

第二章 多媒体基础

周满

15271802577

zhouman@hust.edu.cn

华中科技大学网络空间安全学院

视频多媒体基础

- 一、视频压缩原理
- 二、运动估计
- 三、视频压缩标准H.264
- 四、视频压缩标准H.265

视频压缩的可能性

- ◆ 视频数据中存在着大量的冗余，即图像的各像素数据之间存在极强的相关性。利用这些相关性，一部分像素的数据可以由另一部分像素的数据推导出来，结果视频数据量能极大地压缩，有利于传输和存储



视频压缩的可能性

◆ 1. 空间冗余

- ◆ 视频图像在水平方向相邻像素间、垂直方向相邻像素间的变化一般都很小，存在着极强的空间相关性。特别是同一景物各点的灰度和颜色间往往存在着空间连贯性，从而产生了空间冗余，常称为帧内相关性



视频压缩的可能性

◇ 2. 结构冗余

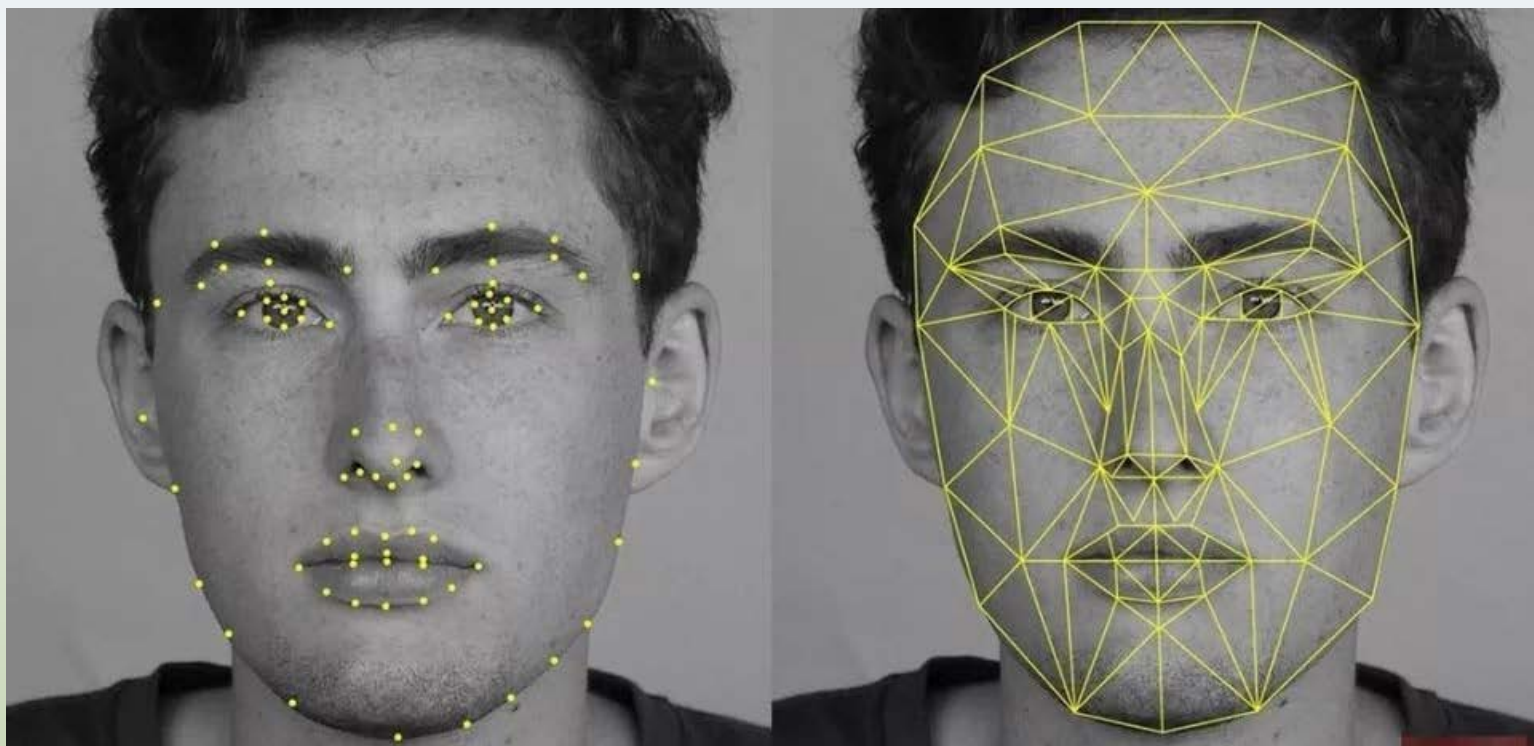
- ◇ 在有些图像的纹理区，图像的像素值存在着明显的分布模式，如方格状的地板图案等。已知分布模式可以通过某一过程生成图像，称为结构冗余



视频压缩的可能性

◆ 3. 知识冗余

- ◆ 有些图像与某些知识有相当大的相关性。如人脸的图像有固定的结构，嘴的上方有鼻子，鼻子的上方有眼睛，鼻子位于脸部图像的中线上。这类规律性的结构可由先验知识得到，此类冗余称为知识冗余



视频压缩的可能性

◆ 4. 视觉冗余

- ◆ 人眼具有视觉非均匀特性，对不敏感信息可适当舍弃。
摄像头在记录原始图像数据时，通常对视觉敏感和不敏感部分同等对待，从而产生了比理想编码更多的数据。
而人眼对图像细节、幅度变化和图像运动并非同时具有最高的分辨能力



视频压缩的可能性

◆ 5. 图像区域的相似性冗余

- ◆ 在图像的两个或多个区域对应的所有像素值相同或相近，从而产生数据重复性存储，这就是图像区域的相似性冗余。在这种情况下，记录一个区域中各个像素的颜色值，与其相同或相近的区域就不再记录各个像素值



视频压缩的可能性

◆ 6. 纹理的统计冗余

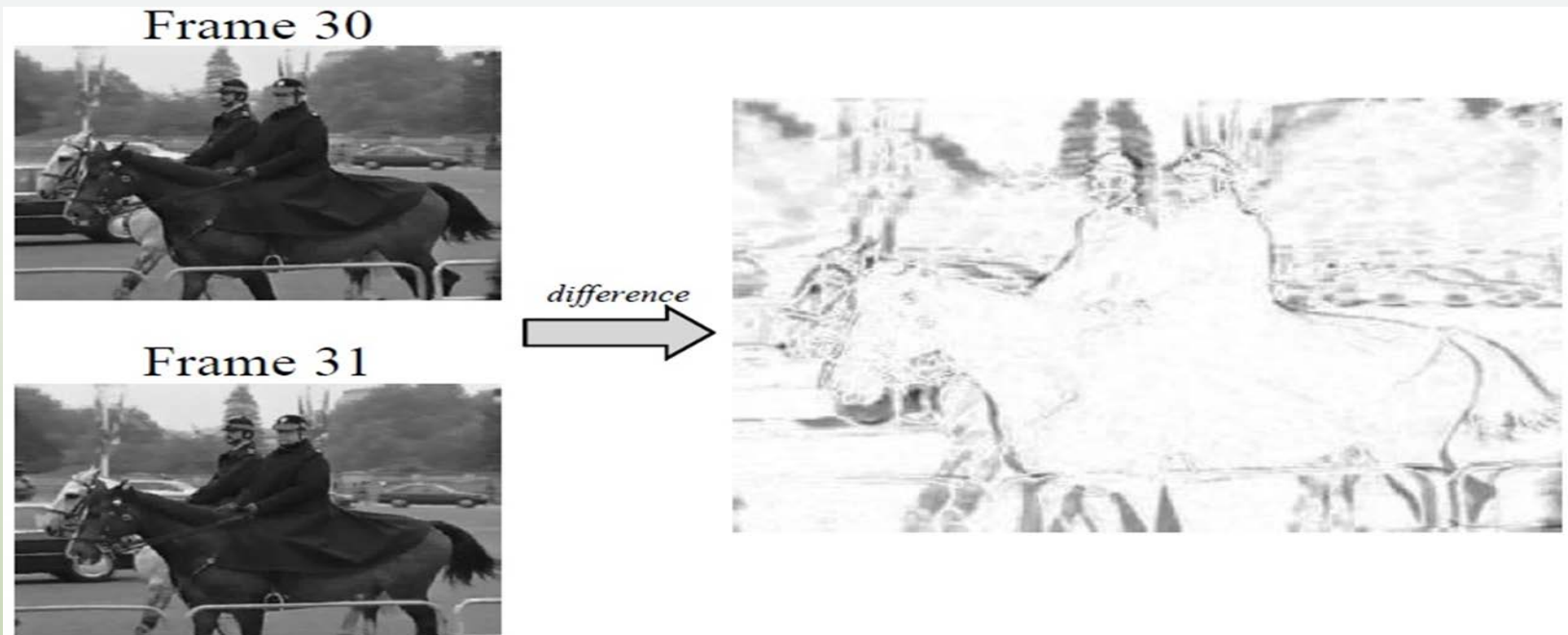
- ◆ 有些图像纹理尽管不严格服从某一分布规律，但是在统计的意义上服从该规律，利用这种性质也可以减少表示图像的数据量，称为纹理的统计冗余



视频压缩的可能性

◆ 7. 时间冗余

- ◆ 视频图像在相邻帧的对应像素间，亮度和色度信息存在极强的相关性。当前帧图像往往具有与前后两帧图像相同的背景和移动物体，只不过移动物体所在的空间位置略有不同。大多数像素的亮度和色度信息基本相同，称为帧间相关性或时间相关性



运动估计-时间冗余

如何去除时间冗余?

