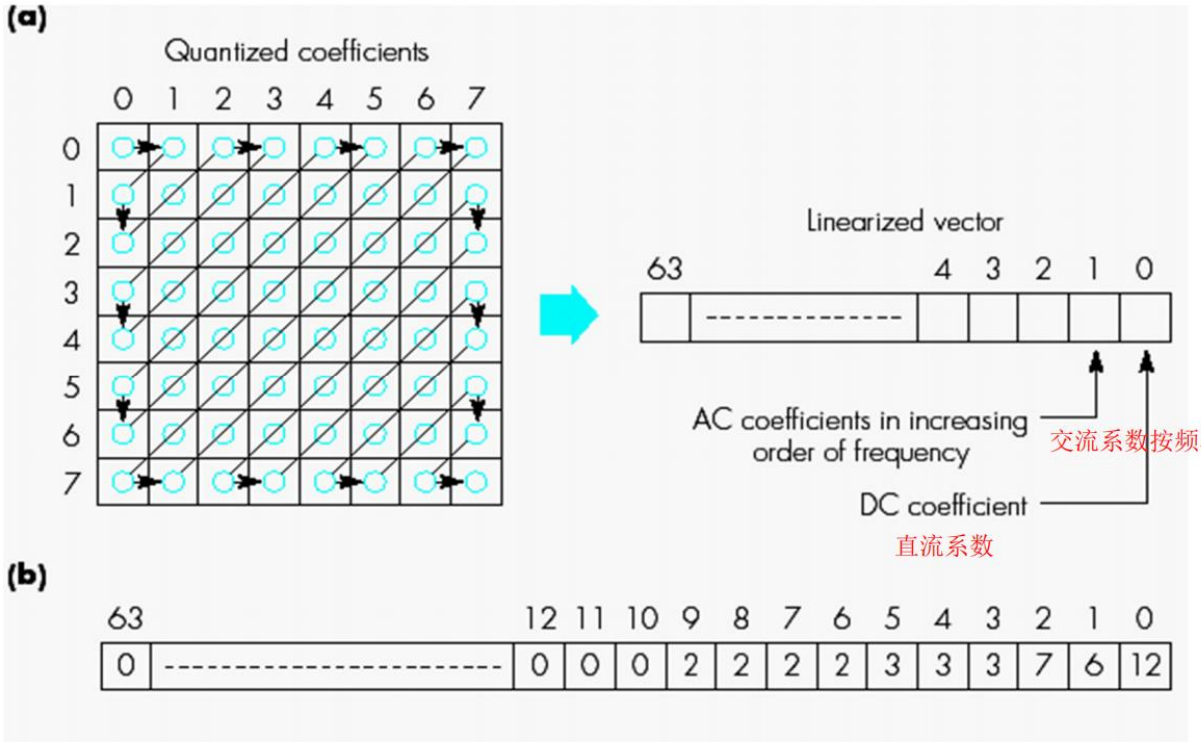
Zig-zag Scanning

锯齿形扫描



2.6.4JPEG编码器

JPEG Encoder

