译 x264 编码器选项分析 (x264 Codec Strong and Weak Points) 1

2014年09月02日 12:55:38　阅读数：6252

文章目录：

==========

本文简单翻译了MSU实验室做的X264的Option（即编码选项，后文称其英文名）分析报告《x264 Codec Strong and Weak Points》。看了之后感觉分析得十分透彻，而且其采用的方法也很有参考价值，因此记录一下其中的要点以作备忘。众所周知，X264的Option非常繁多，如何搭配这些繁多的参数以让编码后的视频体积小，质量好，同时编码速度快，确实是一个很让人头疼的问题。本报告中的实验就是为了解决上述问题而设计的。

一．简介

本文通过使用客观质量评价算法分析X264编码器的选项。下表列出了文中用到的几个词汇。

	Term	Definition	Example
1.	***Option***	Option is the codec parameter. Codec has a number of options.	number of B-frames, motion estimation algorithm, etc.
2.	***Option value***	Each option has a set of option values. Option value influence on the speed and quality of encoding process.	"—me" option (motion estimation algorithm) has values "dia", "hex", "umh", "esa" and "tesa"
3.	***Preset***	Preset is a set of options with fixed values. If option is missing in presets description, its value is equal to default one.	--me 'dia', --ref 4, --subme 6
4.	***Pareto-optimal point (presets)***	Preset is called pareto-optimal, if there are no other presets that simultaneously give better quality and work faster on given sequences. Number of pareto-optimal presets can be selected for each sequence.	see Picture 6
5.	***Envelope line points (presets)***	Presets lying on the convex hull. It corresponds to the best presets (when the ratio λ between relative encoding time and bitrate is fixed) for all possible ratio λ. See Picture 16.	see Picture 7
6.	***Parameter λ***	Represents desired ratio between relative time and bitrate. Common measure of preset quality can be defined as $M=\lambda T+Q$, where T is relative encoding time and Q is relative encoding quality (see section "Preset Analysis Method" for more details.	

简单翻译一下：

名称	定义	例子
Option（选项）	编码选项	B帧数，运动估计算法
Option value（值）	编码选项的值	--me（运动估计算法）包含以下值：”dia”, “hex”, “umh”, “esa”, “tesa”
Preset（预设）	由一系列固定Value（值）的Option（选项）组成的集合	
Pareto-optimal Point (Preset)（帕累托最优 的预设）	没有任何其他Preset（预设）可以比该Preset获得更好的视频质量和更快的编码速度。	
Envelope line Points (Preset)（包络线上的预设）	位于Convex hull（凸包）上的预设。代表了最好的预设。	
Parameter λ（λ参数）	代表期望的编码时间和码率的比值。可以用下式表示： M=λT+Q 其中T代表了相对编码时间，Q代表了相对的视频质量。	