- ◆ 计算前一帧和当前帧的像素值的差异，如果两帧的像素值差异很小，可以直接跳过



运动估计-时间冗余

如何去除时间冗余?

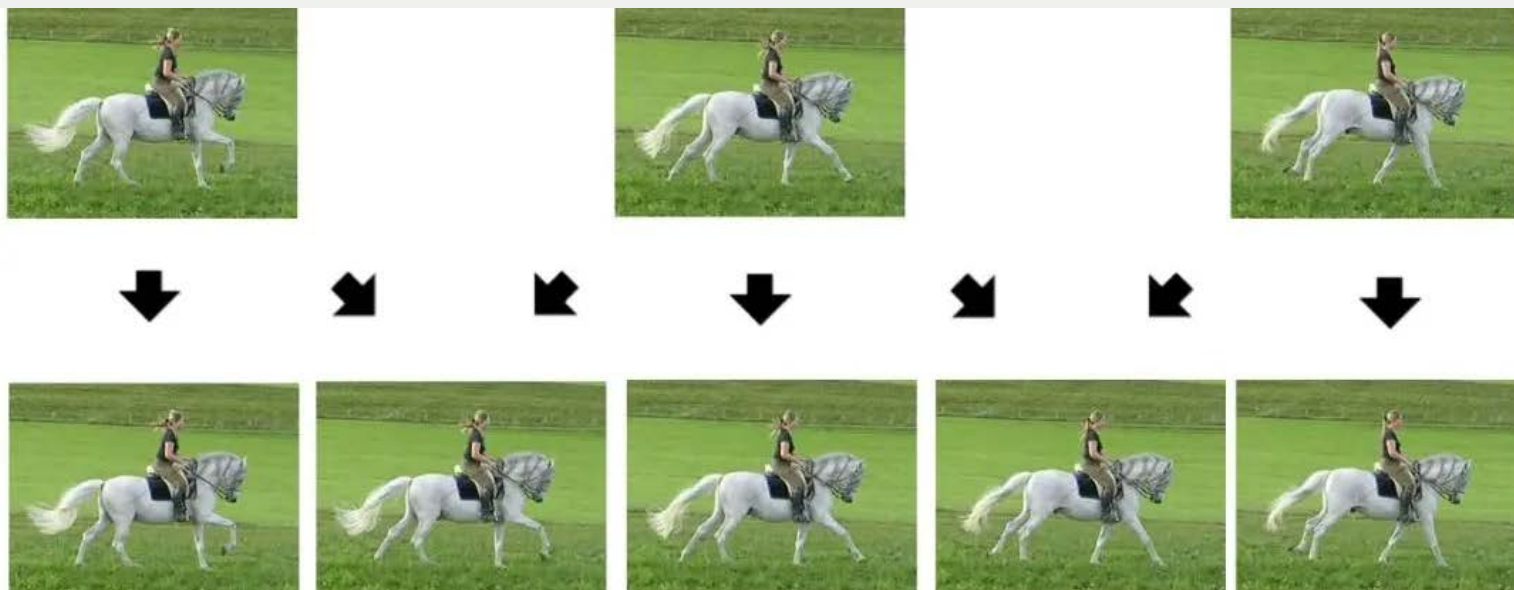
- ◆ 由于视频帧的DCT变换基于 $8*8$ 分块，所以计算和编码前后两帧差异时也需要基于块
 - ◆ 然而由于物体运动，前后两帧的物体所在分块可能不一样，也就是虽然内容相同，但是位置不同



运动估计-时间冗余

如何去除时间冗余?

- ◆ 因此需要捕捉视频里面对象的运动趋势
 - ◆ 从前一帧里寻找匹配块，记录下当前帧和匹配块之间的位移和差异，即为运动估计，可在前一帧图像的基础上进行运动补偿得到当前帧图像



