

清 华 大 学

综 合 论 文 训 练

题目：基于早终止的 AV1 视频编码加速

系 别：计算机科学与技术系

专 业：计算机科学与技术

姓 名：杨皖宁

指导教师：温江涛教授

2018 年 5 月 22 日

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：学校有权保留学位论文的复印件，允许该论文被查阅和借阅；学校可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存该论文。

(涉密的学位论文在解密后应遵守此规定)

签 名：_____ 导师签名：_____ 日 期：_____

中文摘要

在直播行业

论文的摘要是对论文研究内容和成果的高度概括。摘要应对论文所研究的问题及其研究目的进行描述，对研究方法和过程进行简单介绍，对研究成果和所得结论进行概括。摘要应具有独立性和自明性，其内容应包含与论文全文同等量的主要信息。使读者即使不阅读全文，通过摘要就能了解论文的总体内容和主要成果。

论文摘要的书写应力求精确、简明。切忌写成对论文书写内容进行提要的形式，尤其要避免“第 1 章……；第 2 章……；……”这种或类似的陈述方式。

本文介绍清华大学论文模板 THUTHESIS 的使用方法。本模板符合学校的本科、硕士、博士论文格式要求。

本文的创新点主要有：

- 用例子来解释模板的使用方法；
- 用废话来填充无关紧要的部分；
- 一边学习摸索一边编写新代码。

关键词是为了文献标引工作、用以表示全文主要内容信息的单词或术语。关键词不超过 5 个，每个关键词中间用分号分隔。（模板作者注：关键词分隔符不用考虑，模板会自动处理。英文关键词同理。）

关键词：AV1；早终止；视频编码；树剪枝

ABSTRACT

An abstract of a dissertation is a summary and extraction of research work and contributions. Included in an abstract should be description of research topic and research objective, brief introduction to methodology and research process, and summarization of conclusion and contributions of the research. An abstract should be characterized by independence and clarity and carry identical information with the dissertation. It should be such that the general idea and major contributions of the dissertation are conveyed without reading the dissertation.

An abstract should be concise and to the point. It is a misunderstanding to make an abstract an outline of the dissertation and words “the first chapter”, “the second chapter” and the like should be avoided in the abstract.

Key words are terms used in a dissertation for indexing, reflecting core information of the dissertation. An abstract may contain a maximum of 5 key words, with semicolons used in between to separate one another.

Keywords: AV1; early-termination; video coding; tree pruning

目 录

第 1 章 改动。带 English 的标题	1
1.1 封面相关	1
1.2 字体命令	1
1.3 表格样本	2
1.3.1 基本表格	2
1.3.2 复杂表格	3
1.3.3 其它	7
1.4 定理环境	8
1.5 参考文献	11
1.6 公式	12
第 2 章 中华人民共和国	14
2.1 其它例子	14
2.1.1 绘图	14
2.1.2 插图	14
插图索引	17
表格索引	18
公式索引	19
参考文献	20
致 谢	22
声 明	23
附录 A 书面翻译 1	24
A.1 引言	24
A.2 基于树剪枝的快速编码单元推断	25
A.3 实验结果	27
A.4 结论	29

附录 B 书面翻译 2	30
B.1 引言	30
B.2 早终止算法	31
B.2.1 可变块大小零运动推断	31
B.2.2 可变块大小最佳运动推断.....	35
B.3 模拟结果	38
B.4 结论	39

主要符号对照表

HEVC	高效视频编码 (High Efficiency Video Coding)
CU	
PU	
TU	
SKIP	
JCT-VC	(Joint Collaborative Team on Video Coding)
ISO/IEC	国际标准化组织 (International Standardization Organization)
IEC	国际电工委员会 (International Electrotechnical Commission)
MPEG	动态图像专家组 (Moving Picture Experts Group)
ITU-T	国际电信联盟电信标准分局 (ITU Telecommunication Standardization Sector)
VCEG	视频编码专家组 (Video Coding Experts Group)
AVC	高级视频编码 (Advanced Video Coding)
ME	运动估计 (Motion Estimation)
MV	运动向量 (Motion Vector)
DCT	离散余弦变换 (Discrete Cosine Transform)
CAVLC	
QCIF	

第 1 章 改动。带 English 的标题

这是 `ThUTHESIS`^[1] 的示例文档，基本上覆盖了模板中所有格式的设置。建议大家在使用模板之前，除了阅读《`ThUTHESIS` 用户手册》，这个示例文档也最好能看一看。

小老鼠偷吃热凉粉；短长虫环绕矮高粱^②。

1.1 封面相关

封面的例子请参看 `cover.tex`。主要符号表参看 `denation.tex`，附录和个人简历分别参看 `appendix01.tex` 和 `resume.tex`。里面的命令都很直观，一看即会^②。

1.2 字体命令

苏轼（1037-1101），北宋文学家、书画家。字子瞻，号东坡居士，眉州眉山（今属四川）人。苏洵子。嘉佑进士。神宗时曾任祠部员外郎，因反对王安石新法而求外职，任杭州通判，知密州、徐州、湖州。后以作诗“谤讪朝廷”罪贬黄州。哲宗时任翰林学士，曾出知杭州、颖州等，官至礼部尚书。后又贬谪惠州、儋州。北还后第二年病死常州。南宋时追谥文忠。与父洵弟辙，合称“三苏”。在政治上属于旧党，但也有改革弊政的要求。其文汪洋恣肆，明白畅达，为“唐宋八大家”之一。其诗清新豪健，善用夸张比喻，在艺术表现方面独具风格。少数诗篇也能反映民间疾苦，指责统治者的奢侈骄纵。词开豪放一派，对后代很有影响。《念奴娇·赤壁怀古》、《水调歌头·丙辰中秋》传诵甚广。

坡仙擅长行书、楷书，取法李邕、徐浩、颜真卿、杨凝式，而能自创新意。用笔丰腴跌宕，有天真烂漫之趣。与蔡襄、黄庭坚、米芾并称“宋四家”。能画竹，学文同，也喜作枯木怪石。论画主张“神似”，认为“论画以形似，见与儿

① 韩愈（768-824），字退之，河南河阳（今河南孟县）人，自称郡望昌黎，世称韩昌黎。幼孤贫刻苦好学，德宗贞元八年进士。曾任监察御史，因上疏请免关中赋役，贬为阳山县令。后随宰相裴度平定淮西迁刑部侍郎，又因上表谏迎佛骨，贬潮州刺史。做过吏部侍郎，死谥文公，故世称韩吏部、韩文公。是唐代古文运动领袖，与柳宗元合称韩柳。诗力求险怪新奇，雄浑重气势。

② 你说还是看不懂？怎么会呢？

童邻”；高度评价“诗中有画，画中有诗”的艺术造诣。诗文有《东坡七集》等。存世书迹有《答谢民师论文帖》、《祭黄几道文》、《前赤壁赋》、《黄州寒食诗帖》等。画迹有《枯木怪石图》、《竹石图》等。

易与天地准，故能弥纶天地之道。仰以观於天文，俯以察於地理，是故知幽明之故。原始反终，故知死生之说。精气为物，游魂为变，是故知鬼神之情状。与天地相似，故不违。知周乎万物，而道济天下，故不过。旁行而不流，乐天知命，故不忧。安土敦乎仁，故能爱。范围天地之化而不过，曲成万物而不遗，通乎昼夜之道而知，故神无方而易无体。

有天地，然后万物生焉。盈天地之间者，唯万物，故受之以屯；屯者盈也，屯者物之始生也。物生必蒙，故受之以蒙；蒙者蒙也，物之穉也。物穉不可不养也，故受之以需；需者饮食之道也。饮食必有讼，故受之以讼。讼必有众起，故受之以师；师者众也。众必有所比，故受之以比；比者比也。比必有所畜也，故受之以小畜。物畜然后有礼，故受之以履。

履而泰，然后安，故受之以泰；泰者通也。物不可以终通，故受之以否。物不可以终否，故受之以同人。与人同者，物必归焉，故受之以大有。有大者不可以盈，故受之以谦。有大而能谦，必豫，故受之以豫。豫必有随，故受之以随。以喜随人者，必有事，故受之以蛊；蛊者事也。

有事而后可大，故受之以临；临者大也。物大然后可观，故受之以观。可观而后有所合，故受之以噬嗑；嗑者合也。物不可以苟合而已，故受之以贲；贲者饰也。致饰然后亨，则尽矣，故受之以剥；剥者剥也。物不可以终尽，剥穷上反下，故受之以复。复则不妄矣，故受之以无妄。

有无妄然后可畜，故受之以大畜。物畜然后可养，故受之以颐；颐者养也。不养则不可动，故受之以大过。物不可以终过，故受之以坎；坎者陷也。陷必有所丽，故受之以离；离者丽也。

1.3 表格样本

1.3.1 基本表格

模板中关于表格的宏包有三个：booktabs、array 和 longtabular，命令有一个 \hlinewd。三线表可以用 booktabs 提供的 \toprule、\midrule 和 \bottomrule。它们与 longtable 能很好的配合使用。如果表格比较简单的话可以直接用命令 \hlinewd{<width>} 控制。

表 1.1 模板文件。如果表格的标题很长，那么在表格索引中就会很不美观，所以要像 `chapter` 那样在前面用中括号写一个简短的标题。这个标题会出现在索引中。

