编码器根据视频图像帧是由两个隔行场合并而成还是只是一个单个的逐行帧，为这个帧标记Progressive\_frame旗标，旗标的值为Ture或False。如果是True，则这个帧为逐行扫描帧，如果是False，则这个帧是隔行扫描帧。

一帧图像送入到编码器，先经过帧内或帧间预测之后，得到预测值与输入数据的残差，然后进行DCT变化和量化，得到残差系数，然后送入熵编码模块输出码流，同时，残差系数经反量化反变换之后，得到重构图像的残差值，然后和帧内或者帧间的预测值结合，从而得到了重构图像，重构图像再经环内滤波之后，进入参考帧队列，作为下一帧的参考图像，从而一帧帧向后编码。

All pictures that have been decoded that may be used as References for prediction of any subsequent pictures in decoding order must be included in the RPS(Reference Picture Set).

一个参考图像集（RPS）包含一系列用于标识DPB中的图像的POC值。

**GOP Coding Structure**

为了应对不同场合，HEVC设立了GOP的三中编码结构，即帧内编码，低延迟编码，随机访问编码。这三种编码结构在编码结构配置文件中显示了各种结构的参数。

如图所示，帧内编码给出了两个参考的配置文件，main main10 这两个配置文件在参数上的区别，前面一个比特深度为8，后一个为10.（还有别的区别，但是我还不知道）

由帧内编码结构的定义知道，GOP内所有的预测编码都选择帧内预测编码（I帧），

从图中可以看出coding structure中Intraperiod 的值为1，说明所有的帧都选择帧内预测。

同理可以知道低延迟的参数中intraperiod的值应该为-1，即第一帧为I帧，其余的为P B帧。

将下好的HM9.1解压后，发现配置文件有很多：encoder\_intra\_he10.cfg、encoder\_intra\_main.cfg、encoder\_lowdelay\_he10.cfg、encoder\_lowdelay\_main.cfg、encoder\_lowdelay\_P\_he10.cfg、encoder\_lowdelay\_P\_main.cfg、encoder\_randomaccess\_he10.cfg、encoder\_randomaccess\_main.cfg，下面会详细地介绍它们之间的差别。

1. encoder\_xxx\_he10.cfg与encoder\_xxx\_main.cfg之间的差别在于InternalBitDepth不同，main中该参数设置为8，he10中该参数设置为10. 当输入的码流是8bit采样时，若InternalBitDepth设置为10，则每一个样点通过乘以4（即左移两位）变成10bit采样点。当输入的码流是10bit采样时，若InternalBitDepth设置为8，那么每个采样点的值通过公式（x\_2）/4进行下采样。
2. encoder\_intra\_xxx.cfg则是只包含I帧、不包含B帧和P帧的配置文件。
3. encoder\_lowdelay\_xxx.cfg和encode\_lowdelya\_P\_xxx.cfg的差别在于，前者只包含I帧和B帧，而后者是只包含I帧和P帧。
4. encoder\_lowdelay\_xxx.cfg和encoder\_randomaccess\_xxx.cfg的差别在于，randomaccess里面的B帧是分层的B帧。

GOP size is designed to specify the pyramid temporal structure of hierarchical-B structure. Currently, encoder is implemented with assuming the dyadic pyramid, thus GOP size should be a power of 2, e.g. 1, 2, 4, 8, so on.

Non-pyramidal coding (e.g. IBBP) or non-dyadic pyramidal coding can be easily supported by similar fashion to JM. Maybe this kind of functional improvements can be done after some major integration works are finished according to the suitable approval from the SW AhG.

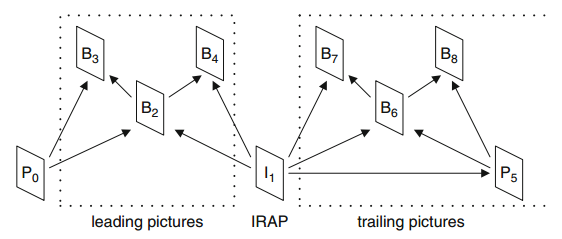
deltaRPS为predictor的POC值减去当前POC的值，deltaRPS的作用是根据当前POC和deltaRPS寻找predictor

每个图像的RPS包含5个不同的参考图像表：

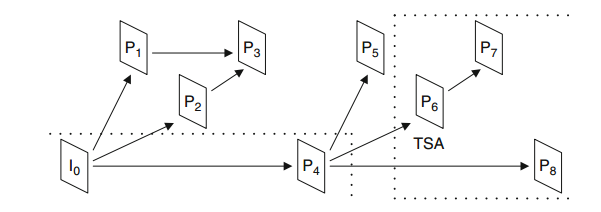
* RefPicSetStCurrBefore：解码和输出顺序均在当前帧之前的短线参考帧，可用于当前帧的帧间预测；
* RefPicSetStCurrAfter：解码顺序在当前帧之前，输出顺序在当前帧之后的短线参考帧，可用于当前帧的帧间预测；
* RefPicSetStFoll：按解码顺序在当前帧之后的图像可用的短线参考帧，不能用于当前帧的帧间预测；
* RefPicSetLtCurr：可用于当前帧的帧间预测的长线参考帧；
* RefPicSetLtFoll：按解码顺序在当前帧之后的图像可用的长线参考帧，不能用于当前帧的帧间预测；

HEVC中的新结构：

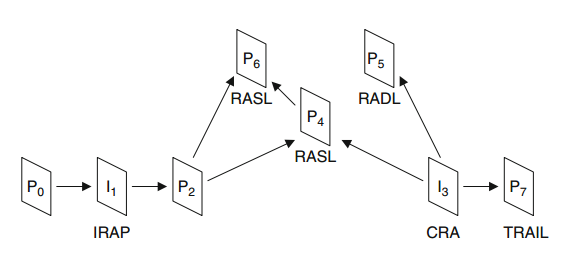
1. 三种基本的图像类型：IRAP、Leading Pics、Trailing Pics