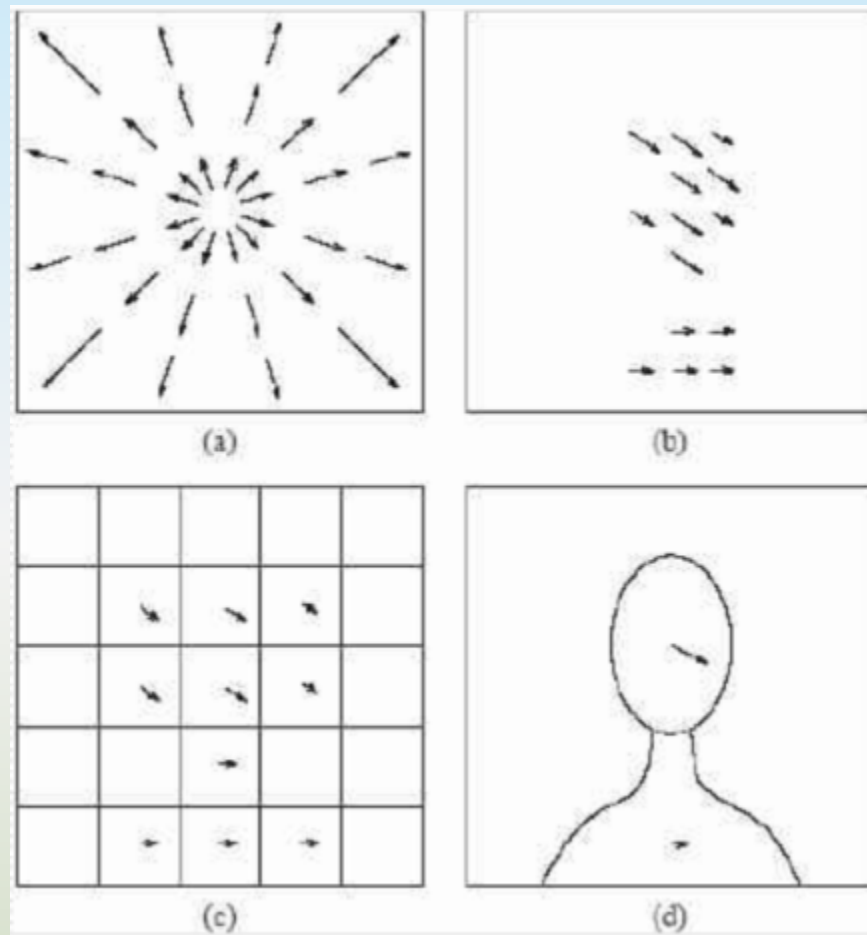
视频多媒体基础

- 一、视频压缩原理
- 二、运动估计
- 三、视频压缩标准H.264
- 四、视频压缩标准H.265

运动估计

◆ 方法分类

- ◆ (a) 全局运动估计
- ◆ (b) 基于像素点的运动估计
- ◆ (c) 基于块的运动估计
- ◆ (d) 基于区域的运动估计

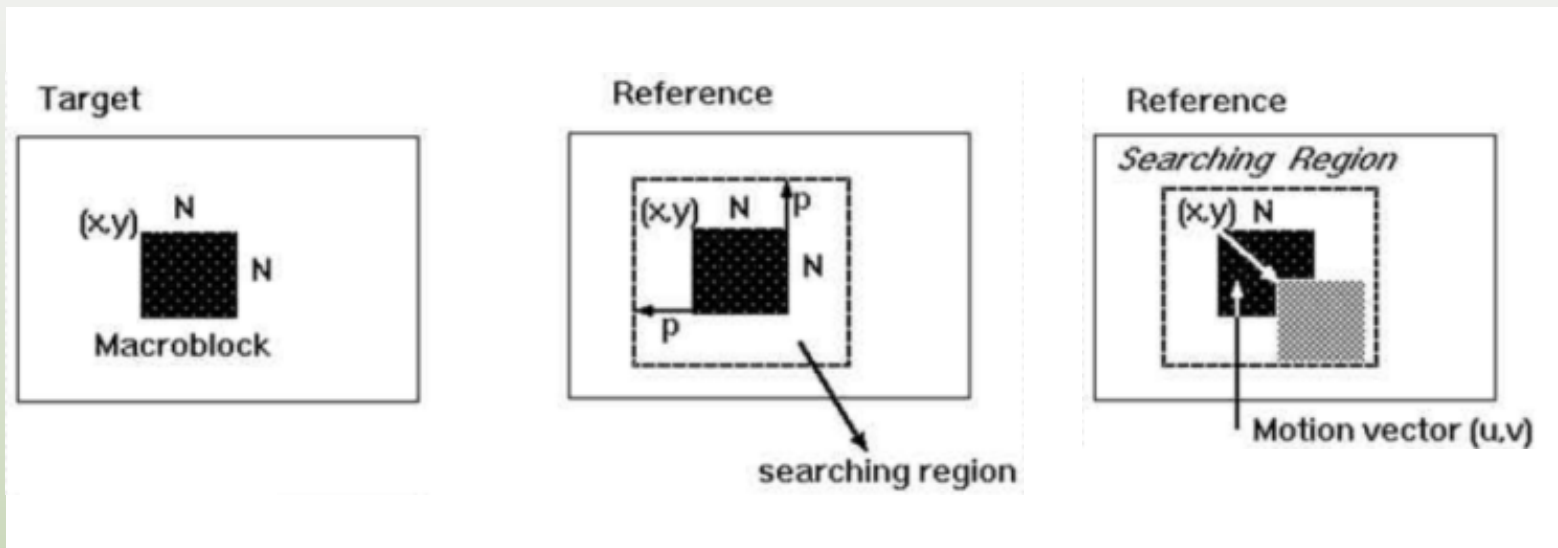


■ 基于块的运动估计因算法简单、便于硬件实现得到了广泛应用

基于块的运动估计

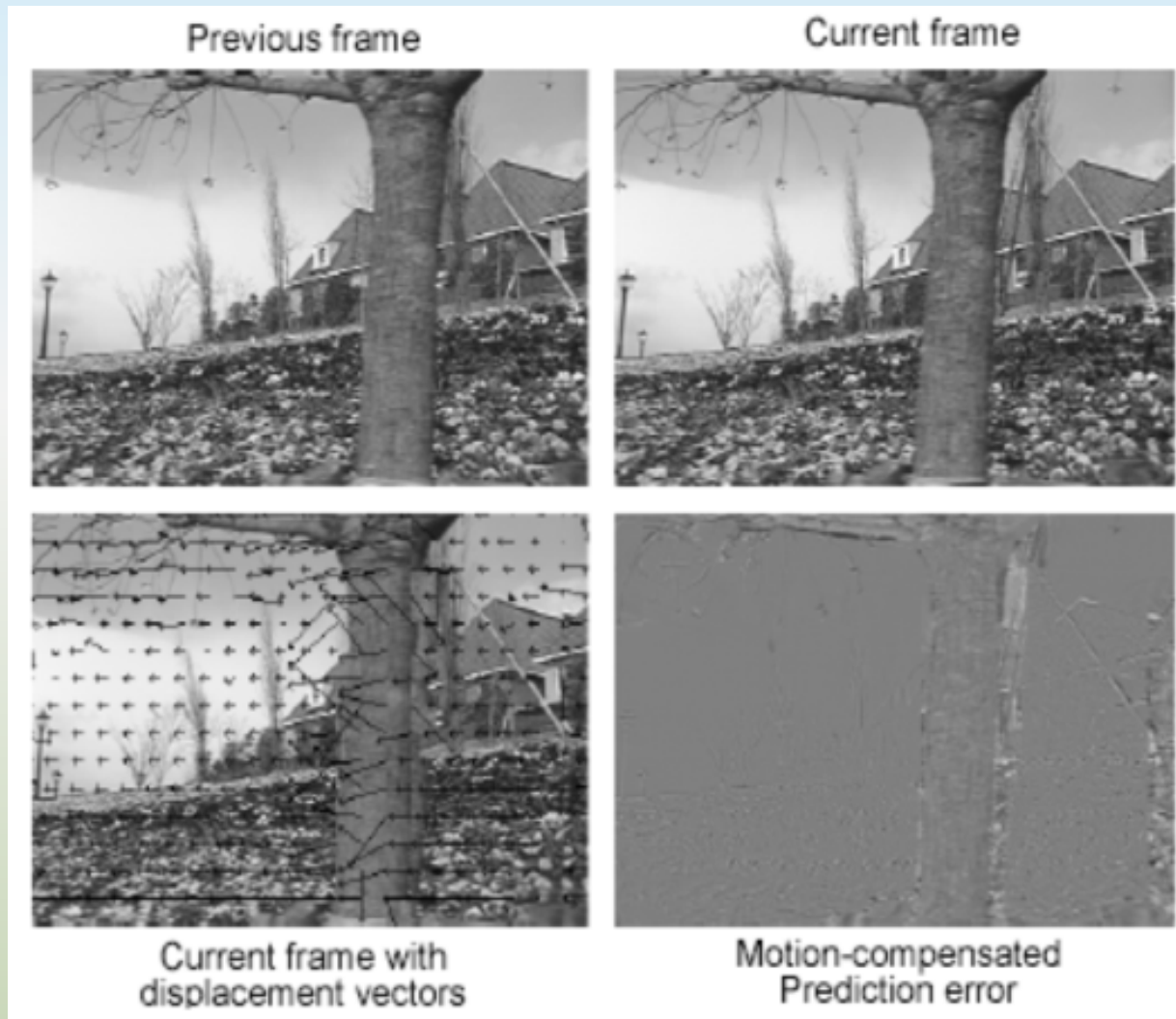
◆ 基本思想

- ◆ 将图像序列的每一帧分成许多互不重叠的宏块，并认为宏块内所有像素的位移量都相同，然后根据匹配准则对每个宏块，在给定搜索范围内找出与当前宏块最相似的宏块，即匹配块，当前块与匹配块之间的相对位移即为运动矢量。视频压缩时，只需保存初始块和运动矢量就可以完全恢复出当前块



基于块的运动估计

◆ 示例



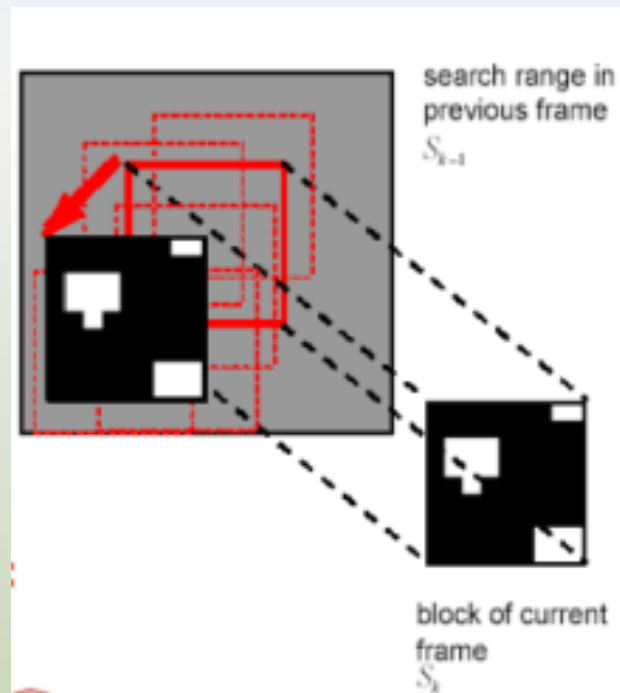
基于块的运动估计

◆ 算法流程

- ◆ 将图像划分为多个子块
- ◆ 为每个子块确定搜索范围
- ◆ 在搜索范围内，为每个子块搜索最佳的匹配块
- ◆ 计算当前块与匹配块之间的相对位移

◆ 注意事项

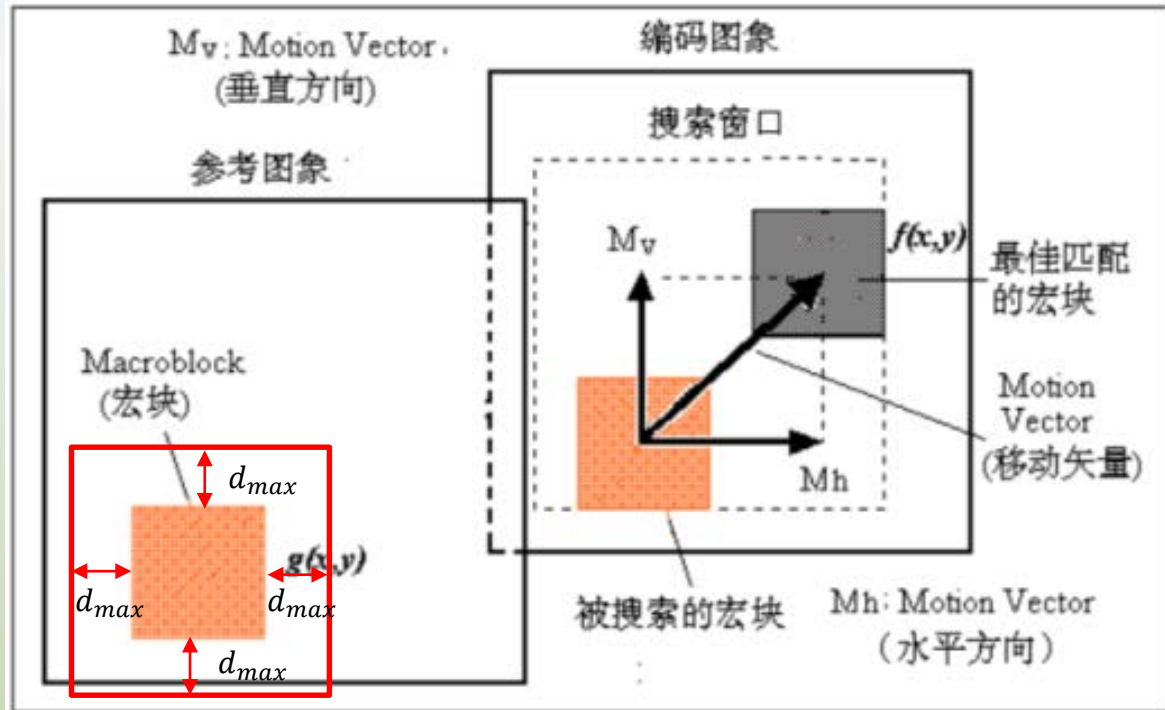
- ◆ 设定适当的块匹配搜索范围
- ◆ 选择合理的块匹配判决准则
- ◆ 减少搜索匹配块的计算次数



基于块的运动估计

◇ 搜索范围

- ◇ 块匹配法把图像分成许多 ($M \times N$ 个像素) 的宏块, 并且假定同一宏块内像素具有相同的位移
- ◇ 图像块在水平和垂直方向的最大位移均为 d_{max} , 则搜索区为: $(M+2d_{max}) \times (N+2d_{max})$



块匹配：判决准则

◆ 常见的块匹配判决准则有三种：

◆ 平均绝对误差(MAD)取最小

$$MAD(d_x, d_y) = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N |S_k(x, y) - S_{k-1}(x + d_x, y + d_y)| \rightarrow \min$$

