing, Communications and Computing, pp. 321-325, Aug. 2014.
18. Nan Hu, and En-hui Yang,“Fast Inter Mode Decision for HEVC Based on Transparent Composite Model，” Natrual Sciences And Engreening Reaearch Council，2015,203035-11
19. T. Zhao, Z. Wang , and S. Kwong，“Flexible mode selection and complexity allocation in high efficiency video coding,” IEEE Journal of Selected Topics in Signal Processing, VOL. 7, NO. 6, DEC 2013
20. M.Grellert, M Shafique and M Usman Karim Khan,“An Adaptive Workload Management Scheme For Hevc Encoding,” in 2013 20th IEEE International Conference on Image Processing (ICIP), 2013.
21. 万帅, 杨付正. 新一代高效视频编码H.265/HEVC原理、标准与实现, 电子工业出版社, 2014.
22. T. Zhao, H. Wang, S. Kwong, and C.-C. J. Kuo, “Fast mode decision based on mode adaptation,”IEEE Trans. Circuits Syst. Video Technol. ,vol. 20, no. 5, pp. 697–705, May 2010.
23. T. Zhao, H. Wang, X. Ji, and S. Kwong, “2-D map-based fast mode decision for H.264/AVC,” in Proc. Pacific-Rim Conf. Multimedia (PCM),LNCS 5353, Dec. 2008, pp. 148–157.
24. S. H. Ri, Y. Vatis, and J. Ostermann, “Fast Inter-Mode Decision in an H.264/AVC Encoder Using Mode and Lagrangian Cost Correlation”, IEEE Trans. Circuits and Systems for Video Technology, vol. 19, pp. 302-306, Feb. 2009.
25. Shuai Li, Ce Zhu,and Yanbo Gao,“Inter-Frame Dependent Rate-Distortion Optimization Using Lagrangian Multiplier Adaption,” Programs Foundation of Ministry of Education of China 20130182110010
26. Thanuja Mallikarachchi, Anil Fernando and Hemantha Kodikara Arachchi ,“Fast Coding Unit Size Selection for HEVC Inter Prediction,” IEEE International Conference on Consumer Electronics (ICCE), 2015.
27. Thierry Luhandjula, Keith L Ferguson,“Sampling Point Path Selection for Fast Intra Mode Prediction,” PCS 2015
28. Muhammad Usman Shahid, Ashfaq Ahmed, Maurizio Martina,“Parallel H.264/AVC Fast Rate-Distortion Optimized Motion Estimation by Using a Graphics Processing Unit and Dedicated Hardware，” IEEE Transactions On Circuits And Systems For Video Technology, VOL. 25, NO. 4, APRIL 2015
29. G. Zhong, X. He, L. Qing, and Y. Li, "A fast inter-prediction algorithm for HEVC based on temporal and spatial correlation," Multimedia Tools and Applications, pp. 1-21, 2014/08/01 2014.
30. Qian Chen and Dapeng Wu, “Delay-Rate-Distortion Model for Real-Time Video Communication,” IEEE Transactions On Circuits And Systems For Video Technology,VOL.25, NO. 8, AUGUST 2015