二．选项（Option）和选项的值（Option value）

本报告分析的X264的Option如下所列。

	Option	Option Values	Comments
1	**Partitions** --partitions x (where 'x' is the partition search types)	"none" "p8x8,b8x8,i8x8,i4x4" "all"	These options determine the partition search types. Default value is "p8x8,b8x8,i8x8,i4x4".
2	**B-Frames** --bframes n (where 'n' is the number of B-frames)	0 2 4	Selects the number of consecutive B-frames between I and P x264 should use. B-frames are frames that are small in size, but when placed correctly, quality loss is insignificant. This can help improve compression effectiveness. Default value is 0.
3	**Reference Frames** --ref n (where 'n' is the number of reference frames)	1 4 8	Selects the maximum number of reference frames that can be used. Reference frames are the frames that refer to other frames (i.e. if both frames are similar) from which they may be predicted. Having a high number of referenced frames will improve quality but slow down encoding. Default value is 1.
4	**Motion Estimation Method** --me x (where 'x' is the motion estimation method)	"dia" "hex" "umh" "tesa"	This option selects the way motion is detected. Motion estimation is a technique to reduce temporal redundancy of a video sequence, and thus it improves compression ratio. It tracks differences between scenes to allocate the various frame types and bitrates. **Diamond (dia)**: Diamond search, radius 1. It has maximum encoding speed. **Hexagon (hex)**: hexagonal search, radius 2. It has worse speed and better quality then the diamond search. **Multi Hex (umh)** (also known as "Uneven Multi-Hexagon"): It is tradeoff between speed and quality. **Hadamard exhaustive(tesa)**: Hadamard exhaustive search. It is slowest method. Default value is "hex".
5	**Subpixel Motion Estimation** --subme n (where 'n' is the estimation value)	1 4 5 6	Also known as "Partition Decision". A very important option that determines how x264 makes decisions about motion estimation. The options are available from 1 to 7, with 1 being the fastest (lowest quality) and 7 being the slowest (best quality). Default value is 5.
6	**Mixed References** --mixed-refs (enables mixed references)	off on	This option allows x264 to have greater control over "Reference Frames". Option only available when at least two reference frames has been set. Default value is "off".
7	**Weighted Prediction** --weightb (enables weighted prediction)	off on	Turns on weighted prediction for B-frames, which results in improved accuracy and therefore in more efficient encoding. Option only available when at least two B-frame has been set. Default value is "off".

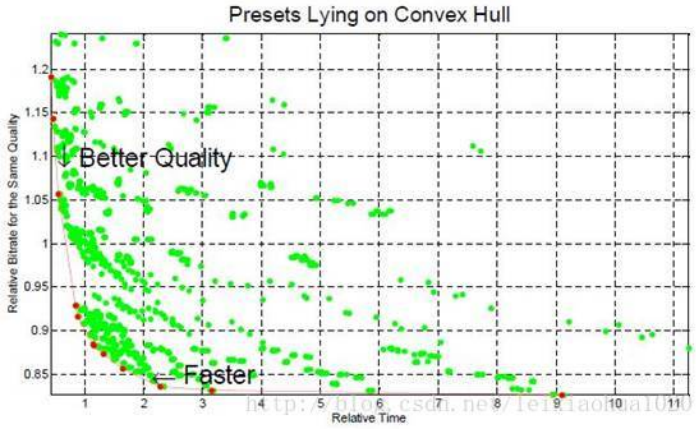
简单翻译一下：

选项	值	备注
Partitions（分块方式） --partitions x	"none" "p8x8,b8x8,i8x8,i4x4" "all"	宏块使用的分块方式。 默认值： "p8x8,b8x8,i8x8,i4x4"
B-Frames（B帧数） --bframes n	0 2 4	I帧和P帧之间连续B帧出现的数量。 默认值：0
Reference Frames（参考帧数） --ref n	1 4 8	参考帧的数量。 默认值：1
Motion Estimation Method（运动估计方法） --me x	“dia” “hex” “umh” “tesa”	运动估计的方法。详细见注释。 默认值：“hex”
Subpixel Motion Estimation（子像素运动估计） --subme n	1 4 5 6	子像素(subpixel)估测复杂度。 默认值：5
Mixed References --mixed-refs	off on	默认值：off
Weighted Prediction --weightb	off on	每一帧对B帧的影响力与其和该B帧的距离相关。 默认值：off