◆ 均方误差(MSE)取最小

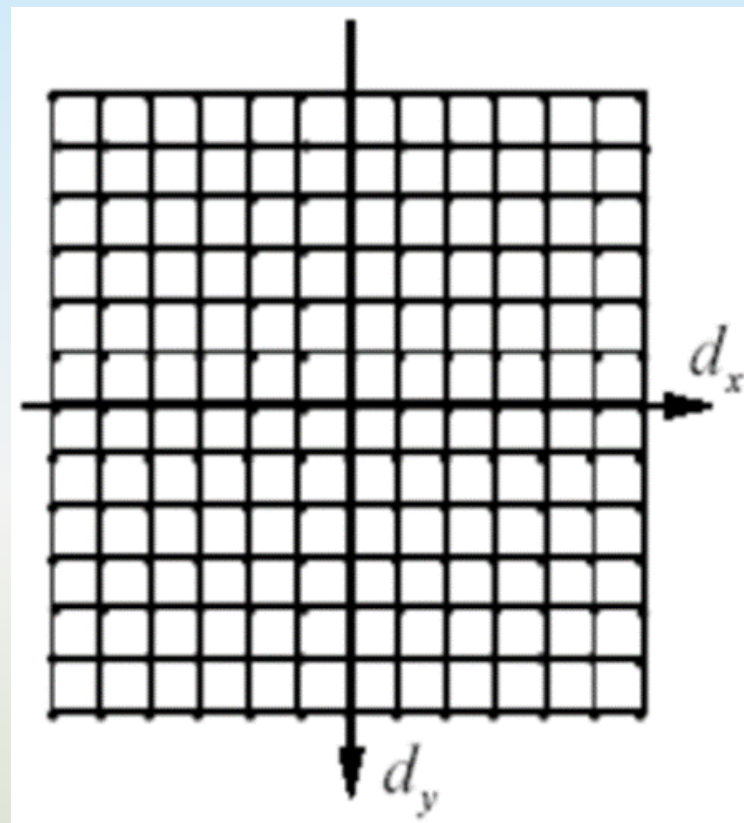
$$MSE = \sum_{x=1}^M \sum_{y=1}^N [S_k(x, y) - S_{k-1}(x + d_x, y + d_y)]^2 \rightarrow \min$$

◆ 归一化互相关函数(NCCF)取最大

$$N_{CCF}(d_x, d_y) = \frac{\sum_{x=1}^M \sum_{y=1}^N S_k(x, y) S_{k-1}(x + d_x, y + d_y)}{\left[\sum_{x=1}^M \sum_{y=1}^N S_k^2(x, y) \right]^{\frac{1}{2}} \left[\sum_{x=1}^M \sum_{y=1}^N S_{k-1}^2(x + d_x, y + d_y) \right]^{\frac{1}{2}}}$$

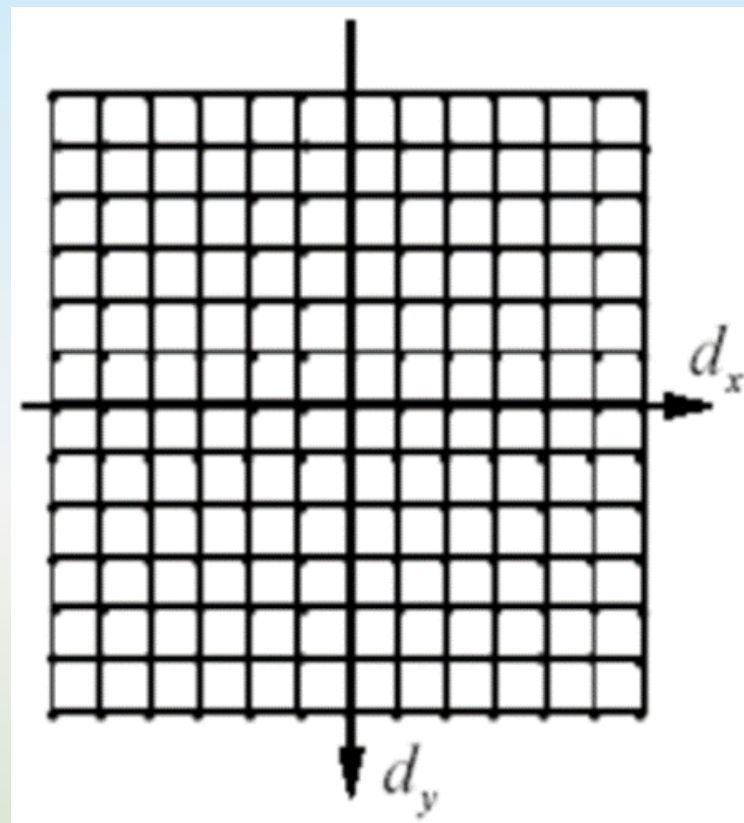
块匹配：全搜索

- ◆ 最细致的搜索方法是在整个搜索区逐点搜索，每搜索一个像素点，计算一次MAD，当MAD达到最小值时，求得最佳匹配像素块
- ◆ 需要计算MAD的次数是 $(M+2d_{max}) \times (N+2d_{max})$ ，而且每个MAD计算的累加次数为 $M \times N$



块匹配：全搜索

- ◆ 当 $M=N=16$, $d_{max}=8$ 时, 需要计算 MAD 的总次数 $=262144$
- ◆ 当图像空间分辨率高 (M 和 N 值更大), 运动速度快时, 需要大范围搜索 (d_{max} 值更大), 这时运算量是相当大的
- ◆ 特点总结:
 - ◆ 精度最高, 高度规则化, 适合并行计算, 但是计算复杂, 难以实时处理



必须研究相应的快速运动估计算法

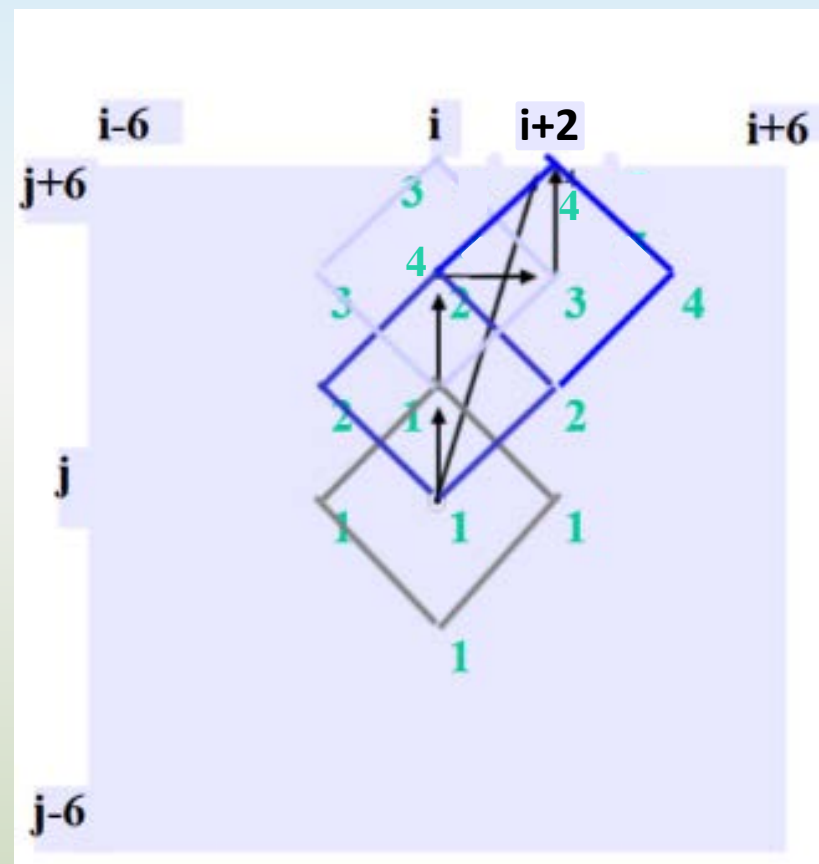
快速搜索算法

- ◆ 包括二维对数法、三步法、分级搜索法、螺旋形搜索法、十字法、菱形法、六边形法等
- ◆ 它们的共同之处在于
 - ◆ 将判决准则MAD趋于最小值的方向视为最小失真方向;
 - ◆ 假定判决准则MAD在靠近最小失真方向时是单调递减的, 即认为它在整个搜索区内是单极点函数, 具有唯一极小值;
 - ◆ 假定快速搜索是从任一猜测点开始, 沿最小失真方向进行的。
- ◆ 这些快速搜索算法在实质上都是统一的梯度搜索法, 不同的是搜索路径和步长有所区别

快速搜索算法

◆ 二维对数法

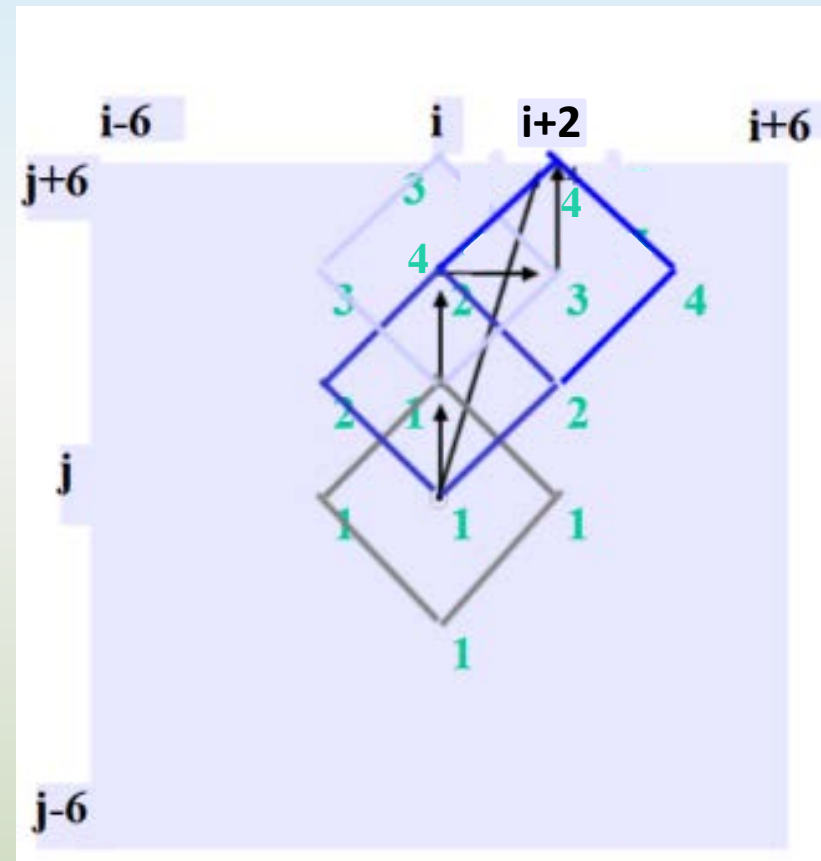
- ◆ 每一步骤中，在中心及其左右上下 5 个相邻位置上计算判决函数值 MAD
- ◆ 取其中最小的 MAD 值作为下一步计算的中心点，重复上面的步骤，直到找到最佳匹配位置，或中心点到达了搜索区域的边界
- ◆ 如右图所示，经过 4 步之后，在 $(i+2, j+6)$ 点得到最小失真方向