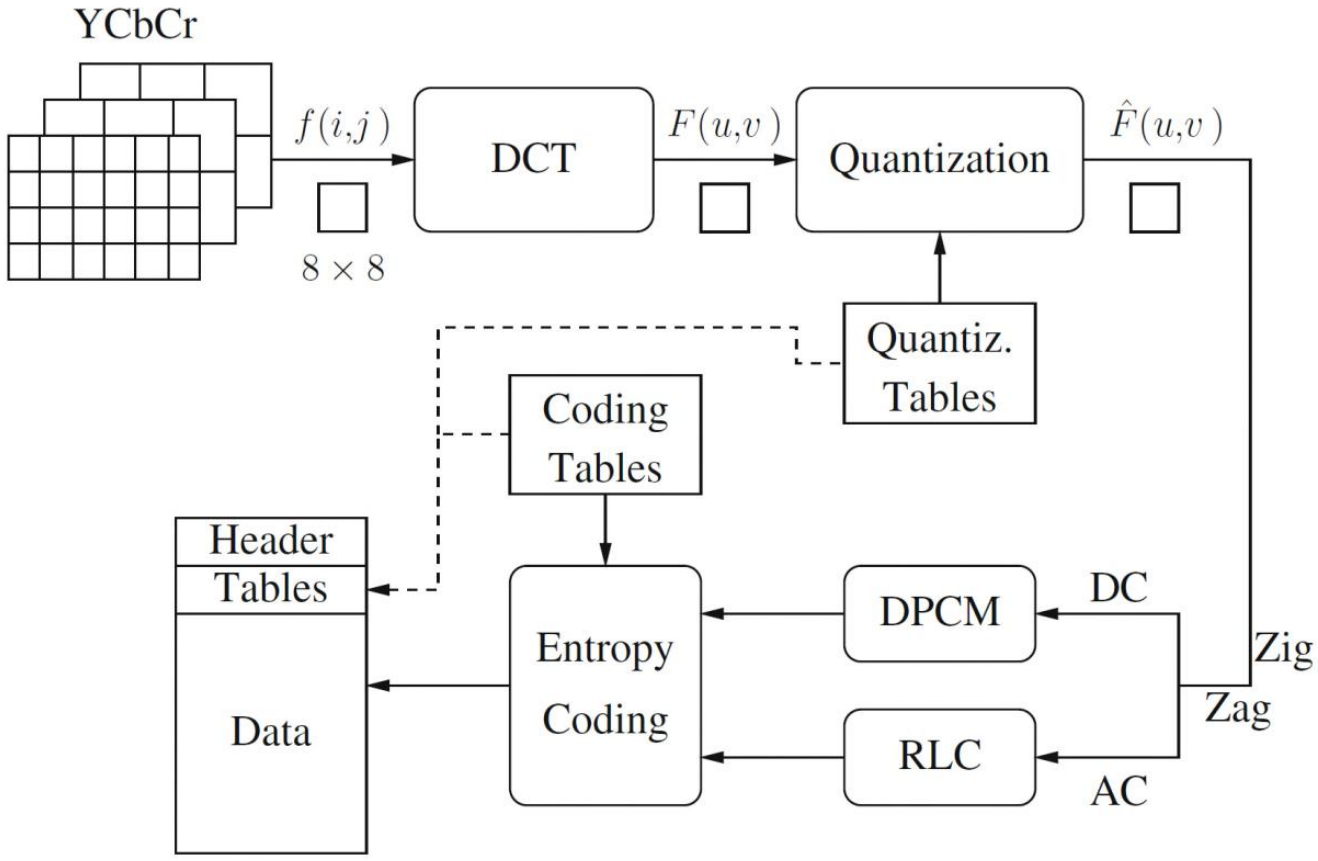
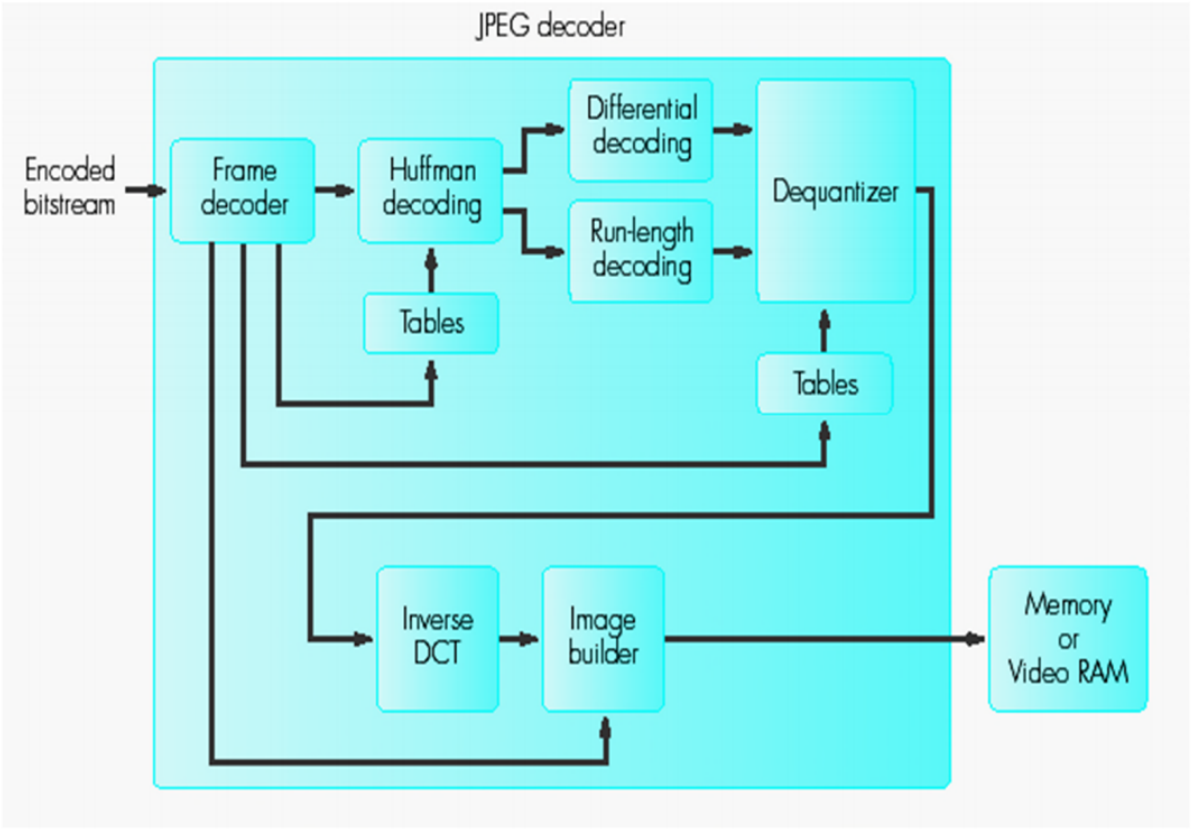