注：
1.运动估计方法简介
dia(diamond菱形搜索) 是最简单的搜索方式，从最优预测值出发，往上、左、下、右一个像素处检测运动向量，挑选最好值，然后重复该步骤，直至找不到更优的运动向量。
hex(hexagon正六角形搜索) 的策略类似，但它对周围六个点进行range-2搜索，因此称为正六角形搜索。此方法效率大大高于dia，且速度相当，因此通常编码常用此项。
umh(uneven multi-hex不对称多六角形搜索)比hex慢很多，但能搜索复杂的多六角形，以避免错过很难找到的运动向量。与hex和dia相似，M.E.半径范围参数直接控制umh的搜索半径，使用者可自行增减搜索的空间尺寸。
esa(exhaustive全面搜索) 在最优预测值附近M.E.半径范围范围内的整个空间内，以高度优化的智能方式搜索运动向量。相当于数学上的穷举法，搜索区域内的每一个运动向量，但是更快些。然而，此方法远远慢于UMH，且好处不多，对于普通编码没有太大用处。
tesa (transformed exhaustive变换全面搜索)算法尝试对各个运动向量近似哈达玛变换比较法。与exhaustive类似，但效果略好，速度略慢。
2. 子像素运动估计
子像素(subpixel)估测复杂度，越大越好。数值1-5单纯控制子像素细化强度。数值6会开启模式决策RDO，数值8将开启运动向量和内部预测模式RDO。RDO模式大幅慢于低级模式。
采用低于2的值，会使用一种较快、但较低质量的lookahead模式，同时会影响--scenecut的决策，因此不推荐。
可选值：
0. fullpel only
1. QPel SAD 1 iteration
2. QPel SATD 2 iterations
3. HPel on MB then QPel
4. Always QPel
5. Multi QPel + bi-directional motion estimation
6. RD on I/P frames
7. RD on all frames
8. RD refinement on I/P frames
9. RD refinement on all frames
10. QP-RD (requires --trellis=2, --aq-mode > 0)
11. Full RD

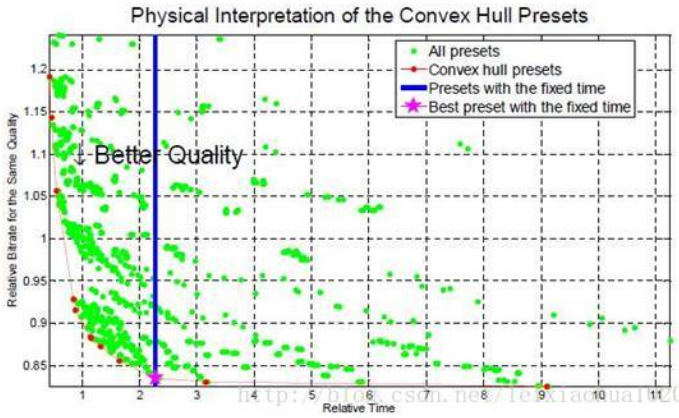
三、最佳Preset

下图显示了本次实验的全部结果。实验系统通过枚举上表所列参数的所有选项，而后进行压缩编码得到的实验结果。每一个绿色的点代表一种Preset压缩的结果，可见实验的数据量很大。



下面简单解释一下这张图的含义。纵坐标代表视频的码率，取值越低，代表同等视频质量的的情况下码率越小。横坐标代表视频的编码时间，取值越低，代表编码时间越低。坐标取值是一个相对值。每一个绿色的点代表一种Preset。因此可知，越是位于左下角的Preset，其编码速度越快，且码率越低。最优的Preset应该是位于凸包(Convex hull)上的点(即那条红颜色线上的点)。需要注意的是，横坐标和纵坐标都是一个相对值而不是绝对的码率和时间。横坐标和纵坐标的值都是相比于X264的Default Preset而言的。X264的Default Preset即X264全部使用默认值的Preset，位于这张图的（1,1）点处。

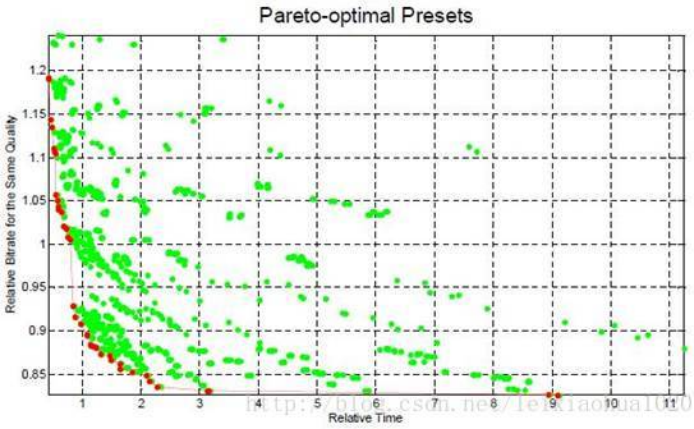
如果固定编码时间，则最优的Preset位于包络线上。如图所示，位于粉红色五角星位置的Preset在相同的编码时间下码率最低。