快速搜索算法

◆ 二维对数法

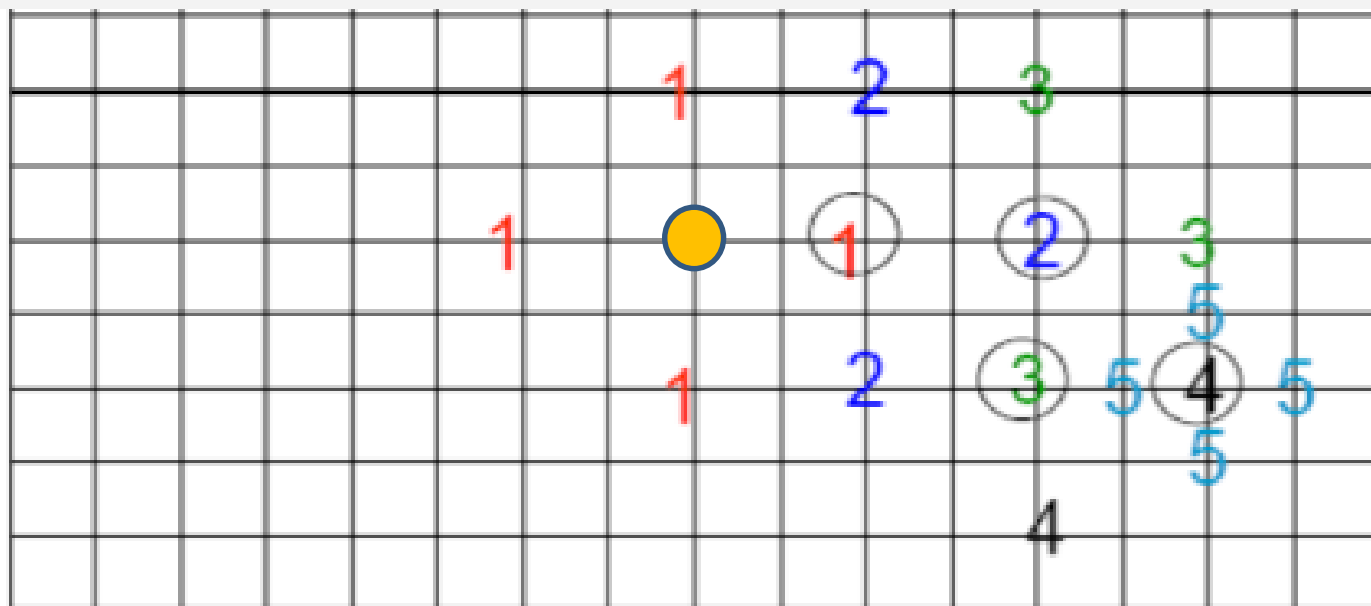
- ◆ 第一步，以起始点 (i, j) 为中心，搜索步长为2个像素，左右上下计算判决函数值 MAD， $(i, j+2)$ 是判决函数值最小的点
- ◆ 第二步，以该点 $(i, j+2)$ 为中心，左右上下计算判决函数值 MAD， $(i, j+4)$ 是判决函数值最小的点
- ◆ 第三步，以该点 $(i, j+4)$ 为中心，左右上下计算判决函数值 MAD， $(i+2, j+4)$ 是判决函数值最小的点
- ◆ 第四步，以该点 $(i+2, j+4)$ 为中心，左右上下计算判决函数值 MAD，已到搜索边界得到最终的运动矢量 $(2, 6)$



快速搜索算法

◆ 十字搜索法

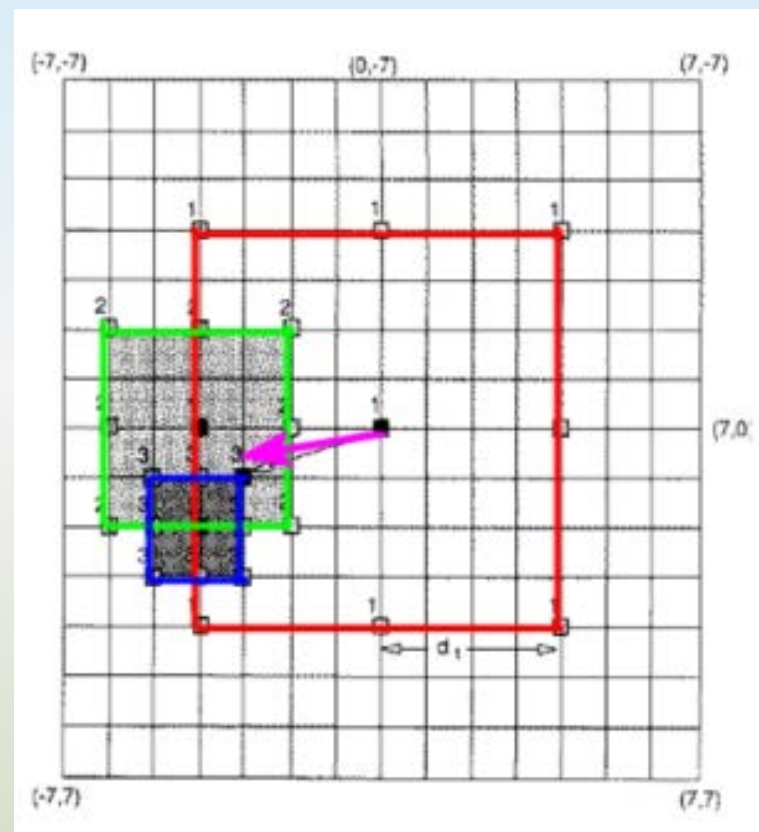
- ◆ 动机: 大多数运动是在水平或竖直方向
- ◆ 初始搜索步长为 $R/2$ ，每步搜索 4 个十字交叉位置
- ◆ 当接近最佳匹配或图像边界时步长减半，也可以在每次循环后步长减半



快速搜索算法

◆ 三步法

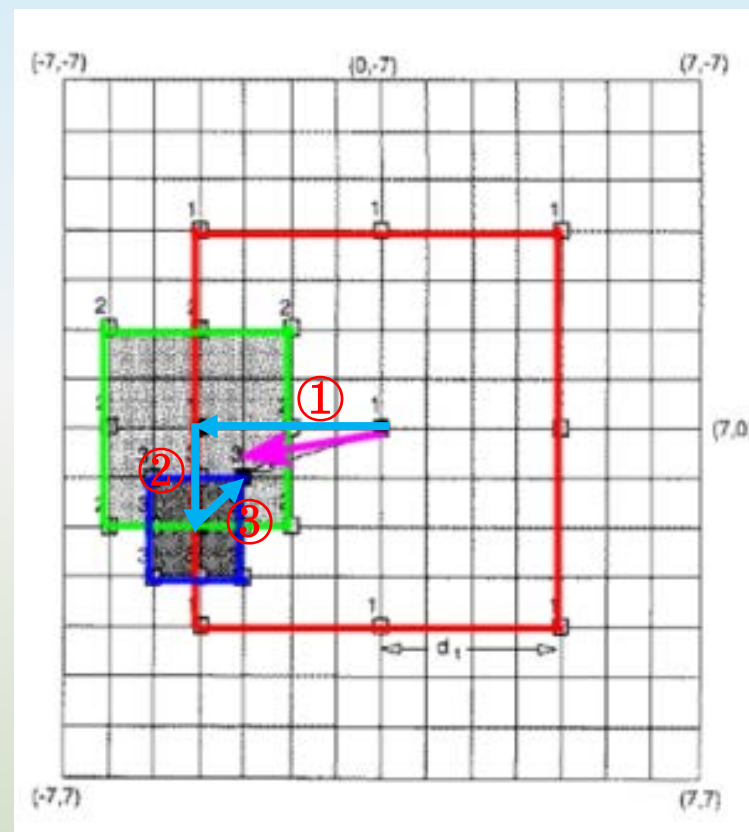
- ◆ 假设匹配误差随着和全局误差最小点的距离增加而单调增加
- ◆ 与二维对数法十分相似，不同的是对8个邻近点进行计算，而且搜索步长随着搜索步骤的增加而减少
- ◆ 优点是运算量显著减少，特别是当搜索范围增大时
- ◆ 缺点是在运动估值的质量方面，和全搜索仍有一定距离，而且对小的运动检测效果不好