Fig. 9.1: Block diagram for JPEG encoder.

Source: Ze-Nian Li, Mark S. Drew, Jiangchuan Liu, Fundamental of Multimedia, Springer 2021

2.6.12 JPEG解码器

JPEG Decoder



3.2.2.3.4分层搜索

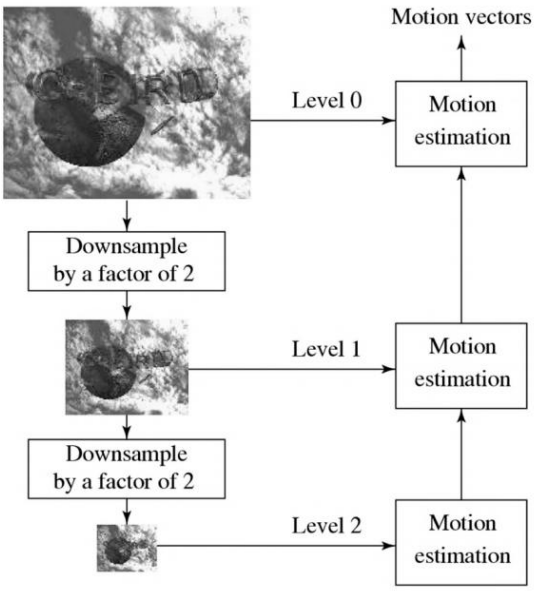
Hierarchical Search

分层搜索；阶层式搜寻

- A three-level hierarchical search, the original image is at Level 0.
- Images at Levels 1 and 2 are obtained by down-sampling from the previous levels by a factor of 2, and the initial search is conducted at Level 2.