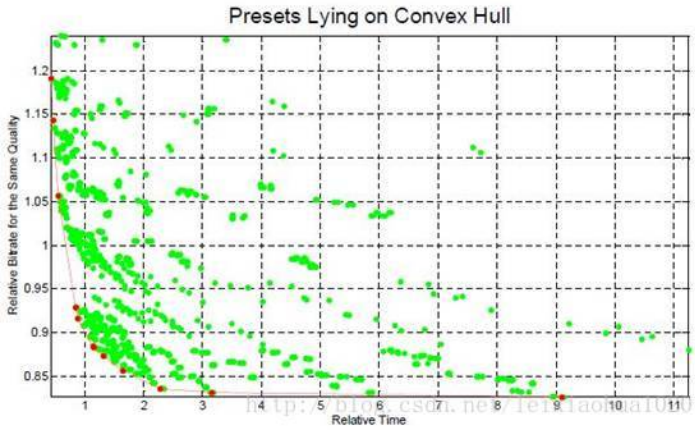
X264默认值如下表所示。

Option	默认值
Partitions	“p8x8,b8x8,i8x8,i4x4”
B-Frames	0
Reference Frames	1
Motion Estimation Method	“hex”
Subpixel Motion Estimation	5
Mixed References	off
Weighted Prediction	off

下图显示了Pareto-optimal Point（帕累托最优点）Preset（即没有任何其他Preset可以比该Preset获得更好的视频质量和更快的编码速度）以及凸包（Convex hull）上的Preset。



凸包（Convex hull）上的Preset数据统计如下。



凸包（Convex hull）上的Preset分析结果如下。表格中一方面列出了凸包（Convex hull）上的Preset使用较多的Option以及使用较少的Option。另一方面列出了消耗时间较长但是质量较高的Preset以及速度较快但是质量较差一些的Preset。

Table 4. List of Convex Hull Presets.

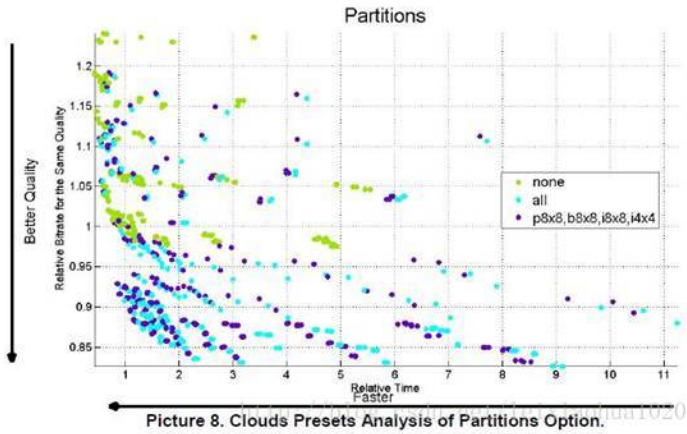
	Time	Bitrate	--partition	--b-frames	--ref	--me
1.	0.423525	1.191290	"none"	0	1	"dia"
2.	0.460319	1.143350	"none"	2	1	"dia"
3.	0.553577	1.056690	"p8x8,b8x8,i8x8,i4x4"	2	1	"dia"
4.	0.844114	0.928896	"p8x8,b8x8,i8x8,i4x4"	2	1	"dia"
5.	0.882376	0.916337	"p8x8,b8x8,i8x8,i4x4"	2	1	"hex"
6.	1.311870	0.873373	"all"	2	1	"umh"
7.	1.145420	0.884237	"p8x8,b8x8,i8x8,i4x4"	2	4	"hex"
8.	1.154520	0.883289	"p8x8,b8x8,i8x8,i4x4"	2	4	"hex"
9.	1.646830	0.856413	"all"	2	1	"umh"
10.	2.275270	0.835395	"all"	2	4	"umh"
11.	3.154320	0.830835	"all"	2	8	"umh"
12.	9.091010	0.826391	"all"	2	8	"tesa"