快速搜索算法

◆ 三步法

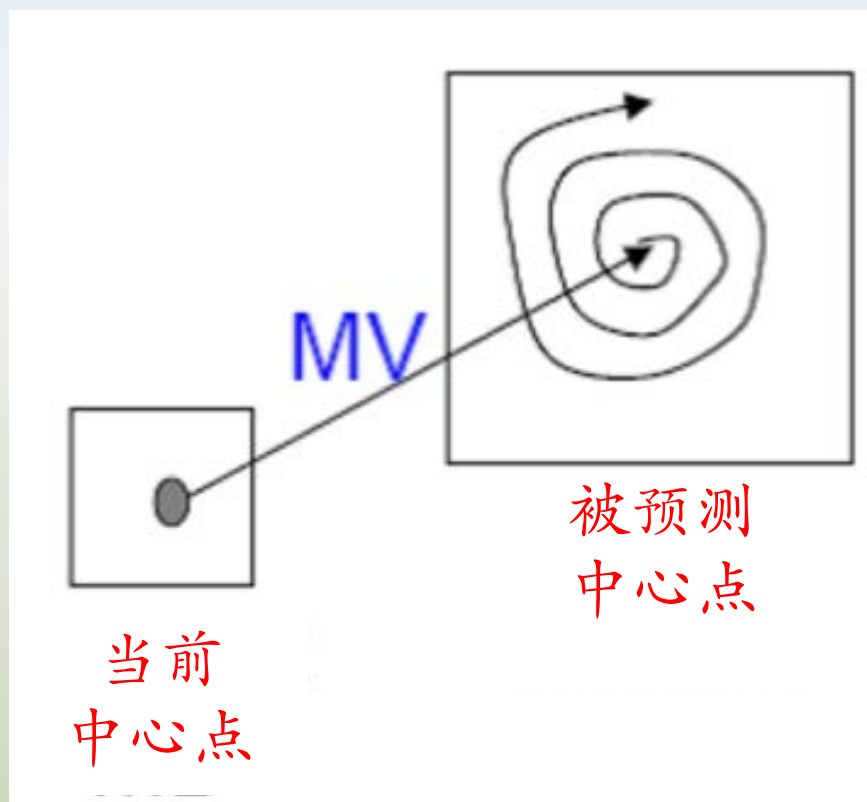
- ◆ 第一步，以起始点 $(0, 0)$ 为中心，搜索步长为4个像素， $(-4, 0)$ 是判决函数值最小的点
- ◆ 第二步，以该点 $(-4, 0)$ 为中心，搜索步长为2个像素，找到第二级近似点 $(-4, -2)$
- ◆ 第三步，对距离中心为1个像素点的邻近点进行搜索，得到最终的运动矢量 $(-3, -1)$



快速搜索算法

◆ 螺旋形搜索法

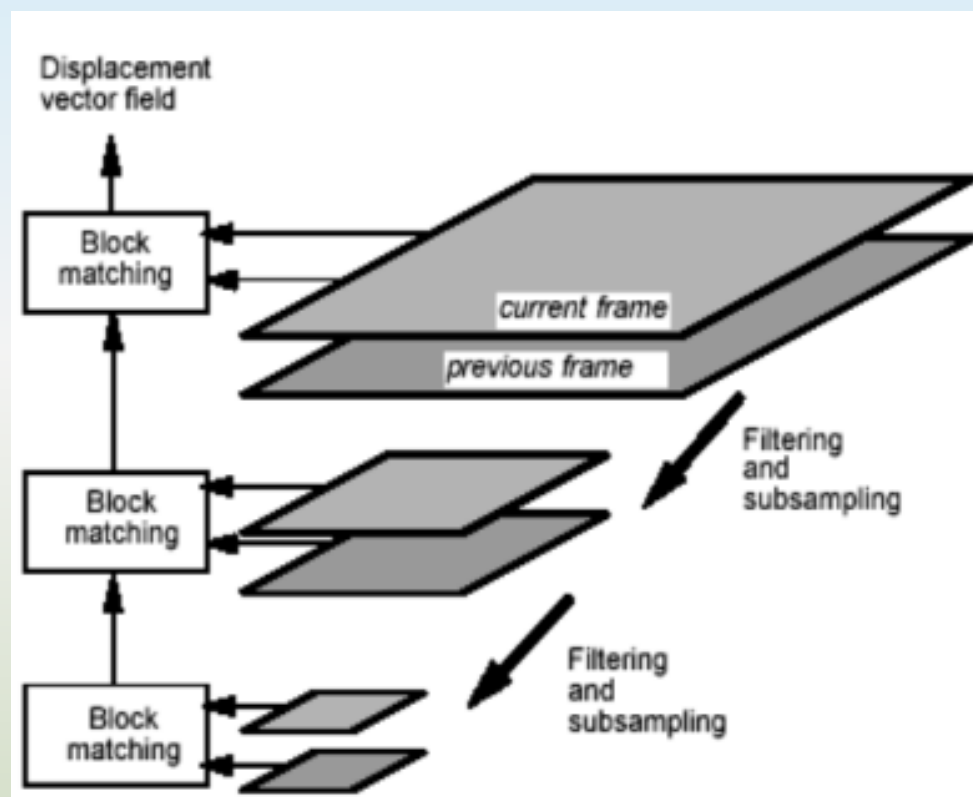
- ◆ 首先需要预测移动的中心点，然后在以被预测移动中心点为中心的区域内，以螺旋形顺序测试候选的移动中心点



快速搜索算法

◆ 分级搜索法

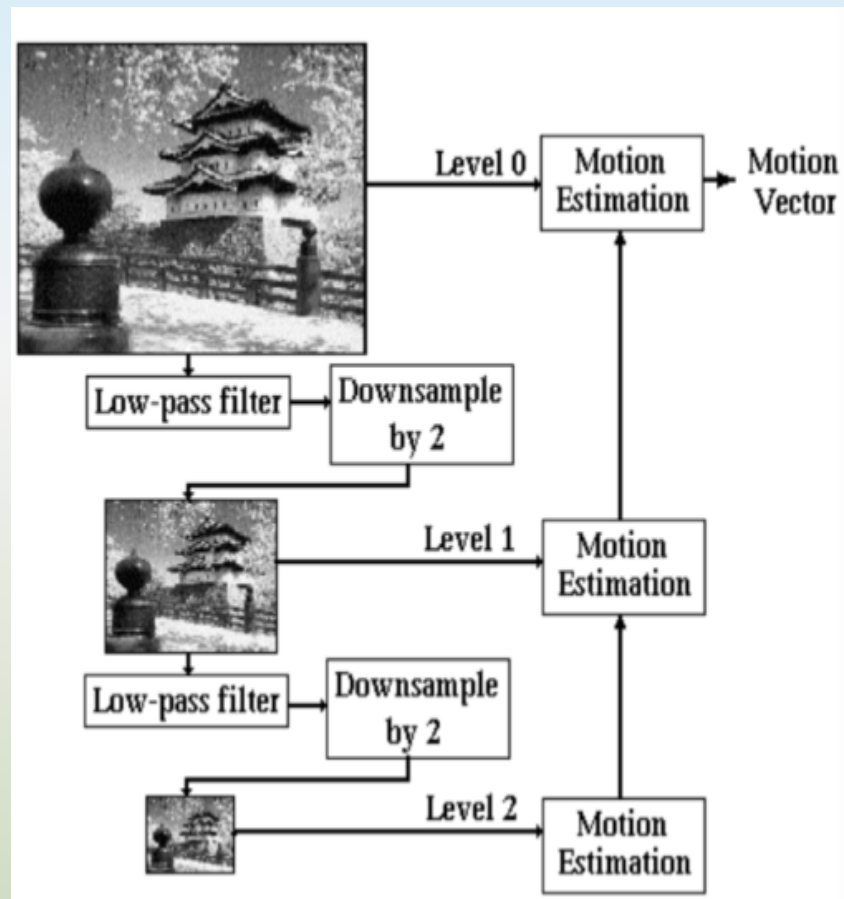
- ◆ 在减少整体运算量的同时，力求接近全搜索的效果
- ◆ 其实质是通过低通滤波和甩点，将原始图像分解成不同分辨率等级的图像序列，然后采用粗、中、细等不同等级的运动搜索



快速搜索算法

◆ 分级搜索法

- ◆ 首先对低分辨率图像进行全搜索。由于分辨率降低，搜索次数成倍减少，这是一种粗搜索
- ◆ 然后以低分辨率图像搜索得到的结果，作为下一步细搜索的起始点
- ◆ 在细搜索时，搜索范围缩小很多，因此，搜索次数也相应减少。经过粗细不同的搜索，得到最终的运动矢量