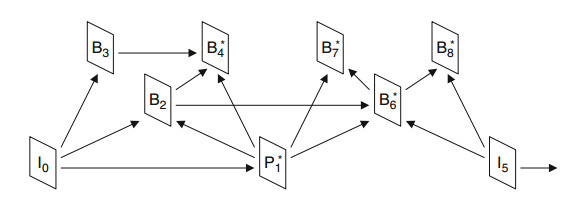
* + 是指使用帧内编码，不参考其他图像。所以IRAP一定是TID为0，编码方式为intra，但是编码方式为intra不一定是IRAP。IRAP解码时不需要依靠其他参考帧，所以在视频随机访问、频道切换、视频编辑等操作时，有重要作用。
  + LPs：解码顺序上在IRAP之后，但是在输出顺序上在IRAP之前的图像
  + Trailing Pics：解码和输出顺序都在IRAP之后的图像，分为以下三种：



* + 1. Temporal Sub-layer Access （TSA）picture，时域分层编码，定义了时域分层切换点；
       1. 一个TSA帧需要满足：TSA帧和其编码顺序（Decoding Order）上随后的TSA帧子序列 不会参考TSA帧之前的、TID大于等于TSA帧的任何帧，如上图，P6和P7就是TSA帧。
    2. Step-wise temporal sub-layer access（STSA）picture，步进时域分层编码。STSA帧要满足STSA帧及其编码顺序后面的相同TID的STSA帧不会参考该STSA帧编码顺序前面、有相同TID的帧。如上图，P2属于STSA帧。
    3. Ordinary Trailing picture（TRAIL）
  + 要求：Trailing pictures只能参考相关的IRAP和相关于同一个IRAP的Trailing pictures，且对于同一个IRAP，必须按照IRAP🡪Leading pictures🡪Trailing pictures的顺序进行编解码
  + IDR：Instantaneous Decoding Refresh，即刻解码刷新帧，解码时彻底清空解码缓冲区，更新解码过程并重新开始一段新的CSV（coding sequence video），因此IDR及其之后的所有帧都不允许参考IDR之前的帧。根据IDR帧是否有可解码的Leading pictures，IDR可分为IDR\_W\_RADL和IDR\_N\_RASL，IDR\_W\_RADL在解码顺序上有RADL pictures，而IDR\_N\_RASL帧没有任何Leading pictures。对于IDR，POC一定是0。
  + CRA：Clean Random Access，对于CRA帧，它的Leading pictures帧可以参考CRA帧之前的帧，所以解码CRA帧时，不会刷新解码器，也不会重新开始一段新的CSV。对于CRA，POC不一定是0。支持开放GOP操作。



* + Random Access Decodable Leading (RADL) picture：RADL帧是IRAP帧的Leading pictures。对于RADL帧，它只能参考关联的IRAP帧和对应的RADL帧，不能参考关联的IRAP帧解码顺序前面的帧。在随机接入点后，可以按顺序解码的图像。随机接入点的图像可能是IDR/CRA/BLA的图像。
* Random Access Skipped Leading (RASL) pictures，RASL帧是CRA帧的Leading pictures。对于RASL帧，它可以参考关联的CRA帧解码顺序前面帧，因此IDR只能有RADL的Leading pictures，而CRA可以有RADL和RASL的Leading pictures。而且，对于CRA帧，RASL要在RADL之前解码。
* Broken Link Access (BLA) pictures：针对视频链接所定义的IRAP类型，因为CRA类型相对IDR类型有更高的编码效率，所以视频序列中更多地使用CRA。但是由于CRA的Leading pictures可以是RASL，所以当从CRA帧开始进行视频拼接、访问时，RASL需要参考CRA编码顺序之前的帧，但是这些帧是无法获得的，所以就定义成BLA帧。遇到BLA帧时，对于其编码顺序后与其关联的RASL，直接舍弃。与IDR帧类似，BLA帧解码时从新开始一个新的CSV，但是不同之处在于：BLA并不会把POC设置为0（IDR会），而是设置成BLA header中的值。有三种类型的BLA：BLA\_N\_LP、BLA\_W\_RADL和BLA\_W\_LP，分别代表禁止所有Leading pictures、仅禁止RASL的Leading pictures和RADL、RASL的Leading pictures都允许。



* Sub-Layer reference VS. Sub-layer non-reference：对于Leading 和Trailing，每一个种类都有\_R和\_N两种类型，分别表示sub-Layer reference和sub-layer non-reference。sub-Layer non-reference是指不能够被相同Temporal Layer的帧参考的帧，如上图中B3-4、B6-8和P1；sub-layer reference是指可以被相同Temporal Layer参考的帧，如上图B2。这些在对视频帧进行选择性丢弃时，可以根据这些帧类型进行判断决策。
* SEI：supplemental enhancement informatyion，辅助增强信息，提供可选的解码支持元数据。在HEVC中，SEI是prefix（SEI必须在Access unit的所有VCL NAL之前）或者suffix（SEI可以在Access unit的VCL NAL之后）的，而且SEI有些是只对当前Access unit有效，有些作用范围可能是多个Access unit甚至整个SVC。VUI是在SPS中的可选信息，VUI不直接影响解码过程，但是提供两个方面的信息：1、解码图像的展示信息，包括宽高比、扫描、分时等信息；2、限制解码端的一些信息，包括tiles、MV、参考图像等。