Table 5. List of Convex Hull Presets (Continuation).

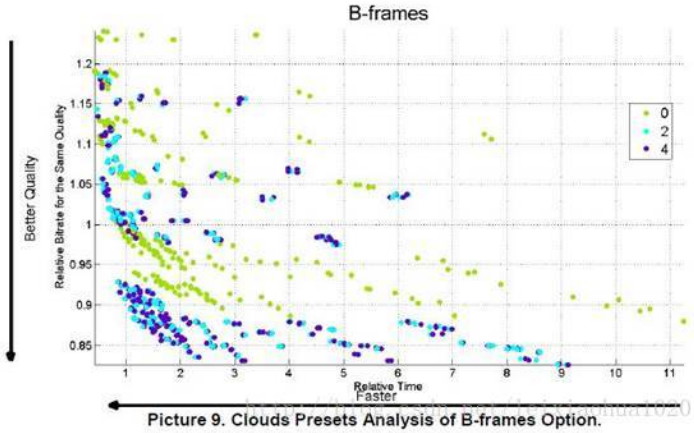
	Time	Bitrate	--subme	--mixed-refs	--weightb[1]
1.	0.423525	1.191290	1	off	off
2.	0.460319	1.143350	1	off	off
3.	0.553577	1.056690	1	off	off
4.	0.844114	0.928896	4	off	off
5.	0.882376	0.916337	4	off	off
6.	1.311870	0.873373	4	off	off
7.	1.145420	0.884237	4	on	on
8.	1.154520	0.883289	4	on	off
9.	1.646830	0.856413	6	off	off
10.	2.275270	0.835395	6	on	off
11.	3.154320	0.830835	6	on	off
12.	9.091010	0.826391	6	on	off

四.
“彩云图”的Preset分析

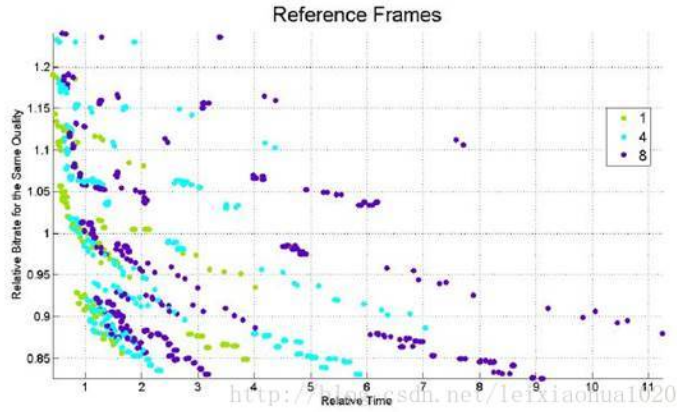
本章中，每张图对应一种关注的Option，包含不同Option Value的 Preset被标记为不同的颜色。实验结果如下所示。