三级分层搜索，原始图像处于o级。

1级和2级的图像是通过对前两级进行2倍的降采样获得的，初始搜索在2级进行。

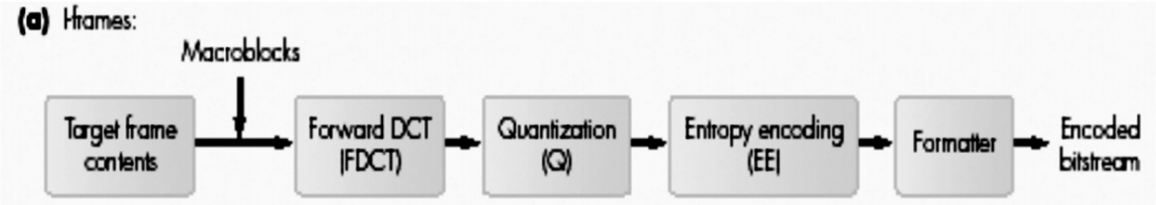


A Three-level Hierarchical Search

3.2.4.2.1 I帧编码流程图

MPEG-1: I-Frame Encoding

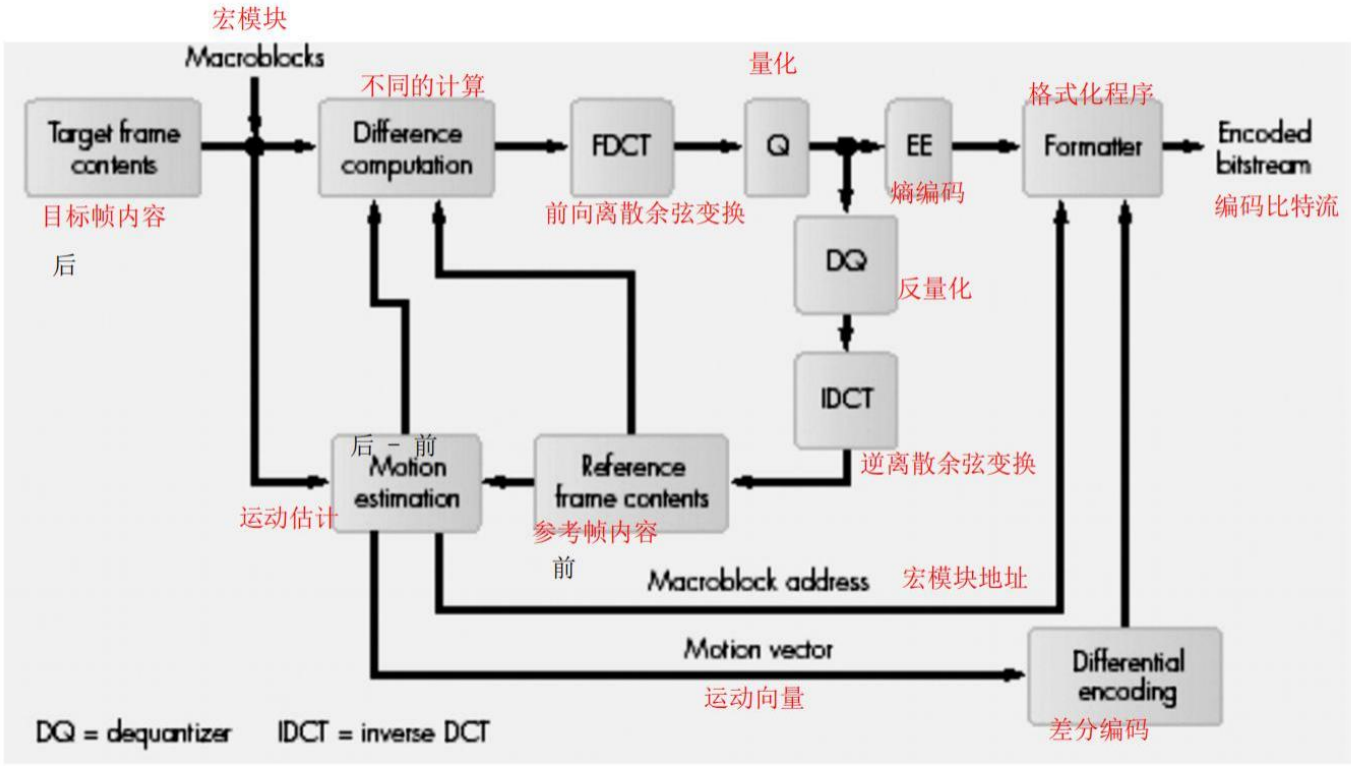
I帧编码



3.2.4.2.2p帧编码流程图

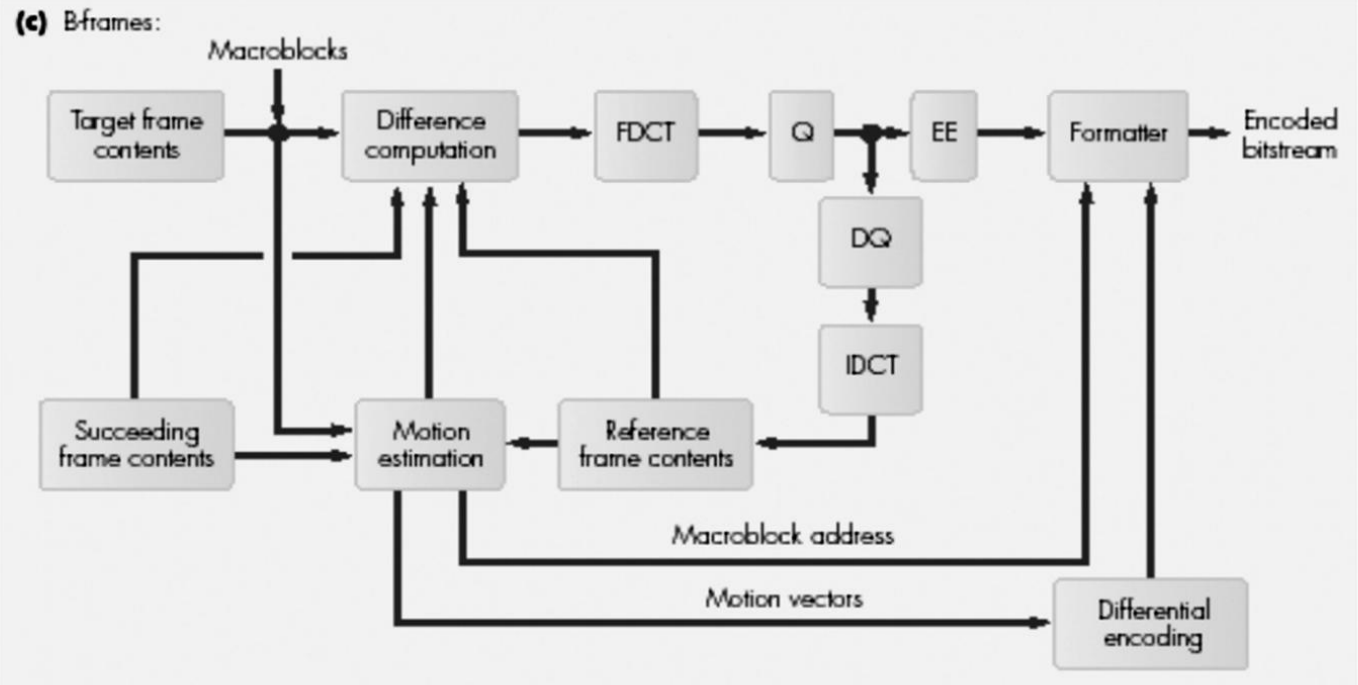
MPEG-1: P-Frame Encoding Flowchart

p帧编码流程图



3.2.4.2.3 B帧编码流程图

MPEG-1: B-Frame Encoding Flowchart