文件名	描述
<code>thuthesis.ins</code>	L ^A T _E X 安装文件，DocStrip ^①
<code>thuthesis.dtx</code>	所有的一切都在这里面 ^② 。
<code>thuthesis.cls</code>	模板类文件。
<code>thuthesis.cfg</code>	模板配置文。 <code>cls</code> 和 <code>cfg</code> 由前两个文件生成。
<code>thuthesis-numeric.bst</code>	参考文献 BIB _T E _X 样式文件。
<code>thuthesis-author-year.bst</code>	参考文献 BIB _T E _X 样式文件。
<code>thuthesis.sty</code>	常用的包和命令写在这里，减轻主文件的负担。

① 表格中的脚注

② 再来一个

首先来看一个最简单的表格。表 1.1 列举了本模板主要文件及其功能。请大家注意三线表中各条线对应的命令。这个例子还展示了如何在表格中正确使用脚注。由于 L^AT_EX 本身不支持在表格中使用 `\footnote`，所以我们不得不将表格放在小页中，而且最好将表格的宽度设置为小页的宽度，这样脚注看起来才更美观。

1.3.2 复杂表格

我们经常会在表格下方标注数据来源，或者对表格里面的条目进行解释。前面的脚注是一种不错的方法，如果不喜欢脚注，可以在表格后面写注释，比如表 1.2。

表 1.2 复杂表格示例 1。这个引用^[2]不会导致编号混乱。

x \ y	First Half		Second Half	
	1st Qtr	2nd Qtr	3rd Qtr	4th Qtr
East*	20.4	27.4	90	20.4
West**	30.6	38.6	34.6	31.6

注：数据来源《ThUThesis 使用手册》。

*：东部

**：西部

此外，表 1.2 同时还演示了另外两个功能：1) 通过 `tabularx` 的 `|x|` 扩展实

现表格自动放大；2）通过命令 \diagbox 在表头部分插入反斜线。

为了使我们的例子更接近实际情况，我会在必要的时候插入一些“无关”文字，以免太多图表同时出现，导致排版效果不太理想。第一个出场的当然是我的最爱：风流潇洒、骏马绝尘、健笔凌云的**李太白**了。

李白，字太白，陇西成纪人。凉武昭王暠九世孙。或曰山东人，或曰蜀人。白少有逸才，志气宏放，飘然有超世之心。初隐岷山，益州长史苏颋见而异之，曰：“是子天才英特，可比相如。”天宝初，至长安，往见贺知章。知章见其文，叹曰：“子谪仙人也。”言于明皇，召见金銮殿，奏颂一篇。帝赐食，亲为调羹，有诏供奉翰林。白犹与酒徒饮于市，帝坐沉香亭子，意有所感，欲得白为乐章，召入，而白已醉。左右以水颰面，稍解，援笔成文，婉丽精切。帝爱其才，数宴见。白常侍帝，醉，使高力士脱靴。力士素贵，耻之，摘其诗以激杨贵妃。帝欲官白，妃辄沮止。白自知不为亲近所容，恳求还山。帝赐金放还。乃浪迹江湖，终日沉饮。永王璘都督江陵，辟为僚佐。璘谋乱，兵败，白坐长流夜郎，会赦得还。族人阳冰为当涂令，白往依之。代宗立，以左拾遗召，而白已卒。文宗时，诏以白歌诗、裴旻剑舞、张旭草书为三绝云。集三十卷。今编诗二十五卷。——《全唐诗》诗人小传

浮动体的并排放置一般有两种情况：1）二者没有关系，为两个独立的浮动体；2）二者隶属于同一个浮动体。对表格来说并排表格既可以像图 1.3、图 1.4 使用小页环境，也可以如图 1.5 使用子表格来做。图的例子参见第 2.1.2.2 节。

表 1.3 第一个并排子表格

111	222
222	333

表 1.4 第二个并排子表格

111	222
222	333

然后就是忧国忧民，诗家楷模杜工部了。杜甫，字子美，其先襄阳人，曾祖依艺为巩令，因居巩。甫天宝初应进士，不第。后献《三大礼赋》，明皇奇之，召试文章，授京兆府兵曹参军。安禄山陷京师，肃宗即位灵武，甫自贼中遁赴行在，拜左拾遗。以论救房琯，出为华州司功参军。关辅饥乱，寓居同州同谷县，身自负薪采椽，舖糒不给。久之，召补京兆府功曹，道阻不赴。严武镇成都，奏为参谋、检校工部员外郎，赐绯。武与甫世旧，待遇甚厚。乃于成都浣花里种竹植树，枕江结庐，纵酒啸歌其中。武卒，甫无所依，乃之东蜀就高适。既至而适卒。是岁，蜀帅相攻杀，蜀大扰。甫携家避乱荆楚，扁舟下峡，未维舟而江陵亦

乱。乃溯沿湘流，游衡山，寓居耒阳。卒年五十九。元和中，归葬偃师首阳山，元稹志其墓。天宝间，甫与李白齐名，时称李杜。然元稹之言曰：“李白壮浪纵恣，摆去拘束，诚亦差肩子美矣。至若铺陈终始，排比声韵，大或千言，次犹数百，词气豪迈，而风调清深，属对律切，而脱弃凡近，则李尚不能历其藩翰，况堂奥乎。”白居易亦云：“杜诗贯穿古今，尽工尽善，殆过于李。”元、白之论如此。盖其出处劳佚，喜乐悲愤，好贤恶恶，一见之于诗。而又以忠君忧国、伤时念乱为本旨。读其诗可以知其世，故当时谓之“诗史”。旧集诗文共六十卷，今编诗十九卷。

表 1.5 并排子表格

(a) 第一个子表格		(b) 第二个子表格	
111	222	111	222
222	333	222	333

不可否认 \LaTeX 的表格功能没有想象中的那么强大，不过只要足够认真，足够细致，同样可以排出来非常复杂非常漂亮的表格。请参看表 1.6。

表 1.6 复杂表格示例 2

Network Topology		# of nodes	# of clients			Server
GT-ITM	Waxman Transit-Stub	600	2%	10%	50%	Max. Connectivity
Inet-2.1		6000				
Xue	Rui	Ni	ThUThESIS			
	ABCDEF					

最后就是清新飘逸、文约意赅、空谷绝响的王大侠了。王维，字摩诘，河东人。工书画，与弟缙俱有俊才。开元九年，进士擢第，调太乐丞。坐累为济州司仓参军，历右拾遗、监察御史、左补阙、库部郎中，拜吏部郎中。天宝末，为给事中。安禄山陷两都，维为贼所得，服药阳喑，拘于菩提寺。禄山宴凝碧池，维潜赋诗悲悼，闻于行在。贼平，陷贼官三等定罪，特原之，责授太子中允，迁中庶子、中书舍人。复拜给事中，转尚书右丞。维以诗名盛于开元、天宝间，宁薛诸王驸马豪贵之门，无不拂席迎之。得宋之问辋川别墅，山水绝胜，与道友裴迪，浮舟往来，弹琴赋诗，啸咏终日。笃于奉佛，晚年长斋禅诵。一日，忽索笔

作书数纸，别弟缙及平生亲故，舍笔而卒。赠秘书监。宝应中，代宗问缙：“朕常于诸王坐闻维乐章，今存几何？”缙集诗六卷，文四卷，表上之。敕答云，卿伯氏位列先朝，名高希代。抗行周雅，长揖楚辞。诗家者流，时论归美。克成编录，叹息良深。殷璠谓维诗词秀调雅，意新理惬。在泉成珠，著壁成绘。苏轼亦云：“维诗中有画，画中有诗也。”今编诗四卷。

要想用好论文模板还是得提前学习一些 $\text{T}_{\text{E}}\text{X}/\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ 的相关知识，具备一些基本能力，掌握一些常见技巧，否则一旦遇到问题还真比较麻烦。我们见过很多这样的同学，一直以来都是使用 Word 等字处理工具，以为 $\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ 模板的用法也应该类似，所以就沿袭同样的思路来对待这种所见非所得的排版工具，结果被折腾的焦头烂额，疲惫不堪。

如果您要排版的表格长度超过一页，那么推荐使用 `longtable` 或者 `supertabular` 宏包，模板对 `longtable` 进行了相应的设置，所以用起来可能简单一些。表 1.7 就是 `longtable` 的简单示例。

表 1.7 实验数据

测试程序	正常运行 时间 (s)	同步 时间 (s)	检查点 时间 (s)	卷回恢复 时间 (s)	进程迁移 时间 (s)	检查点 文件 (KB)
CG.A.2	23.05	0.002	0.116	0.035	0.589	32491
CG.A.4	15.06	0.003	0.067	0.021	0.351	18211
CG.A.8	13.38	0.004	0.072	0.023	0.210	9890
CG.B.2	867.45	0.002	0.864	0.232	3.256	228562
CG.B.4	501.61	0.003	0.438	0.136	2.075	123862
CG.B.8	384.65	0.004	0.457	0.108	1.235	63777
MG.A.2	112.27	0.002	0.846	0.237	3.930	236473
MG.A.4	59.84	0.003	0.442	0.128	2.070	123875
MG.A.8	31.38	0.003	0.476	0.114	1.041	60627
MG.B.2	526.28	0.002	0.821	0.238	4.176	236635
MG.B.4	280.11	0.003	0.432	0.130	1.706	123793
MG.B.8	148.29	0.003	0.442	0.116	0.893	60600
LU.A.2	2116.54	0.002	0.110	0.030	0.532	28754
LU.A.4	1102.50	0.002	0.069	0.017	0.255	14915
LU.A.8	574.47	0.003	0.067	0.016	0.192	8655
LU.B.2	9712.87	0.002	0.357	0.104	1.734	101975
LU.B.4	4757.80	0.003	0.190	0.056	0.808	53522
LU.B.8	2444.05	0.004	0.222	0.057	0.548	30134