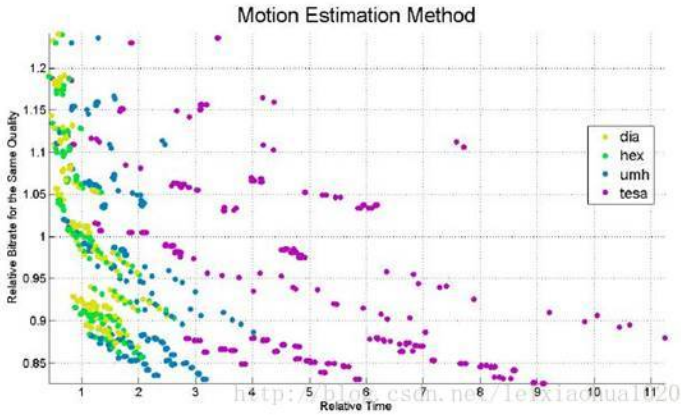


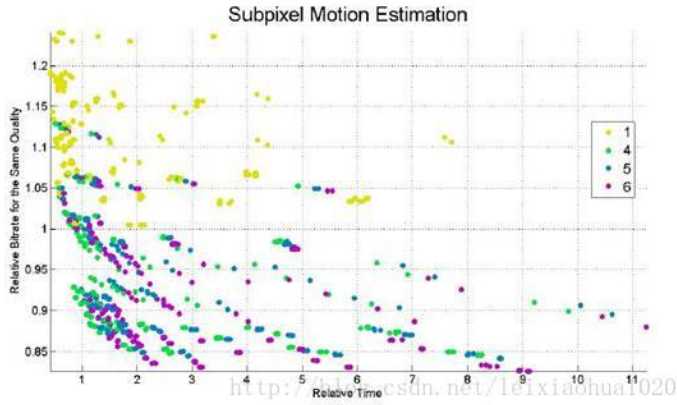
Picture 8. Clouds Presets Analysis of Partitions Option.

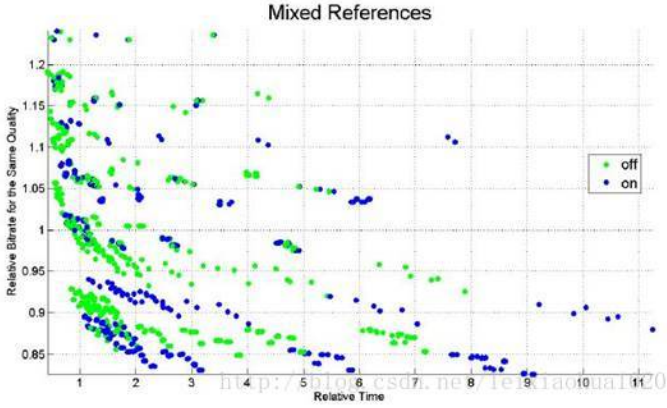


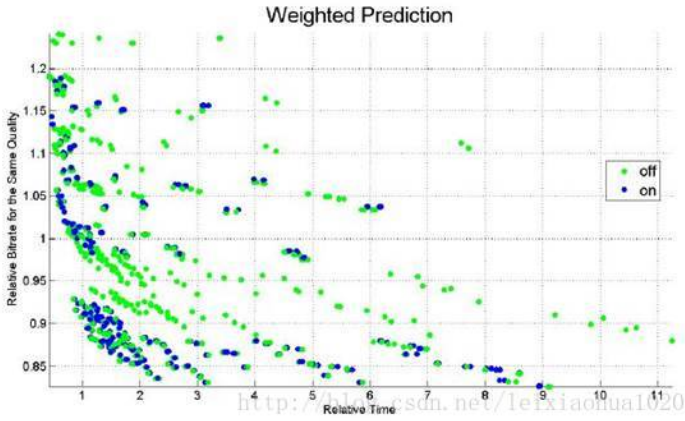
Picture 9. Clouds Presets Analysis of B-frames Option.