视频多媒体基础

- 一、视频压缩原理
- 二、运动估计
- 三、视频压缩标准H.264
- 四、视频压缩标准H.265

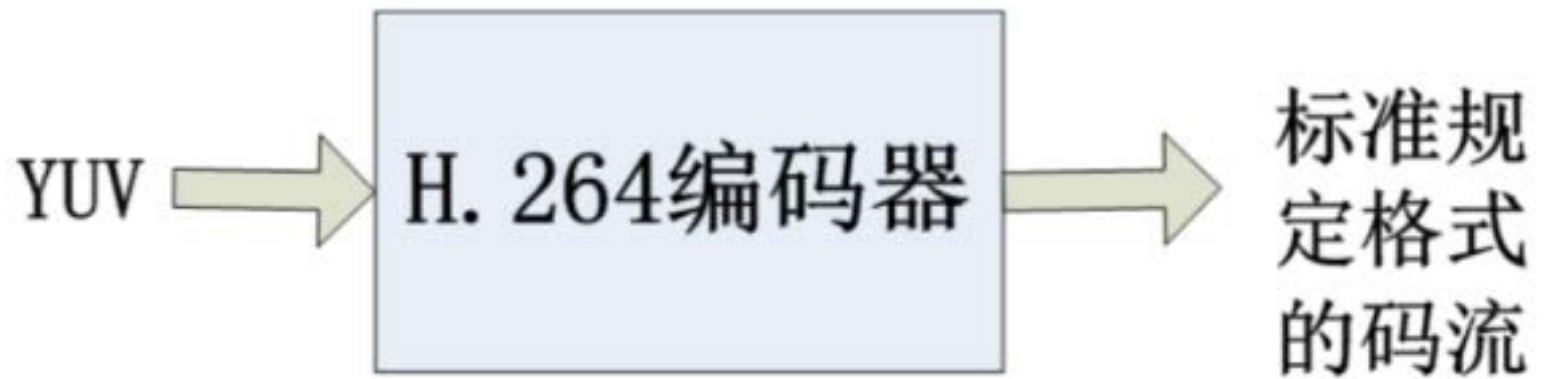
H.264的出现

- ◆ 两大视频编码组织
 - ◆ ISO/IEC下的动态图像专家组MPEG (Moving Picture Experts Group)
 - ◆ ITU-T下的视频编码专家组VCEG (Video Coding Experts Group)
- ◆ 2001年，VCEG 与MPEG联合组成了联合视频组 (JVT, Joint Video Team) ，继而推出了新数字视频编码标准H.264

H.264的特点

- ◆ 具有良好的网络亲和性，适用于各种网络传输
- ◆ 在混合编码器的基本框架下，对其主要关键部件都做了重大改进，如多模式运动估计、帧内预测、多帧预测、基于内容的变长编码等
- ◆ 在关注H.264的优越性能的同时必然需要衡量其实现难度。从总体上说，H.264性能的改进是以增加复杂性为代价而获得的
 - ◆ 据估计，H.264编码的计算复杂度大约相当于上一代数字视频编码标准H.263的3倍，解码复杂度大约相当于H.263的2倍

编码器的输入和输出



编码器的输入和输出

什么是YUV颜色模型

- 人眼视觉系统对亮度的感知比对色度的感知更加敏感，因此将彩色空间分解为亮度(Y)、色度(U)和浓度(V)三个基本部分，色度和浓度通道(UV)的采样率可以低于亮度通道(Y)，而不会显著降低感知质量
- 根据对Y、U、V三个分量的采样比率不同，数字视频图像可分为4:1:1，4:2:2，4:4:4等格式



4:1:1

4:2:2

4:4:4

视频的冗余

◆ 空间冗余

- ◆ 每帧图像上，每个像素点与它相邻的像素点具有相关性

◆ 时间冗余

- ◆ 序列图像中，相邻帧之间具有很强的相关性

◆ 信息熵冗余

- ◆ 也称编码冗余，每个像素点只要按照信息熵大小分配相应比特数即可，如果对每个像素值采用相同比特数则存在信息熵冗余

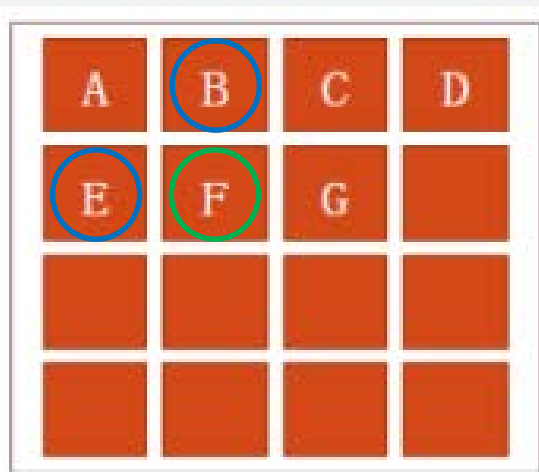
视频的冗余

- ◆ 针对不同的冗余，H.264分别提出消除各种冗余的编码方法
 - ◆ 针对空间冗余：帧内(intra)预测编码，变换编码
 - ◆ 针对时间冗余：帧间(inter)预测编码
 - ◆ 针对信息熵冗余：使用上下文自适应的变长编码（CAVLC），将各种编码方法有效组合在一起
- ◆ 这些技术在编码器中的次序关系如下图：



预测技术

- ◆ H.264处理图像时，不以像素为单位，而以宏块(MB)为最小编码单位，宏块为 16×16 像素的小块
- ◆ 视频压缩处理时，每一帧图像被划分成许多这样的宏块，从左到右，从上到下依次处理



图像被划分成宏块

H.264的预测参考左方或者上方已编码的宏块，例如F块的编码，会参考E块或者B块的编码

帧内预测技术

4×4的亮度块

模式

说明

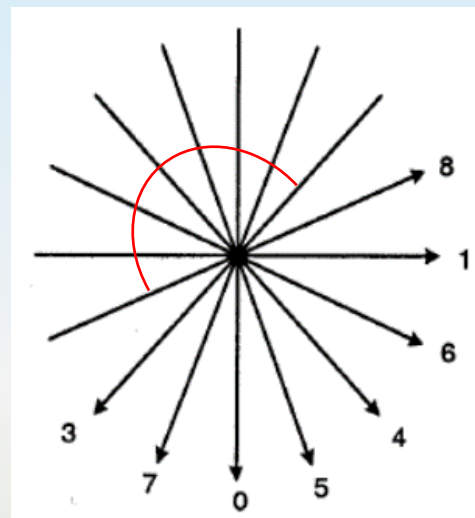
Mode 0 由上方的A、B、C、D进行垂直推算

Mode 1 由左侧的I、J、K、L进行水平推算

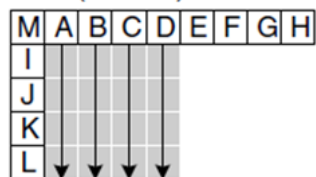
Mode 2 P中的所有样本预测值都等于A~D及I~L的平均值

Mode 3 由45度角往左下方向进行推算

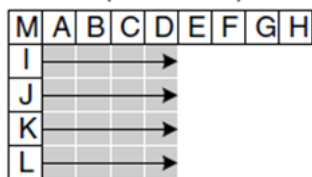
Mode 4 由45度角往右下方向进行推算



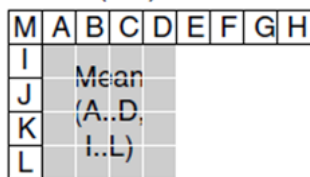
0 (vertical)



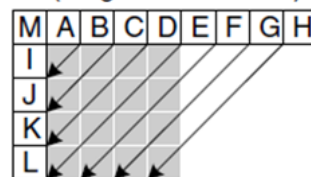
1 (horizontal)



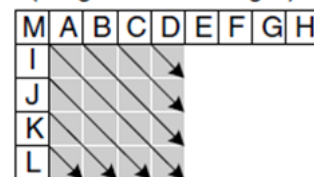
2 (DC)



3 (diagonal down-left)



4 (diagonal down-right)



帧内预测技术

模式

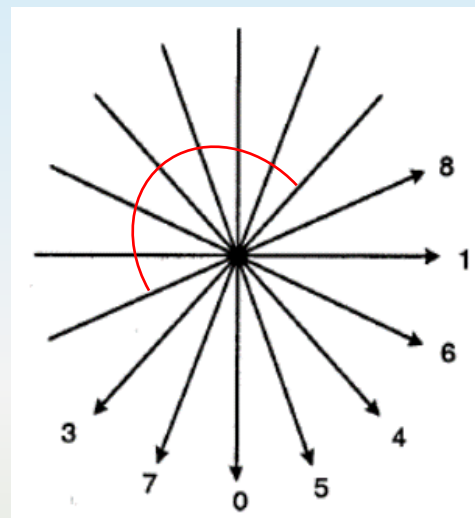
说明

Mode 5 以垂直向下方向右偏大约26.6度角(即宽: 高= 1: 2)的方向推行推算

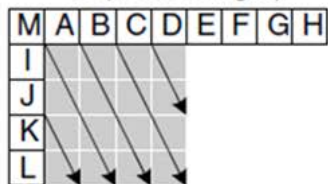
Mode 6 以水平向右方向下偏大约26.6度角(即宽: 高= 2: 1)的方向进行推算

Mode 7 以垂直向下方向左偏大约26.6度角(即宽: 高= 1: 2)的方向推行推算

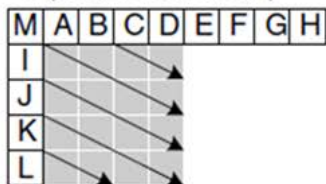
Mode 8 以水平向右方向上偏大约26.6度角(即宽: 高= 2: 1)的方向进行推算



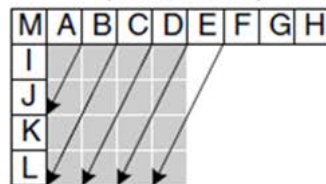
5 (vertical-right)



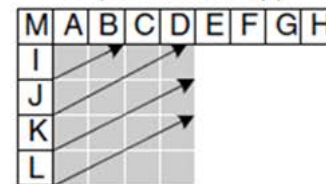
6 (horizontal-down)



7 (vertical-left)



8 (horizontal-up)