续下页

续表 1.7 实验数据

测试程序	正常运行 时间 (s)	同步 时间 (s)	检查点 时间 (s)	卷回恢复 时间 (s)	进程迁移 时间 (s)	检查点 文件 (KB)
EP.A.2	123.81	0.002	0.010	0.003	0.074	1834
EP.A.4	61.92	0.003	0.011	0.004	0.073	1743
EP.A.8	31.06	0.004	0.017	0.005	0.073	1661
EP.B.2	495.49	0.001	0.009	0.003	0.196	2011
EP.B.4	247.69	0.002	0.012	0.004	0.122	1663
EP.B.8	126.74	0.003	0.017	0.005	0.083	1656

1.3.3 其它

如果不想让某个表格或者图片出现在索引里面,请使用命令 `\caption*`。这个命令不会给表格编号,也就是出来的只有标题文字而没有“表 XX”,“图 XX”,否则索引里面序号不连续就显得不伦不类,这也是 \LaTeX 里星号命令默认的规则。

有这种需求的多是本科同学的英文资料翻译部分,如果觉得附录中英文原文中的表格和图片显示成“表”和“图”不协调的话,一个很好的办法就是用 `\caption*`, 参数随便自己写,比如不守规矩的表 1.111 和图 1.111 能满足这种特殊需要(可以参看附录部分)。

表 1.111 这是一个手动编号,不出现在索引中的表格。

ThUThesis

Figure 1.111 这是一个手动编号,不出现在索引中的图。

薛瑞尼

如果的确想让它编号,但又不想让它出现在索引中的话,目前模板上不支持。

最后,虽然大家不一定会独立使用小页,但是关于小页中的脚注还是有必要提一下。请看下面的例子。

柳宗元，字子厚（773-819），河东（今永济县）人^①，是唐代杰出的文学家，哲学家，同时也是一位政治改革家。与韩愈共同倡导唐代古文运动，并称韩柳^②。

① 山西永济水饺。

② 唐宋八大家之首二位。

唐朝安史之乱后，宦官专权，藩镇割据，土地兼并日渐严重，社会生产破坏严重，民不聊生。柳宗元对这种社会现实极为不满，他积极参加了王叔文领导的“永济革新”，并成为这一运动的中坚人物。他们革除弊政，打击权奸，触犯了宦官和官僚贵族利益，在他们的联合反扑下，改革失败了，柳宗元被贬为永州司马。

1.4 定理环境

给大家演示一下各种和证明有关的环境：

假设 1.1： 待月西厢下，迎风户半开；隔墙花影动，疑是玉人来。

$$c = a^2 - b^2 \quad (1-1)$$

$$= (a + b)(a - b) \quad (1-2)$$

千辛万苦，历尽艰难，得有今日。然相从数千里，未曾哀戚。今将渡江，方图百年欢笑，如何反起悲伤？（引自《杜十娘怒沉百宝箱》）

定义 1.1： 子曰：「道千乘之国，敬事而信，节用而爱人，使民以时。」

千古第一定义！问世间、情为何物，只教生死相许？天南地北双飞客，老翅几回寒暑。欢乐趣，离别苦，就中更有痴儿女。君应有语，渺万里层云，千山暮雪，只影向谁去？

横汾路，寂寞当年箫鼓，荒烟依旧平楚。招魂楚些何嗟及，山鬼暗啼风雨。天也妒，未信与，莺儿燕子俱黄土。千秋万古，为留待骚人，狂歌痛饮，来访雁丘处。

命题 1.1： 曾子曰：「吾日三省吾身——为人谋而不忠乎？与朋友交而不信乎？传不习乎？」

多么凄美的命题啊！其日牛马嘶，新妇入青庐，奄奄黄昏后，寂寂人定初，我命绝今日，魂去尸长留，揽裙脱丝履，举身赴清池，府吏闻此事，心知长别离，徘徊庭树下，自挂东南枝。

注释 1.1： 天不言自高，水不言自流。

$$\begin{aligned}\varphi(x, z) &= z - \gamma_{10}x - \gamma_{mn}x^m z^n \\ &= z - Mr^{-1}x - Mr^{-(m+n)}x^m z^n\end{aligned}$$

$$\zeta^0 = (\xi^0)^2, \quad (1-3)$$

$$\zeta^1 = \xi^0 \xi^1, \quad (1-4)$$

$$\zeta^2 = (\xi^1)^2, \quad (1-5)$$

天尊地卑，乾坤定矣。卑高以陈，贵贱位矣。动静有常，刚柔断矣。方以类聚，物以群分，吉凶生矣。在天成象，在地成形，变化见矣。鼓之以雷霆，润之以风雨，日月运行，一寒一暑，乾道成男，坤道成女。乾知大始，坤作成物。乾以易知，坤以简能。易则易知，简则易从。易知则有亲，易从则有功。有亲则可久，有功则可大。可久则贤人之德，可大则贤人之业。易简，而天下矣之理矣；天下之理得，而成位乎其中矣。

公理 1.1： 两点间直线段距离最短。

$$x \equiv y + 1 \pmod{m^2} \quad (1-6)$$

$$x \equiv y + 1 \pmod{m^2} \quad (1-7)$$

$$x \equiv y + 1 \pmod{m^2} \quad (1-8)$$

《彖曰》：大哉乾元，万物资始，乃统天。云行雨施，品物流形。大明始终，六位时成，时乘六龙以御天。乾道变化，各正性命，保合大和，乃利贞。首出庶物，万国咸宁。

《象曰》：天行健，君子以自强不息。潜龙勿用，阳在下也。见龙再田，德施普也。终日乾乾，反复道也。或跃在渊，进无咎也。飞龙在天，大人造也。亢龙有悔，盈不可久也。用九，天德不可为首也。

引理 1.1： 《猫和老鼠》是我最爱看的动画片。

$$\begin{aligned} \int_a^b \left\{ \int_a^b [f(x)^2 g(y)^2 + f(y)^2 g(x)^2] - 2f(x)g(x)f(y)g(y) dx \right\} dy \\ = \int_a^b \left\{ g(y)^2 \int_a^b f^2 + f(y)^2 \int_a^b g^2 - 2f(y)g(y) \int_a^b fg \right\} dy \end{aligned}$$

行行重行行，与君生别离。相去万余里，各在天一涯。道路阻且长，会面安可知。胡马依北风，越鸟巢南枝。相去日已远，衣带日已缓。浮云蔽白日，游子不顾返。思君令人老，岁月忽已晚。弃捐勿复道，努力加餐饭。

定理 1.1: 犯我强汉者，虽远必诛

——陈汤（汉）

$$y = 1 \quad (1-9a)$$

$$y = 0 \quad (1-9b)$$

道可道，非常道。名可名，非常名。无名天地之始；有名万物之母。故常无，欲以观其妙；常有，欲以观其徼。此两者，同出而异名，同谓之玄。玄之又玄，众妙之门。上善若水。水善利万物而不争，处众人之所恶，故几于道。曲则全，枉则直，洼则盈，敝则新，少则多，多则惑。人法地，地法天，天法道，道法自然。知人者智，自知者明。胜人者有力，自胜者强。知足者富。强行者有志。不失其所者久。死而不亡者寿。

证明 燕赵古称多感慨悲歌之士。董生举进士，连不得志于有司，怀抱利器，郁郁适兹土，吾知其必有合也。董生勉乎哉？

夫以子之不遇时，苟慕义强仁者，皆爱惜焉，矧燕、赵之士出乎其性者哉！然吾尝闻风俗与化移易，吾恶知其今不异于古所云邪？聊以吾子之行卜之也。董生勉乎哉？

吾因子有所感矣。为我吊望诸君之墓，而观于其市，复有昔时屠狗者乎？为我谢曰：“明天子在上，可以出而仕矣！” ——韩愈《送董邵南序》 □

推论 1.1: 四川话配音的《猫和老鼠》是世界上最好看最好听最有趣的动画片。

$$V_i = v_i - q_i v_j, \quad X_i = x_i - q_i x_j, \quad U_i = u_i, \quad \text{for } i \neq j; \quad (1-10)$$

$$V_j = v_j, \quad X_j = x_j, \quad U_j u_j + \sum_{i \neq j} q_i u_i. \quad (1-11)$$

迢迢牵牛星，皎皎河汉女。纤纤擢素手，札札弄机杼。终日不成章，泣涕零如雨。河汉清且浅，相去复几许。盈盈一水间，脉脉不得语。

例 1.1：大家来看这个例子。

$$\begin{cases} \nabla f(\mathbf{x}^*) - \sum_{j=1}^p \lambda_j \nabla g_j(\mathbf{x}^*) = 0 \\ \lambda_j g_j(\mathbf{x}^*) = 0, \quad j = 1, 2, \dots, p \\ \lambda_j \geq 0, \quad j = 1, 2, \dots, p. \end{cases} \quad (1-12)$$

练习 1.1：请列出 Andrew S. Tanenbaum 和 W. Richard Stevens 的所有著作。

猜想 1.1: *Poincare Conjecture* If in a closed three-dimensional space, any closed curves can shrink to a point continuously, this space can be deformed to a sphere.