Reference Frames
Relative Bitrate for the Same Quality
Relative Time
1
4
8


Motion Estimation Method
Relative Bitrate for the Same Quality
Relative Time
dia
hex
umh
tesa


Subpixel Motion Estimation
Relative Bitrate for the Same Quality
Relative Time
1
4
5
6


Mixed References
Relative Bitrate for the Same Quality
Relative Time
off
on



从以上图表可以得出的各个Option的分析结果如下表所示。

	Option	Preset	Comments
1	**Partitions** --partitions x	• "none" • "all" • p8x8,b8x8,i8x8,i4x4"	Presets with partitions equal to "none" works well when encoding speed is high. Partitions value "all" is the best when it is required high quality. If it is important both speed and quality the best choice is "p8x8,b8x8,i8x8,i4x4".
2	**B-Frames** --bframes n	• 0 • 2 • 4	If maximum encoding speed is required the best choice is 0. In other cases the 2 or 4 B-frame is better. Values 2 and 4 of option B-frames does not significantly differ.
3	**Reference Frames** --ref n	• 1 • 4 • 8	Presets with 1 reference frame is better when the speed is more important than quality. If speed is not the most important factor, but still important 4 reference frames are more preferable. 8 reference frames is optimal when maximum quality is required.
4	**Motion Estimation Method** --me x	• "dia" • "hex" • "umh" • "tesa"	Presets with "dia" and "hex" algorithms are optimal if you want to get high speed. "umh" algorithm is a good tradeoff between speed and quality. "tesa" algorithm is optimal when maximum quality is required.
5	**Subpixel Motion Estimation** --subme n	• 1 • 4 • 5 • 6	Presets with subme 5 are not optimal. Among the best presets with high speed all have subme 1. High quality presets with sumbe 6 have better speed than high quality presets with other subme value. If it is important both speed and quality the best choice is subme 4.
6	**Mixed References** --mixed-refs	• off • on	Presets with turned off mixed references are optimal for high speed encoding. If the maximum quality is required the best choice is to use mixed references.
7	**Weighted Prediction** --weightb	• off • on	Optimal presets have both values of weighted prediction option. There are slightly more presets with weighted prediction "off" among the best presets according to the maximum speed values.

简单翻译一下：

选项	值	结论
Partitions（分块方式） --partitions x	"none" "p8x8,b8x8,i8x8,i4x4" "all"	要求速度快的时候用"none"，要求视频质量高的时候用"all"。要求速度质量均衡考虑的时候使用"p8x8,b8x8,i8x8,i4x4"
B-Frames（B帧数） --bframes n	0 2 4	要求速度极高的时候用"0"。其他情况下用"2"，"4"，它们二者之间区别不大。
Reference Frames（参考帧数） --ref n	1 4 8	要求速度快的时候用"1"，要求视频质量极高的时候用"8"。要求速度质量均衡考虑的时候使用"4"
Motion Estimation Method（运动估计方法） --me x	"dia" "hex" "umh" "tesa"	要求速度快的时候用"dia"和"hex"，要求视频质量极高的时候用"tesa"。要求速度质量均衡考虑的时候使用"umh"。
Subpixel Motion Estimation（子像素运动估计） --subme n	1 4 5 6	要求速度快的时候用"1"，要求视频质量高的时候用"6"。要求速度质量均衡考虑的时候使用"4"
Mixed References --mixed-refs	off on	

Weighted Prediction --weightb	off on	

注1：“Weighted prediction”作用不明显
注2:“bframes”中的2和 4 B-frames差别不大。
“彩云图”的分析结论如下表所示。该表列出了不同的使用环境下Option应该使用的值（Option value）。分成三种情况：速度最重要，速度和视频质量平均考虑，视频质量最重要。

	Option	Time Is More Important than Quality	Time/Quality tradeoff	Quality Is More Important than Time
1.	--partitions	“none”	“p8x8,b8x8,i8x8,i4x4”	“all”
2.	--bframes	0	2, 4	2, 4
3.	--ref	1	4	8
4.	--me	“dia”, “hex”	“umh”	“tesa”
5.	--subme	1	4	6
6.	--mixed-refs	off	off, on	on

原文地址：

http://compression.ru/video/codec_comparison/pdf/x264_options_analysis_08.pdf

文章标签： x264 编码 优化 选项

个人分类： 视频编码 x264