

HEVC 变换和量化

变换是视频编解码中用来去除块内部相关性的方法。在 H.264/AVC 中，变换在 main profile 中只能使用 4x4 DCT 变换，在 high profile 中则可以对 8x8 以上的块选择是否使用 8x8 DCT 变换。也就是说，一个 8x8 以上的块有两种可选的变换方式。在 HEVC 中，一个块可以通过二叉树的形式划分成若干个 TU，每个 TU 的大小可以从 4x4 到 32x32。一个 16x16 的块，可以对前三个块用 8x8 变换，最后一个块划分成 4 个 TU，使用 4x4 的变换。与 CU 类似，TU 的二叉树也增加了编码的 RDO 开销，同时对解码的开销影响较小。需要指出的是复杂度的问题。与预测不同，假设一个方块的边长为 n，预测是对每个点算一遍，复杂度是 $\Theta(n^2)$ ，所以大块和其分成的小块解码复杂度相同。但是对于变换，做的是矩阵乘法，复杂度是 $\Theta(n^3)$ （忽略哪些阶更小但常数项特别大的复杂算法吧），所以使用大块解码复杂度是要更高的。

HEVC 在变换上除了使用更大的变换，另一个显著特点就是每个尺寸的块除了 DCT 变换，在 4x4 的 intra 亮度块的时候还会使用 DST 变换。笔者也不知道为什么使用 DST，有高手知道的话还请赐教。HEVC 中变换的系数比 H.264/AVC 复杂不少，不再是只有 1 或 2 了。另外，HEVC 中的 DCT 系数虽然有 4x4 到 32x32 这 4 组，但是都可以从 32x32 的系数数组中取出，这是为了在硬件实现中节约 ROM。

HEVC 中新加入了 transform skip 技术。在人造序列（如动画片、游戏视频等）中，图像平缓区域无变化、边界区域剧烈变化的趋势，预测之后的残差与自然图像有明显的区别。这些残差本来数量较少，但是值很大。经过变换，这些残差会形成大量的高频信息，反而造成了冗余。transform skip 技术在编码端决定是否进行变换，在上述现象发生是不进行变换，从而使得人造序列的压缩率得到了极大地提升。

变换完的系数需要量化，然后扫描成一维数组进行熵编码（当然先扫描再量化也行）。在 H.264/AVC 中，扫描方式只有 zig-zag，而在 HEVC 中，根据模式不同，如图 5-1 所示，使用水平、竖直、对角三种方式之一。对于超过 4x4 的块，先将块分成若干个 4x4 的 CG（Coefficient Group），CG 之间和每个 CG 内部都使用选定的扫描顺序进行扫描。

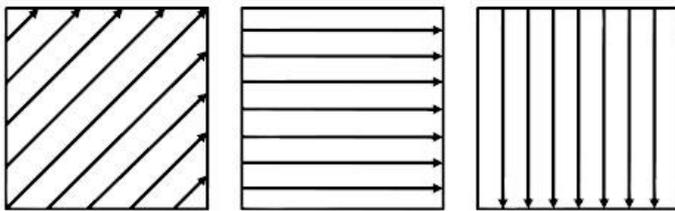


图 5-1

对于量化，有一个很有意思的新技术和读者们分享一下，叫做 sign data hiding，该技术也被认为是一种熵编码技术。其在编码端发生在量化过程中（也可认为是量化之后），解码端发生在熵解码过程中。对于一个扫描过的 CG，也就是一个 16 个元素的一维系数数组，如果其第一个非零元素和最后一个非零元素的距离大于等于 4，则用该 CG 的系数之和的奇偶性决定第一个元素的符号。在编码过

程中，可以通过将某一个系数增减 1 来改变 CG 系数和的奇偶性。这一过程发生在每次量化的时候。该技术增加了编码的复杂性，解码的复杂性没有影响。