问题 1.1：回答还是不回答，是个问题。

如何引用定理 1.1 呢？加上 \label 使用 \ref 即可。妾发初覆额，折花门前剧。郎骑竹马来，绕床弄青梅。同居长千里，两小无嫌猜。十四为君妇，羞颜未尝开。低头向暗壁，千唤不一回。十五始展眉，愿同尘与灰。常存抱柱信，岂上望夫台。十六君远行，瞿塘滟滪堆。五月不可触，猿声天上哀。门前迟行迹，一一生绿苔。苔深不能扫，落叶秋风早。八月蝴蝶来，双飞西园草。感此伤妾心，坐愁红颜老。

1.5 参考文献

当然参考文献可以直接写 \bibitem，虽然费点功夫，但是好控制，各种格式可以自己随意改写。

本模板推荐使用 **BIB_TE_X**，分别提供数字引用(thuthesis-numeric.bst)和作者年份引用(thuthesis-author-year.bst)样式，基本符合学校的参考文献格式(如专利等引用未加详细测试)。看看这个例子，关于书的^[2-4]，还有这些^[5-7]，关于杂志的^[8-10]，硕士论文^[11-12]，博士论文^[13-14]，标准文件^[15]，会议论文^[16-17]，技术报告^[18]，电子文献^[19-20]。中文参考文献^[21]应增加 language="chinese" 字段，以便进行相应处理。另外，本模板对中文文献^[22]的支持并不是十全十美，如果有不如意的地方，请手动修改 bbl 文件。

有时候不想要上标，那么可以这样 [13]，这个非常重要。

有时候一些参考文献没有纸质出处，需要标注 URL。缺省情况下，URL 不会在连字符处断行，这可能使得用连字符代替空格的网址分行很难看。如果需要，可以将模板类文件中

```
\RequirePackage{hyperref}
```

一行改为：

```
\PassOptionsToPackage{hyphens}{url}
\RequirePackage{hyperref}
```

使得连字符处可以断行。更多设置可以参考 url 宏包文档。

1.6 公式

贝叶斯公式如式 (1-13)，其中 $p(y|\mathbf{x})$ 为后验； $p(\mathbf{x})$ 为先验；分母 $p(\mathbf{x})$ 为归一化因子。

$$p(y|\mathbf{x}) = \frac{p(\mathbf{x}, y)}{p(\mathbf{x})} = \frac{p(\mathbf{x}|y)p(y)}{p(\mathbf{x})} \quad (1-13)$$

论文里面公式越多， \TeX 就越 happy。再看一个 `amsmath` 的例子：

$$\det \mathbf{K}(t = 1, t_1, \dots, t_n) = \sum_{I \in \mathbf{n}} (-1)^{|I|} \prod_{i \in I} t_i \prod_{j \in I} (D_j + \lambda_j t_j) \det \mathbf{A}^{(\lambda)}(\bar{I}|\bar{I}) = 0. \quad (1-14)$$

前面定理示例部分列举了很多公式环境，可以说把常见的情况都覆盖了，大家在写公式的时候一定要好好看 `amsmath` 的文档，并参考模板中的用法：

$$\begin{aligned} \int_a^b \left\{ \int_a^b [f(x)^2 g(y)^2 + f(y)^2 g(x)^2] - 2f(x)g(x)f(y)g(y) dx \right\} dy \\ = \int_a^b \left\{ g(y)^2 \int_a^b f^2 + f(y)^2 \int_a^b g^2 - 2f(y)g(y) \int_a^b fg \right\} dy \end{aligned}$$

其实还可以看看这个多级规划：

$$\left\{ \begin{array}{l} \max_x F(x, y_1^*, y_2^*, \dots, y_m^*) \\ \text{subject to:} \\ G(x) \leq 0 \\ (y_1^*, y_2^*, \dots, y_m^*) \text{ solves problems } (i = 1, 2, \dots, m) \\ \left\{ \begin{array}{l} \max_{y_i} f_i(x, y_1, y_2, \dots, y_m) \\ \text{subject to:} \\ g_i(x, y_1, y_2, \dots, y_m) \leq 0. \end{array} \right. \end{array} \right. \quad (1-15)$$

这些跟规划相关的公式都来自于刘宝碇老师《不确定规划》的课件。

第 2 章 中华人民共和国

2.1 其它例子

在第 1 章中我们学习了贝叶斯公式 (1-13)，这里我们复习一下：

$$p(y|\mathbf{x}) = \frac{p(\mathbf{x}, y)}{p(\mathbf{x})} = \frac{p(\mathbf{x}|y)p(y)}{p(\mathbf{x})} \quad (2-1)$$

2.1.1 绘图

本模板不再预先装载任何绘图包（如 `pstricks` , `pgf` 等），完全由用户来决定。个人觉得 `pgf` 不错，不依赖于 Postscript。此外还有很多针对 \LaTeX 的 GUI 作图工具，如 XFig(jFig), WinFig, Tpx, Ipe, Dia, Inkscape, LaTeXPiX, jPicEdt, jaxdraw 等等。

2.1.2 插图

强烈推荐《 $\text{\LaTeX} 2_{\epsilon}$ 插图指南》！关于子图形的使用细节请参看 `subcaption` 宏包的说明文档。

2.1.2.1 一个图形

一般图形都是处在浮动环境中。之所以称为浮动是指最终排版效果图形的位置不一定与源文件中的位置对应^①，这也是刚使用 \LaTeX 同学可能遇到的问题。如果要强制固定浮动图形的位置，请使用 `float` 宏包，它提供了 `[H]` 参数，比如图 2.1。



图 2.1 利用 Xfig 制图

^① This is not a bug, but a feature of \LaTeX !

大学之道，在明明德，在亲民，在止于至善。知止而后有定；定而后能静；静而后能安；安而后能虑；虑而后能得。物有本末，事有终始。知所先后，则近道矣。古之欲明明德于天下者，先治其国；欲治其国者，先齐其家；欲齐其家者，先修其身；欲修其身者，先正其心；欲正其心者，先诚其意；欲诚其意者，先致其知；致知在格物。物格而后知至；知至而后意诚；意诚而后心正；心正而后身修；身修而后家齐；家齐而后国治；国治而后天下平。自天子以至于庶人，壹是皆以修身为本。其本乱而未治者否矣。其所厚者薄，而其所薄者厚，未之有也！

——《大学》

2.1.2.2 多个图形

如果多个图形相互独立，并不共用一个图形计数器，那么用 `minipage` 或者 `parbox` 就可以。否则，请参看图 2.2，它包含两个小图，分别是图 2.2(a)和图 2.2(b)。推荐使用 `\subcaptionbox`，因为可以像图 2.2 那样对齐子图的标题，也可以使用 `subcaption` 宏包的 `\subcaption`（放在 `minipage` 中，用法同 `\caption`）或是 `subfigure`、`subtable` 环境，像图 2.3，不要再用 `\subfloat`、`\subfigure` 和 `\subtable`。



(a) 第一个小图形



(b) 第二个小图形，注意这个图略矮些。如果标题很长的话，它会自动换行

图 2.2 包含子图形的大图形 (`subcaptionbox` 示例)

古之学者必有师。师者，所以传道受业解惑也。人非生而知之者，孰能无惑？惑而不从师，其为惑也，终不解矣。生乎吾前，其闻道也固先乎吾，吾从而师之；生乎吾後，其闻道也亦先乎吾，吾从而师之。吾师道也，夫庸知其年之先後生於吾乎！是故无贵无贱无长无少，道之所存，师之所存也。

嗟乎！师道之不传也久矣，欲人之无惑也难矣。古之圣人，其出人也远矣，犹且从师而问焉；今之众人，其下圣人也亦远矣，而耻学於师。是故圣益圣，愚



(a) 第一个小图形

清华大学

(b) 第二个小图形，注意这个图略矮些。sub-figure 中同一行的子图在顶端对齐。

图 2.3 包含子图形的大图形 (subfigure 示例)



清华大学
Tsinghua University

图 2.4 并排第一个图



清华大学
Tsinghua University

图 2.5 并排第二个图

益愚。圣人之所以为圣，愚人之所以为愚，其皆出於此乎？爱其子，择师而教之，於其身也，则耻师焉，惑焉。彼童子之师，授之书而习其句读者，非吾所谓传其道、解其惑者也。句读之不知，惑之不解，或师焉，或不焉，小学而大遗，吾未见其明也。巫医、乐师、百工之人不耻相师，士大夫之族曰“师”曰“弟子”之云者，则群聚而笑之。问之，则曰：彼与彼年相若也，道相似也，位卑则足羞，官盛则近谀。呜呼！师道之不复，可知矣。巫医、乐师、百工之人。吾子不齿，今其智乃反不能及，其可怪也欤！圣人无常师。孔子师郯子、苌子、师襄、老聃。郯子之徒，其贤不及孔子。孔子曰：“三人行，必有我师。”是故弟子不必不如师，师不必贤於弟子。闻道有先後，术业有专攻，如是而已。

如果要把编号的两个图形并排，那么小页就非常有用：

李氏子蟠，年十七，好古文、六艺，经传皆通习之，不拘於时，学於余。余嘉其能行古道，作师说以贻之。

——韩愈（唐）

插图索引

图 2.1	利用 Xfig 制图	14
图 2.2	包含子图形的大图形 (subcaptionbox 示例).....	15
图 2.3	包含子图形的大图形 (subfigure 示例)	16
图 2.4	并排第一个图	16
图 2.5	并排第二个图	16
图 A-1	根节点 CU 为 64 x 64、深度为 4 的编码树结构示意图	25
图 A-2	基于树剪枝的 CU 推断方法的步骤	27
图 A-3	测试环境与软件参考配置	28
图 A-4	文中所提出的算法与 HM3.0 在编码时间上的对比: (a) 低延迟方案; (b) 随机访问方案	28
图 B-1	当量化参数为 32 时, 七种模式下零运动块的比例	31
图 B-2	H.264 中的七种预测模式	32
图 B-3	对于 Foreman 序列, 不同准确率的备选阈值 (QP=32)	33
图 B-4	VBZMD 整个过程	34
图 B-5	参考阈值下 VBZMD 的表现	35
图 B-6	当前块与其相邻块	36
图 B-7	预测向量块比例	36
图 B-8	最优阈值下, VBBMD 的表现 (QP=32)	37
图 B-9	在不同条件下 VBBMD 的表现	38