帧内预测技术

16×16的亮度块

模式

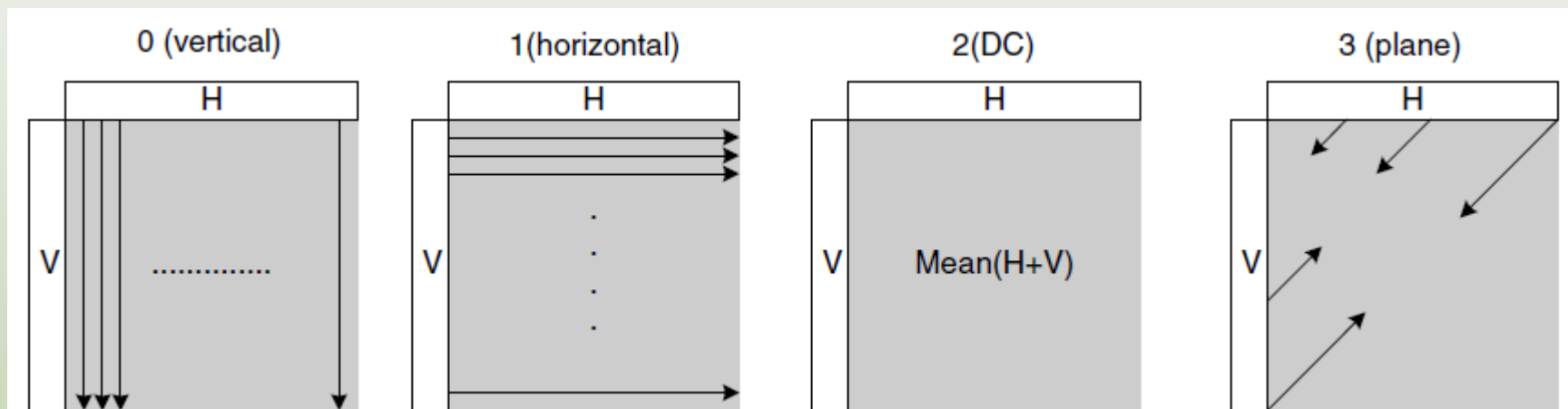
说明

Mode 0 由上方的样本(H)垂直推算

Mode 1 由左侧的样本(V)水平推算

Mode 2 上方样本(H)和左侧样本(V)的平均值

Mode 3 根据上方的样本(H)和左边的样本(V)同时推算得出，在亮度平滑变化的区域效果很好



帧间预测技术

- ◆ 帧间预测就是时域预测，旨在消除时域冗余信息，简单地说就是利用之前编码过的图像来预测之后要编码的图像。其中涉及到两个重要的概念：运动估计和运动补偿
 - ◆ 运动估计即寻找当前编码的宏块在参考帧中的最佳对应块，并且计算出对应块的偏移（运动矢量）
 - ◆ 运动补偿是根据运动矢量和帧间预测方法，求得当前帧的估计值，其实就是将运动矢量参数贴到参考帧上获取当前帧

H.264的应用

- ◆ H.264 支持三个不同档次的应用
 - ◆ 基本档次：主要用于“视频会话”，如会议电视，可视电话，远程医疗、远程教学等
 - ◆ 主要档次：主要用于消费电子应用，如数字电视广播，数字视频存储等
 - ◆ 扩展档次：主要用于网络的视频流，如视频点播



H.264的应用

在视频会议中的应用

- ◆ 视频会议系统对编解码的要求包括：
 - ◆ a. 编解码高效，压缩率高
 - ◆ b. 网络适应性好
 - ◆ c. 抗丢包性能和抗误码性能好



H.264的应用

在视频会议中的应用

◆ H.264中采用的措施

- ◆ 通过对传统帧内预测、帧间预测、变换编码和熵编码等算法的改进，进一步提高编码效率和图像质量
- ◆ 编解码系统首次定义了视频编解码层（VCL）和网络提取层（NAL）
- ◆ 利用参数集、片、FMO、冗余片等关键技术大大提高了系统抗丢包和抗误码的能力



H.264的应用

在视频监控中的应用

- ◆ 与1999年投入使用的国际标准MPEG-4相比，H.264具有以下优势：
 - ◆ 压缩比提高了近30%，大大提高了存储和网络传输带宽
 - ◆ 采用新的算法极大抑制了由于摄像机噪声导致的图像失真和背景流动现象，使得图像质量更加清晰，无马赛克出现
 - ◆ 具备本地实时监视、音视频同步压缩存储、组合报警、有线/无线网络传输、管理权限设置等多种功能



H.264的应用

在数字电视中的应用

- ◆ 数字电视节目一般通过IP包的形式进行接收，在节目播放时，CPU进行的操作除了解压图像数据之外，还需要进行图像显示、音频数据解压、IP数据接收等操作，会造成帧率过慢，图像质量差的问题，达不到实时传输的效果
- ◆ H.264在同等图像质量下，比MPEG-4配置节省一半码率，同时在低码率下方块效应很小，此外具有良好的网络接口特性



视频多媒体基础

- 一、视频压缩原理
- 二、运动估计
- 三、视频压缩标准H.264
- 四、视频压缩标准H.265

H.265的出现

- ◆ H.265是视频编码专家组ITU-VCEG继H.264之后制定的新一代视频编码标准
 - ◆ 2012年8月，爱立信公司推出了首款H.265编解码器
 - ◆ 2013年4月，国际电联（ITU）正式批准通过了H.265标准，标准全称为高效视频编码（HEVC）
- ◆ H.265在有限带宽下能够传输更高质量的网络视频，仅需原先一半带宽即可播放相同质量的视频



H.265的出现

- ◆ H.265围绕前一代视频编码标准H.264，保留原来的大多数技术，同时对一些相关技术加以改进
- ◆ 新技术对某些方面进行改进优化，如码流、编码质量、延时和算法复杂度等，从而提高压缩效率、增强鲁棒性和错误恢复能力、减少实时的时延、降低复杂度等



H.265的优点

◆ 更高的图像质量

- ◆ 智能手机、平板机等移动设备将能够直接在线播放1080p的全高清视频
- ◆ 同时支持4K (4096×2160) 和8K (8192×4320) 的超高清视频



H.265的优点

◆ 更高的图像帧率

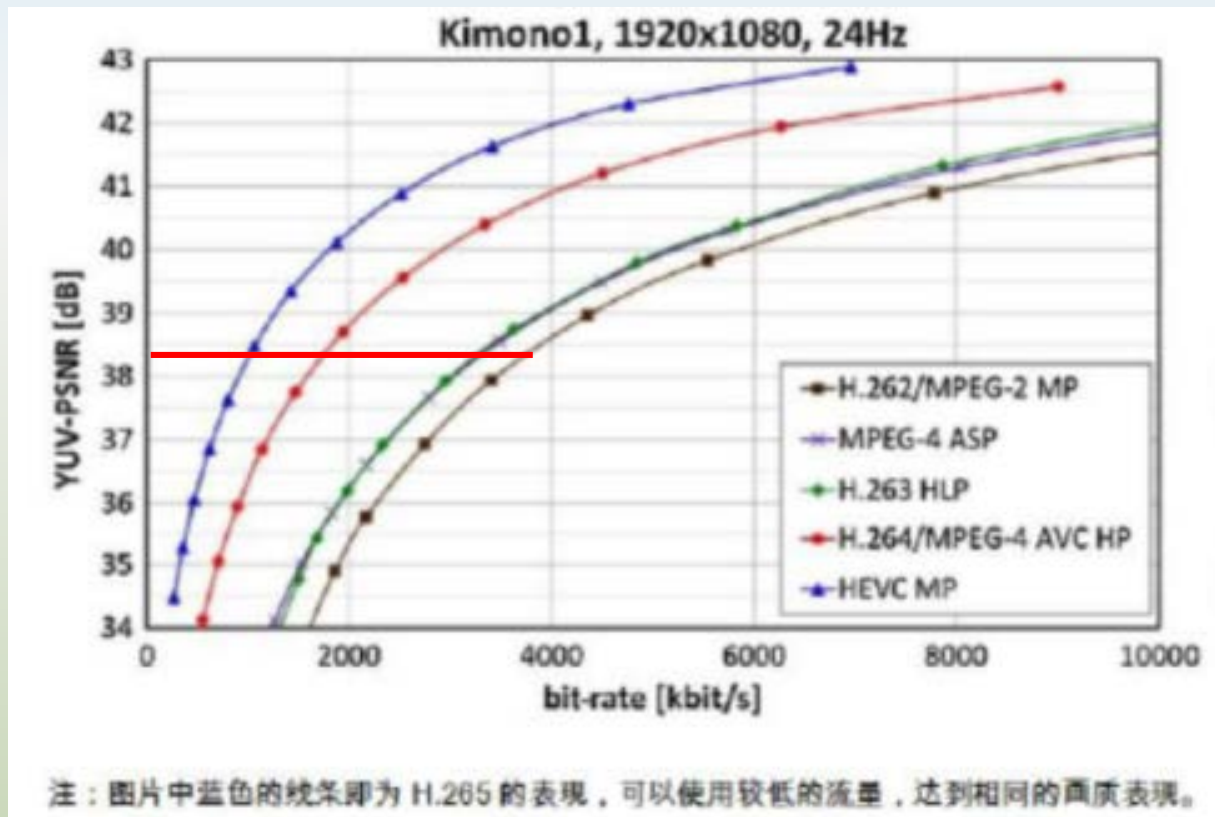
- ◆ 视频帧率从30 帧/秒转向60帧/秒，支持120帧/秒甚至240帧/秒的超高帧率



H.265的优点

◇ 更低的传输码率

- ◇ 在同等图像质量下，数据量大约只有MPEG2的1/16，MPEG4的1/6，H264的1/2



H.265的优点

◇ 更强的容错能力

- ◇ 在高达45%丢包率的不稳定网络环境下仍然能够稳定传输，适用于各种恶劣环境



H.265的优点

- ◆ 更清晰的声音体验
 - ◆ 实现高保真的立体声传输，达到环绕声的效果



H.265的优点

- ◆ 更全的视频接口
 - ◆ 支持HDSDI、DVI、HDMI、CVBS等多种视频接口



H.265的优点

◆ 更好的网络适应性

- ◆ 可以通过IP、无线、卫星等多种网络传输图像，适用各种工作场景