表格索引

表 1.1	模板文件	3
表 1.2	复杂表格示例 1。这个引用 ^[2] 不会导致编号混乱。	3
表 1.3	第一个并排子表格	4
表 1.4	第二个并排子表格	4
表 1.5	并排子表格.....	5
表 1.6	复杂表格示例 2.....	5
表 1.7	实验数据	6

公式索引

公式 1-1	8
公式 1-2	8
公式 1-3	9
公式 1-4	9
公式 1-5	9
公式 1-6	9
公式 1-7	9
公式 1-8	9
公式 1-9a	10
公式 1-9b	10
公式 1-10	10
公式 1-11	10
公式 1-12	11
公式 1-13	12
公式 1-14	12
公式 1-15	13
公式 2-1	14
公式 A-1	26
公式 A-2	26
公式 A-3	26
公式 B-1	32
公式 B-2	32

参考文献

- [1] 薛瑞尼. THUTHESIS: 清华大学学位论文模板[EB/OL]. 2017. <https://github.com/xueruini/thuthesis>.
- [2] Knuth D E. The \TeX book[M]. 15th ed. Reading, MA: Addison-Wesley Publishing Company, 1989
- [3] Goossens M, Mittelbach F, Samarin A. The \LaTeX companion[M]. Reading, MA: Addison-Wesley Publishing Company, 1994: 112-125
- [4] Gröning P, Nilsson L, Ruffieux P, et al. Encyclopedia of nanoscience and nanotechnology: volume 1[M]. [S.l.]: American Scientific Publishers, 2004: 547-579
- [5] Krasnogor N. Towards robust memetic algorithms[M]//Hart W, Krasnogor N, Smith J. Studies in Fuzziness and Soft Computing: volume 166 Recent Advances in Memetic Algorithms. New York: Springer Berlin Heidelberg, 2004: 185-207
- [6] 阎真. 沧浪之水[M]. [出版地不详]: 人民文学出版社, 2001: 185-207
- [7] 班固. 苏武传[M]//郑在瀛, 汪超宏, 周文复. 新古文观止丛书: 第2卷 传记散文英华. 武汉: 湖北人民出版社, 1998: 65-69
- [8] Chafik El Idrissi M, Roney A, Frigon C, et al. Measurements of total kinetic-energy released to the $N = 2$ dissociation limit of H_2 — evidence of the dissociation of very high vibrational Rydberg states of H_2 by doubly-excited states[J]. Chemical Physics Letters, 1994, 224(10): 260-266.
- [9] Mellinger A, Vidal C R, Jungen C. Laser reduced fluorescence study of the carbon-monoxide nd triplet Rydberg series-experimental results and multichannel quantum-defect analysis[J]. J. Chem. Phys., 1996, 104(5): 8913-8921.
- [10] Shell M. How to use the IEEEtran \LaTeX class[J]. Journal of \LaTeX Class Files, 2002, 12(4): 100-120.
- [11] 猪八戒. 论流体食物的持久保存[D]. 北京: 广寒宫大学, 2005.
- [12] Jeyakumar A R. Metamori: A library for incremental file checkpointing[D]. Blacksburg: Virginia Tech, 2004.
- [13] 沙和尚. 论流沙河的综合治理[D]. 北京: 清华大学, 2005.
- [14] Zadok E. FiST: A System for Stackable File System Code Generation[D]. USA: Computer Science Department, Columbia University, 2001.
- [15] IEEE Std 1363-2000. IEEE standard specifications for public-key cryptography[M]. New York: IEEE, 2000

- [16] Kim S, Woo N, Yeom H Y, et al. Design and Implementation of Dynamic Process Management for Grid-enabled MPICH[C]//the 10th European PVM/MPI Users' Group Conference. Venice, Italy: [s.n.], 2003.
- [17] Kocher C, Jaffe J, Jun B. Differential power analysis[C]//Wiener M. Lecture Notes in Computer Science: volume 1666 Advances in Cryptology (CRYPTO '99). [S.l.]: Springer-Verlag, 1999: 388-397.
- [18] Woo A, Bailey D, Yarrow M, et al. The NAS parallel benchmarks 2.0[R/OL]. The Pennsylvania State University CiteSeer Archives, 1995. <http://www.nasa.org/>.
- [19] 萧钰. 出版业信息化迈入快车道[EB/OL]. <http://www.creader.com/news/200112190019.htm>.
- [20] Online Computer Library Center, Inc. History of OCLC[EB/OL]. <http://www.oclc.org/about/history/default.htm>.
- [21] 贾宝玉, 林黛玉, 薛宝钗, 等. 论刘姥姥食量大如牛之现实意义[J]. 红楼梦杂谈, 1800, 224: 260-266.
- [22] 王重阳, 黄药师, 欧阳峰, 等. 武林高手从入门到精通[C]//第 N 次华山论剑. 西安, 中国: [出版者不详], 2006.

致 谢

衷心感谢导师 xxx 教授和物理系 xxx 副教授对本人的精心指导。他们的言传身教将使我终生受益。

在美国麻省理工学院化学系进行九个月的合作研究期间，承蒙 xxx 教授热心指导与帮助，不胜感激。感谢 xx 实验室主任 xx 教授，以及实验室全体老师和同学们的热情帮助和支持！本课题承蒙国家自然科学基金资助，特此致谢。

感谢 L^AT_EX 和 Th_UT_HESIS^[1]，帮我节省了不少时间。

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。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：_____ 日 期：_____

附录 A 书面翻译 1

基于编码树剪枝的 HEVC 快速编码单元推断方法

摘要：本文为 HEVC 提出了一种快速决定编码单元划分的方法，这是一种通过基于树剪枝的策略，来更早推断出编码单元的尺寸的方法。在 HEVC 中，最为有效的是可变的编码单元尺寸，这是一个全新的概念。在推断最佳的编码单元尺寸的过程中，作为参考的 HEVC 编码器会逐个测试每种可能的编码单元尺寸，以便估测每种情况下，由编码单元尺寸所决定的编码器的表现。在整个编码过程中，这是主要的计算复杂度来源，因此，为了实现一个高效快速的编码器，这一过程应是攻坚的重点。一种简单的树剪枝算法是，若当前节点的编码模式已经足够好，那么其子节点的计算均可以被跳过（例如 SKIP 模式）。实验结果显示，本文所提出的方法，相较于测试模型 3.0 下的 HEVC 编码器，最多有可能达到 40 % 的时间缩减，而其代价仅仅是一点可以忽略不计的编码效果上的损失。在第六次 JCT-VC 会议上，本文所提出的方法被 HEVC 编码器测试模型 4.0 所采用。

关键词：HEVC；编码单元早推断；快速高效视频编码；视频编码。

A.1 引言

2010 年，ISO/IEC 旗下的组织 MPEG 和 ITU-T 旗下的 VCEG 拟定了一份新的视频编码标准，该标准被称为 HEVC。这份新标准被期望能够满足人们对于视频日益增长的需求，包括压缩效率、视频分辨率、帧率、计算复杂度等诸多方面。尽管 HEVC 标准尚处于发展阶段，迄今为止，其压缩效率已经达到了前一标准（MPEG-4 AVC/H.264）的二倍之高。

虽然 HEVC 遵循传统的编码结构，包括基于块的运动补偿和变换编码，但实质性改进来自于基于四叉树结构的新的分层编码概念。在 HEVC 中，视频编码和解码过程由以下三个单元组成：编码单元（CU），用于作为变换四叉树的根，以及用于帧间/跳过/帧内预测；预测单元（PU），用于决定模式，包括运动估计和率失真优化；变换单元（TU），用于变换编码和熵编码。在三种编码单元中，CU 对于提升压缩效率是至关重要、最为关键的，因为它决定了初始编码块的大

小，而初始编码块的大小对其他处理单元（例如，PU 和 TU）的性能影响很大。

当使用 CU、PU 和 TU 三种编码单元时，提升压缩效率是可能的，但同时，这也会显著地增加计算的复杂度。通常，在编码过程中会生成一个四叉树，其中根节点表示最大的 CU 大小，并且具有对应于下一个 CU 大小的四个子节点。这棵树被称为编码树，且其最大深度为 4。编码单元与编码树（例如，当根节点的 CU 大小为 64×64 且树的深度设定为 4 的时候，CU0 大小为 64×64 ，CU1 为 32×32 ，...，CU3 为 8×8 。）被包含在 CU 尺寸优化的计算与选择中，这一点导致了 CU 数量的指数增长（例如， $1 \times \text{CU}_0 + 4 \times \text{CU}_1 + 16 \times \text{CU}_2 + 64 \times \text{CU}_3$ ），这一点被直观地展示在 A-1 中。由于每一个 CU 都对应着一个相应的 PU 和 TU 的计算，因此，日益增长的计算复杂对给实时编码器的设计带来了严峻的挑战。HEVC 参考编码器（即，3.0 版）大约比 MPEG-4 AVC/H.264 参考编码器（即，JM 17.0 高配置版本）慢三倍。为了克服 HEVC 编码器在计算上的复杂性，我们提出了基于 CU 早终止的树剪枝策略，并在 JCT-VC 第六会议上介绍了该方法。

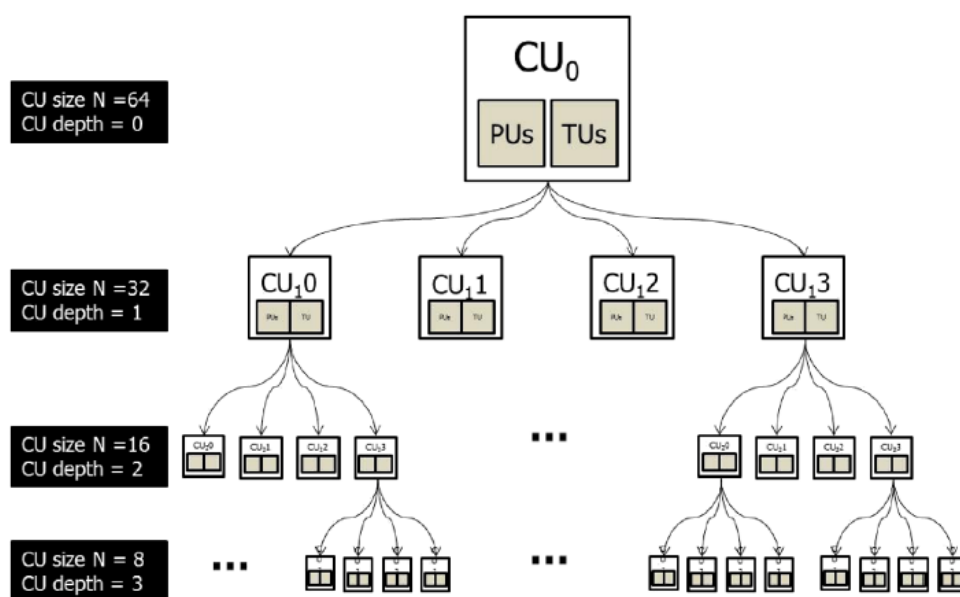


图 A-1 根节点 CU 为 64×64 、深度为 4 的编码树结构示意图

A.2 基于树剪枝的快速编码单元推断

基于可变 CU 大小的编码树推断过程计算复杂度可以描述如下：

$$\begin{cases} f(n) = f(n-1) + 4^n * M_n, f(0) = M_0 \text{ and} \\ M_i = (\frac{1}{4})^i * M_{i-1} \end{cases} \quad (\text{A-1})$$

这里， $f(n)$ 是当编码树的最大深度被设定为 n 时所需操作的总数，而 M_i 则是对于第 i 层给定的 CU 大小所需的操作总数。在 A-1 中，我们假定 M_i 是 M_{i-1} 的四分之一是因为 CU 的大小比起前一层会减少四分之一。总的操作数可以被整合为如下的公式：

$$f(n) = \sum_{i=0}^n 4^i * M_i \cong O(M_0 * n) \quad (\text{A-2})$$

如 A-2 所展示，计算的复杂度随着 CU 深度的增加单调增加。当考虑到 M_i 代表着包括对每个 CU 大小的运动估计在内的整个编码过程这一事实，最大 CU 深度就自然成为了决定编码时间的主要因素。

本文研究了一种通过更早地确定 CU 的深度来加速 CU 深度推断的策略。我们利用了这样一个事实，当当前 CU 节点的代价低于该节点子节点的代价时，我们就不必再对子树进行进一步的处理（例如， $RDCost(CU_t) < \sum_{i=0}^3 RDCost(CU_{t-1}^i)$ ）。这一方法的唯一问题是，子树的代价必须是已知的，这一要求妨碍了子树计算复杂度的减少程度。我们可以避免子树代价的计算，如果当前节点的代价已经是最小的了（例如，SKIP 模式），而这正是本方法的核心概念。为了进一步减少计算复杂度，我们可以定义剪枝的条件如下：

Pruning condition :

$$(m' = \underset{m \in Mode}{argmin} RDCost(CU_t | PU = m)) \leq Threshold \quad (\text{A-3})$$

$$Mode = \{SKIP, Inter2Nx2N, Inter2NxN, InterNx2N, InterNxN\}$$

这里，Mode 表示按根据复杂度确定的预测模式的有序集合，而 m' 则是为当前 CU 深度选择的预测模式，且 Threshold 将根据 Mode 选择。在 A-3 中，通过从 SKIP 到 InterNxN，不断改变 Threshold 的值，剪枝过程的可能性将被推断。

若 Threshold 被设置为 SKIP 时，剪枝过程将不会对压缩效率产生影响；否则，以降低压缩效率为代价，更多的子树将被剪去。根据上述分析，我们提出了 ?? 所示的剪枝过程。

```

Algorithm proposed CU decision
Recursive_CU_Processing(depth,index){
    parent.cost = CU_processing(depth,index)
    if (selected prediction_mode ≤ Threshold
        Best.CU = CU(depth)
        pruning remaining processes
    else
        for from index = 0 to to index = 3 do
            children.cost+ = Recursive_CU_Processsing(depth + 1,
            index)
        end
        if(parent.Cost ≤ children.cost)
            Best.CU = CU(depth)
        else
            Best.CU = CU(depth+1)
        if(leaf node)
            return
        }
    }

```

图 A-2 基于树剪枝的 CU 推断方法的步骤

A.3 实验结果

为了进行性能评测，我们测试了我们的方法与 HM 3.0 两种方法的总执行时间，以确认计算复杂度的降低程度。编码的表现以 $\Delta Bitrate[(B_{PRO} - B_{REF})/B_{REF} \times 100]$ 和 $\Delta PSNR(P_{PRO} - P_{REF})$ 来衡量，并且时间的减少以 $\Delta Time[(T_{PRO} - T_{REF})/T_{REF} \times 100]$ 来衡量。在实验中，我们把 Threshold 值设置为 SKIP 模式。有关编码环境的更多细节在 A-3 被展示。

Test Sequences	<ul style="list-style-type: none"> Class A (2560 × 1600): Traffic and People On Street Class B (1920 × 1080): Kimono, ParkScene, Cactus, and BQTerrace Class C (832 × 480): BasketballDrill, BQMall, PartyScene, and RaceHorsesC Class E (416 × 240): BasketballPass, BQSquare, BlowingBubbles, RaceHorses Class E (1280 × 720): Vidyo1, Vidyo3, and Vidyo4
Total Frames to be Coded	<ul style="list-style-type: none"> Class A: 5 seconds of video duration Other Classes: 10 seconds of video duration
Software	<ul style="list-style-type: none"> HM 3.0
Quantization Parameter	<ul style="list-style-type: none"> 22, 27, 32 and 37
Others	<ul style="list-style-type: none"> High efficiency setting

图 A-3 测试环境与软件参考配置

(a)	HM 3.0			Proposed Method			Comparison		
Class	Bitrate (kb/s)	PSNR (dB)	Time (ms)	Bitrate (kb/s)	PSNR (dB)	Time (ms)	Bitrate (%)	PSNR (dB)	Time (%)
Class B Average	10,587	36.74	697.29	10,553	36.72	450.76	-0.34	-0.02	-38
Class C Average	3,803	35.04	141.45	3,794	35.02	104.81	-0.33	-0.02	-29
Class D Average	1,031	34.55	34.58	1,027	34.52	25.86	-0.44	-0.03	-28
Class E Average	1,862	40.19	246.60	1,847	40.16	104.67	-0.78	-0.03	-58
		Average					-0.44	-0.03	-37

(b)	HM 3.0			Proposed Method			Comparison		
Class	Bitrate (kb/s)	PSNR (dB)	Time (ms)	Bitrate (kb/s)	PSNR (dB)	Time (ms)	Bitrate (%)	PSNR (dB)	Time (%)
Class A Average	14,423	37.11	1,116.5	14,363	37.05	690.30	-0.61	-0.06	-41
Class B Average	9,620	36.74	548.42	9,564	36.70	300.70	-0.55	-0.04	-47
Class C Average	3,533	35.09	112.66	3,523	35.04	74.15	-0.51	-0.05	-37
Class D Average	942	34.78	27.46	940	34.74	18.51	-0.44	-0.04	-36
		Average					-0.52	-0.05	-41

图 A-4 文中所提出的算法与 HM3.0 在编码时间上的对比：(a) 低延迟方案；(b) 随机访问方案

A-4 展示了在低延迟方案与随机访问方案两种环境下，我们评测了我们的方法与 HM 3.0 两种方法的总执行时间。尽管我们测试了每一种 A-3 中所列举的不同的量化参数下的结果，但 A-4 中，我们只展示每一类的平均结果。A-4 显示了我们的方法相较于 HM 3.0，在低延迟方案的环境下减少了 63.34% 的时间，而

在随机访问方案的环境下，则减少了 59.43 % 的时间。以图像质量的轻微退化为代价所获得的编码增益，是所提出的方法的自然结果，因为编码树修剪过程会使图像质量上的比特减少。通过调整 **Threshold** 的值，我们可以节约更多的编码时间，但其代价是图片质量的进一步下降。目前，我们所提出的方法，在模式设置为 **SKIP** 时，能得到编码时间与图片质量相平衡的最好结果。

A.4 结论

在本文中，我们为 HEVC 提出了一种快速的 CU 决策方法，这种方法基于编码树修剪从而更早确定 CU 的大小，与 HM 3.0 编码器相比，仅以可以忽略不计的 BD-比特率损失为代价，就可以使编码时间减少约 40 %。实验结果表明，当你设计一个快速的 HEVC 编码器时，我们所提出的编码树剪枝方法应当被考虑。在 JCT-VC 第六次会议上，HEVC 测试模型 4.0 编码器采用了我们所提出的方法。

附录 B 书面翻译 2

一种有效的 H.264 上的可变块大小早终止算法

摘要： H.264 视频编码标准比以前的标准提供了更高的编码效率，但同时，其复杂度也显著提高。在 H.264 编码器中，最耗时的组成部分是可变块大小的运动估计。为了降低运动估计的复杂度，本文提出了一种早终止算法。它通过检查仅仅一个搜索点来预测最佳的运动矢量。使用该方法，一些运动搜索可以被提前停止，由此，大量的搜索点可以被跳过。本文所提出的方法适用于任何快速运动估计算法。实验采用的是 H.264 所采用的一种快速运动估计算法。结果表明，本算法显著降低了复杂度，而代价仅仅是可以忽略不计的视频质量的降低。

关键词： 早终止；H.264；运动估计；可变块大小。

B.1 引言

近年来，视频编码技术高速发展。理所当然地，压缩性能的提升伴随着计算代价的增长。作为 ITU-T 的 VCEG 和 ISO/IEC 的 MPEG 所最新制定标准的 H.264，相较于之前的 H.261、H.263、MPEG-1/2/4 等标准，在编码效率上有着巨幅的提升。然而，其过高的计算复杂度使其无法被广泛应用到实时业务中。

通常来说，一个视频编码器最为耗时的部分是运动估计。减少其计算代价的方法有两种。第一种是加速算法本身，对 ME 过程来说，数种快速算法被提出，例如六边形搜索（HBS）、增强型预测分区搜索（EPZS）和混合非对称交叉多六边形网格搜索（UMHexagonS）等。另一种方法就是更早地终止 ME 的计算。通过预测那些 DCT 系数会被量化为 0 的块，一些方法有效地减少了 ME 的计算量。另一方面，一些块的重要部分在 ME 之后具有零运动矢量（MV）。据此，有算法提出的零运动检测算法（ZMD），通过将它们的差值的绝对值求和，并与预先定义的阈值进行比较来检测这样的块，然后跳过其余的搜索点。

然而，上述早终止方法都是针对以前的编码方案（如 H.263）而开发的。它们不能再应用于 H.264。这是因为相比于仅有两种块大小（16x16 和 8x8）的 H.263，H.264 拥有从 16x16 到 4x4 变化的七种块大小。

通过扩展 ZMD 的概念，我们提出了一种用于 H.264 视频编码的可变块大小最佳运动检测（VBBMD）算法。该方法在三个方面与 ZMD 不同：

1. 基于不同块大小的检测精度，分别获得 VBBMD 中使用的阈值，同时考虑了复杂度的降低，使得阈值的决策更加合理；
2. 在 VBBMD 中，除零运动点外，预测的移动到的点也会被检查，从而更多的搜索点可以被跳过；
3. 对于 16x16 的块，VBBMD 使用双阈值。使用较低的阈值跳过较小的块的运动搜索。

我们的方法可以适用于任何运动搜索算法，并且不需要额外的计算。注意，整个讨论集中于帧间帧（P 帧）和整数像素运动搜索。文章的剩余部分如下进行组织。B.2 介绍了早期终止算法及其改进。B.3 中给出了仿真结果。最后，在 B.4，我们总结了本文，并给出了未来的方向。

B.2 早终止算法

B.2.1 可变块大小零运动推断

通常来说，对大多数序列，在 ME 后相当数量的块会有这零运动矢量，正如 B-1 展示的一样。

ZMB RATES (%) OF SEVEN MODES (QP = 32)							
Sequence	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6	Mode 7
Akiyo	98.03	97.51	97.18	97.28	97.14	96.47	97.04
Salesman	96.80	95.92	95.69	95.15	94.78	94.13	94.33
News	91.90	91.72	91.98	92.35	91.93	91.31	91.78
Silent	85.01	84.65	84.15	85.01	84.75	84.23	85.06
Coastguard	32.44	33.51	32.80	30.91	31.40	30.71	31.06
Foreman	43.97	44.15	43.57	44.25	44.81	43.73	45.42

图 B-1 当量化参数为 32 时，七种模式下零运动块的比例

表中，模式 1-7 代表从 16x16 到 4x4 的七种帧间预测模式，如 B-2 所展示。

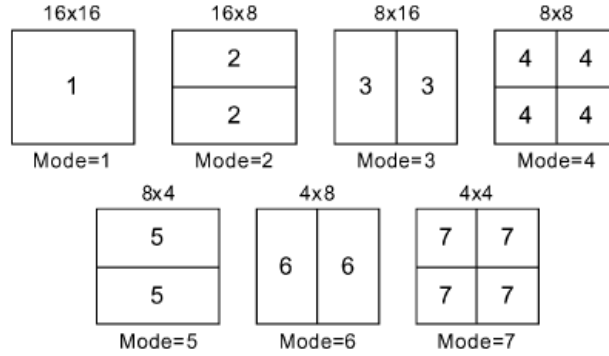


图 B-2 H.264 中的七种预测模式

可以看出，不论是运动较为活跃亦或是不太活跃的视频序列，有 30.71 % 到 98.03% 的块存在零运动。如果我们可以预测零运动块（ZMB），我们可以提前停止 ME 的计算，从而消除计算的一部分。通过扩展 ZMD 的概念，我们开发了一个可变块大小的 ZMD（VBZMD）算法。

在以前的 ZMD 方法中，当块的 SAD 小于给定阈值时，块可以被看作是 ZMB。在 H.264 中选择最佳匹配块时，使用代价函数 J 而不是 SAD 作为预测误差的度量。

$$J(m, \lambda) = SAD(s, c(m)) + \lambda R(m - p) \quad (\text{B-1})$$

这里， $m = (m_x, m_y)^T$ 是当前的 MV， $p = (p_x, p_y)^T$ 是预测的 MV，而 λ 则是拉格朗日乘数， $R(m - p)$ 是编码 MV 所用的比特数。

如果一个块是零运动的，那么 $MV(0, 0)$ 处极有可能有一个很小的代价值。因此，我们为七种模式分别定义阈值 $THZ_i (i = 1, \dots, 7)$ 。在整个运动估计过程中， $MV(0, 0)$ 首先被检查。如果 $MV(0, 0)$ 处的代价满足 B-2，我们就认为这个块是一个 ZMB，由此剩余的搜索过程既可以被停止。

$$J_i < THZ_i, \text{ for } i = 1, \dots, 7. \quad (\text{B-2})$$

H.264 中 ZMD 的关键是如何去决定不同块大小下的阈值。很明显，阈值越大，检测到的 ZMB 越多，更多的搜索点就可以被跳过。然而同时，更多的块被

错误地选择，这会导致更大的图像质量损失。因此，在性能和复杂性之间存在一个平衡。在实践中，防止视频质量的损失可能比复杂度的轻微增加更重要。

因此，我们将检测精度作为确定阈值的指导。这里，精度表示当代价小于阈值时，块为 ZMB 的概率。我们的方法是，在实验中基于不同的准确率获得多个候选阈值集，并选择一个最优集合，在实践中，我们需要在质量和复杂度之间取一个较好的折衷。许多序列的实验表明，在相同的阈值下，低运动场景中的准确率和检测率都高于有较多运动的场景中的准确率和检测率。由于 Foreman 序列可以代表具有较大运动的场景，所以我们可以选择基于 Foreman 的阈值，并将它们应用到其他序列。B-3 显示了我们获得的候选阈值。

CANDIDATE THRESHOLDS FOR DIFFERENT ACCURACY RATES OF FOREMAN (QP = 32)							
Accuracy	Mode1	Mode2	Mode3	Mode4	Mode5	Mode6	Mode7
80%	600	550	570	400	315	300	250
75%	960	650	660	480	360	350	270
70%	1650	920	890	600	410	400	300
65%	2270	1200	1160	700	470	460	350
60%	2800	1500	1430	880	530	530	400

图 B-3 对于 Foreman 序列，不同准确率的备选阈值 (QP=32)

在我们的方法中，如果块的零运动代价小于相应的阈值，则我们将其视为块的最佳 MV 并跳到同一宏块中的下一个块。如果当前块是宏块中的最后一个块，则跳转到下一个模式。在所有模式被检查之后，拥有最小成本的模式被选择，并应用于后续的任一为 H.264 设计的 ME 算法。在模式 1 中，由于一个宏块只包含一个块，如果零运动成本足够小，那么不仅可能是最好的 MV，而且模式 1 也很可能是宏块的最佳模式。因此，我们为模式 1 定义了额外的低阈值 THS。如果零运动代价小于 THS，则整个宏块的运动搜索停止。VBZMD 的整个过程概述如下：


```

Assume the current macroblock is MB( $x, y$ )
For each mode  $i$  ( $i = 1, \dots, 7$ ) of the MB( $x, y$ )
  For each block in the macroblock
    Calculate the cost at (0, 0)
    If cost < THZ $i$ 
      Set MV = (0, 0)
      If mode = 1 and cost < THS
        Set best mode = 1
        Exit two loops
      End if
      Exit one loop
    Else
      Do the normal motion search process
    End if
  End for
End for
If best mode is not set
  Set best mode to be the mode with the minimal cost

```

图 B-4 VBZMD 整个过程

我们所使用的 H.264 编码器是 6.1e 版本的 JVT 软件。我们将所提出的方法与 H.264 标准所采用的快速 ME 算法 UMHexagonS 相结合。与完全搜索相比，快速 ME 方法可以将整数像素 ME 的复杂度降低 90%，而通过使用所提出的方法可以进一步减少计算。因为 800 的阈值对应于模式 1 中相对较高的 78% 的准确率，并且它在实验中似乎工作得很好，所以我们选择它作为 THS。正如 B-5 所展示的，对于代表低运动场景的 Akiyo 序列，每个宏块（PPMB）的搜索点中，有高达 93.47% 的部分被减去，而其代价仅为平均峰值信号（PSNR）降级不超过 0.05 分贝，且比特率的增加是微不足道的。即使对于具有较大的面部运动和摄像机平移的 Foreman 序列，计算量的节省也是相当明显的，且 PSNR 丢失依旧轻微。这里，搜索点的数目是在宏块级别上计算的。例如，模式 4 中，大小为 8×8 的块中的一个搜索点，等同于 1/4 PPMB。这里，对应于 65% 的准确率的阈值集在性能和复杂度之间提供了良好的折衷。Akiyo 和 Foreman 序列的 PSNR 退化分别为 0.04 和 0.13，而计算量的节省分别为 91.9% 和 59.64%。

PERFORMANCE OF VBZMD WITH CANDIDATE THRESHOLDS (QP = 32)

Sequence	Method	Accuracy	PSNR (dB)	PPMB	Bit-rate (kbps)
Akiyo	Fast ME		35.22	63.86	14.52
	Fast ME +VBZMD	80%	35.18	29.02	14.65
		75%	35.20	21.41	14.67
		70%	35.19	7.96	14.70
		65%	35.18	5.17	14.59
		60%	35.17	4.17	14.67
Foreman	Fast ME		32.90	160.48	79.34
	Fast ME +VBZMD	80%	32.87	124.97	79.00
		75%	32.86	112.85	79.55
		70%	32.82	88.75	80.49
		65%	32.77	64.77	82.19
		60%	32.75	49.39	84.73

图 B-5 参考阈值下 VBZMD 的表现

直观地，高精度相对应的 PSNR 应该大于与低精度相对应的 PSNR。然而，在这个表中可以看出，Akiyo 序列中，精度为 75% 对应的 PSNR 值和 70% 对应的 PSNR 值，比精度 80% 对应的 PSNR 高。这种情况是可能的，因为较大的阈值不仅减少了更多的计算，而且减少了编码 MV 所需的比特，然后节省的比特提高了总体 PSNR。

B.2.2 可变块大小最佳运动推断

在减少计算量上，尽管 VBZMD 已经取得了相当好的表现，但我们仍可以将其改进以得到更好的表现。在 H.264 的中，当前块的左、上和右上（或上左）相邻块的 MV 将用作预测的当前块的 MV 等，如 B-6 所说明。由于相邻块的 MV 通常是相关的，预测向量很可能正是在 ME 之后最好的 MV。为方便描述，我们定义最佳 MV 为预测向量的块为预测向量块（PVB）。正如 B-7 所示，不同序

列的 PVB 率序列是大于相应的、如 B-1 所示的 ZMB 率得。特别是，对于运动较为激烈的场景，如 Coastguard 和 Foreman 序列，PVB 率较 ZMB 率高 23% 到 122%。另一方面，另一个实验显示，对大多数序列来说，PVB 率较 ZMB 率高 97%。我们所观察到的现象启示我们，应当使用预测向量来预测最佳 MV，而不是零 MV。此外，在 H.264 中，进行编码和传输是当前 MV 和预测矢量之间的差异，而不是当前的 MV 本身。因此，使用预测向量来预测最佳 MV 可以节省用于编码 MV 的比特。

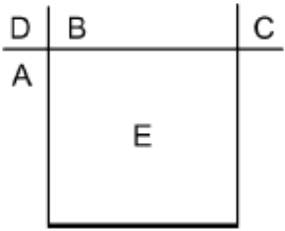


图 B-6 当前块与其相邻块

Sequence	Mode	Mode	Mode	Mode	Mode	Mode	Mode
	1	2	3	4	5	6	7
Akiyo	98.05	97.51	97.23	97.37	97.27	96.61	97.24
Salesman	96.84	95.71	95.46	95.40	95.11	94.54	95.03
News	91.80	91.70	91.66	92.83	92.71	92.22	93.16
Silent	85.09	85.04	84.57	86.32	86.80	86.36	88.28
Coastguard	58.68	62.19	62.41	64.32	66.17	67.09	69.04
Foreman	54.19	56.26	55.46	57.79	60.00	59.15	63.15

图 B-7 预测向量块比例

基于前面的 VBZMD，我们提出了一种 VBBMD 算法。关键思想是将预测向量的成本与给定阈值进行比较。如果成本小于阈值，则将预测向量视为最佳 MV，然后可以跳过剩余的搜索点。

同样，如何选择不同块大小情况下的阈值仍然是问题的关键。使用与

VBZMD 中相同的方法，我们可以在不同的准确率下获得多个候选阈值集。VBBMD 的步骤与 VBZMD 几乎相同，唯一的区别在于，我们检查预测矢量而不是零 MV。我们用候选阈值集编码不同的视频序列，并选择最佳集合。结果表明，阈值的集合（2500, 1500, 1400, 920, 625, 570, 500）在准确度为 66% 时提供了一个很好的折衷。此外，模式 1 的低阈值 THS 也是 800，因为它工作得很好。B-8 展示出了具有最佳阈值集的 VBBMD 结果。与 VBZMD 在同一表中的结果相比，我们可以看到，更多的搜索点被进一步减少，而 PSNR 和比特率几乎一样。这验证了 VBBMD 比 VBZMD 更有效。

Sequence	Method	PSNR (dB)	PPMB	Bit-rate (kbps)
Akiyo	Fast ME	35.22	63.86	14.52
	Fast ME+VBZMD	35.18	5.17	14.59
	Fast ME+VBBMD	35.18	4.17	14.59
Foreman	Fast ME	32.90	160.48	79.34
	Fast ME+VBZMD	32.77	64.77	82.19
	Fast ME+VBBMD	32.77	30.12	82.34

图 B-8 最优阈值下，VBBMD 的表现（QP=32）

如前所述，阈值是在 $QP = 32$ 的条件下确定的。然而实验表明，不同的 QP 值将得到不同的阈值。这是因为代价值 J 与 QP 有关。在 B-1 中， J 的值与 λ 直接相关，而 λ 的值是从一个以 QP 为索引的表中查询得到的。当一个块被视为 PVB 时，当前 MV m 与预测 MV p 相差最多为 $3/4$ 像素，这可以用 10 比特来编码。因此，我们近似地设置 $R(m - p)$ 为 10。我们定义 $QP = 32$ 条件下的阈值为标准阈值，对于其他的 QP 取值，相应的阈值应设置为标准阈值加上 $(\lambda_{QP} - \lambda_{32}) \times 10$ 。

B.3 模拟结果

我们将 VBBMD 实现在了 H.264 参考编码器上，并使用 UMHexagonS 快速 ME 算法。我们设置运动搜索范围为 32，并设置参考帧数为 1。率失真优化被禁止并且 CAVLC 熵编码开启。

为了验证所提出的方法在不同实验条件下的有效性，我们选择了六个序列，其运动活动从小到大变化。他们是 Akiyo, Salesman, News, Silent, Coastguard 和 Foreman 序列。所有序列均为 QCIF 格式，并以 30 帧每秒编码。除第一帧以外的所有帧都被编码为 P 帧。此外，为了检查在不同比特率下的性能，在我们的实验中使用了三个 QP 值：28, 32, 36。如 B.2.2 所述，在 VBBMD 中使用的阈值根据 QP 进行调整。

为便于比较，我们在 B-9 中显示了平均 PSNR 增益 ($\Delta PSNR$)、PPMB 平均降低率 ($\Delta PPMB$) 和比特率的增长率 ($\Delta Bit-rate$)。从结果可以看出，所提出的方法是稳定的，并且在不同的条件下工作得很好。与仅使用快速 ME 算法相比，我们的方法进一步减少了 76.10% 到 94.12% 的整数像素搜索点，而 PSNR 衰减平均仅为 0.06 分贝，并且比特率只是略微增加。

QP	Performance	Akiyo	Sales- man	News	Silent	Coast- guard	Fore- man
28	$\Delta PSNR$ (dB)	-0.05	-0.02	-0.06	-0.04	-0.05	-0.12
	$\Delta PPMB$ (%)	94.12	90.22	83.30	81.95	87.67	81.76
	$\Delta Bit-rate$ (%)	0.04	1.60	1.33	2.42	3.72	4.60
32	$\Delta PSNR$ (dB)	-0.03	-0.02	-0.06	-0.04	-0.06	-0.13
	$\Delta PPMB$ (%)	93.39	91.07	83.48	83.46	86.37	80.61
	$\Delta Bit-rate$ (%)	1.03	0.80	1.26	2.95	4.63	3.43
36	$\Delta PSNR$ (dB)	-0.05	-0.06	-0.08	-0.04	-0.06	-0.13
	$\Delta PPMB$ (%)	92.63	90.43	82.17	83.72	77.46	76.10
	$\Delta Bit-rate$ (%)	0.66	0.37	1.41	-0.07	3.00	3.45

图 B-9 在不同条件下 VBBMD 的表现

B.4 结论

本文提出了一种 H.264 视频编码的早终止方法 VBBMD，该方法通过进一步减少计算总量来帮助现有的快速运动搜索算法。该方法通过检查一个搜索点来预测最佳的 MV。通过这种方法，一些运动搜索可以被提前停止，因此与这些搜索相关联的计算可以被减少。在该方法中，考虑了复杂度的降低的同时，阈值通过基于不同块大小的检测精度来确定，这使得阈值决策更加合理。实验结果表明，我们提出的方法实现了显著的复杂度降低，而视频质量的退化是可以忽略不计的。未来的方向可能包括使阈值根据量化级别和运动级别自适应调整，并将早终止方法应用于子像素的运动估计。