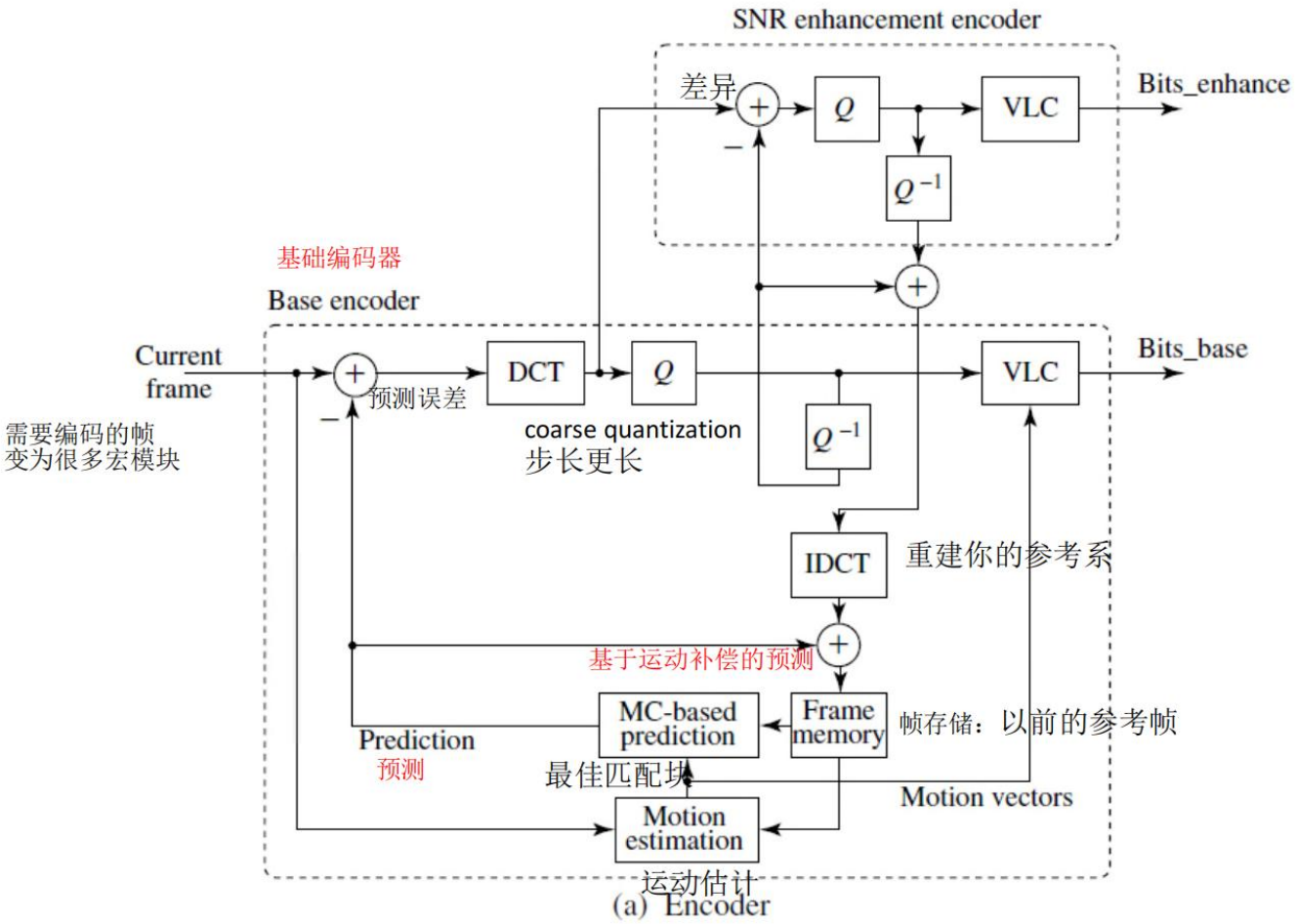


3.2.5.3.1.信噪比可扩展

3.2.5.3.1.信噪比可扩展

MPEG-2: SNR Scalability 2:13:41

步长要小得多，因为你希望它覆盖 更多细节

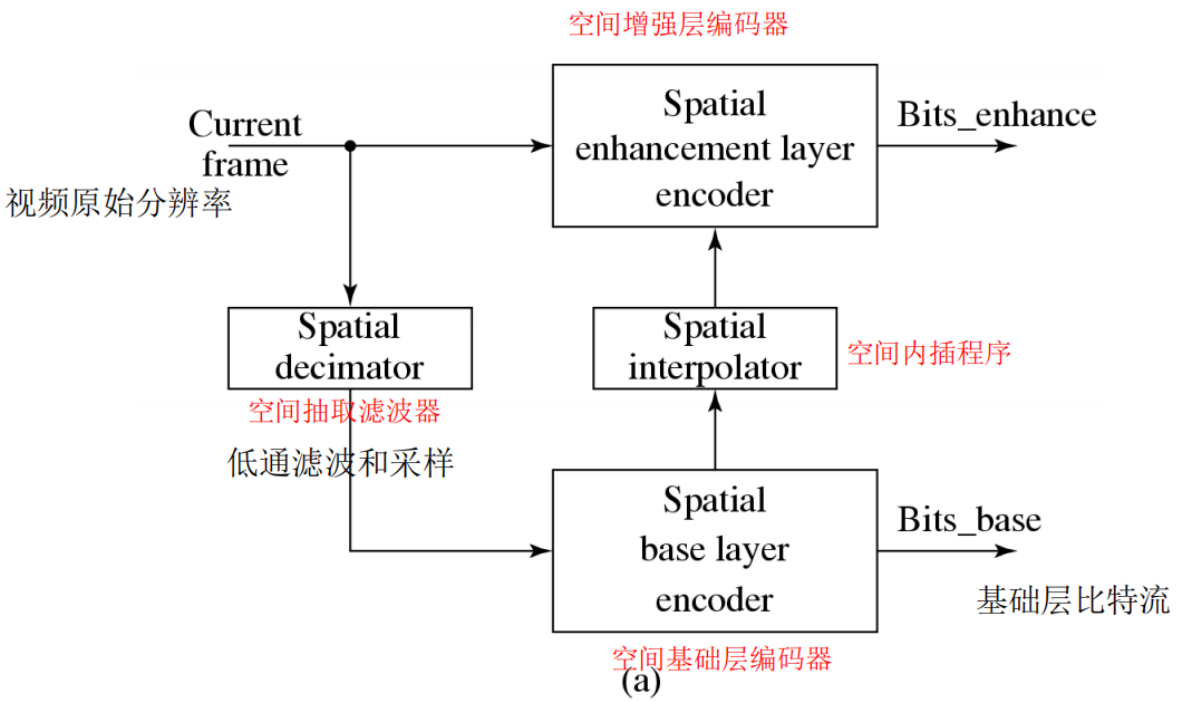


(a) Encoder

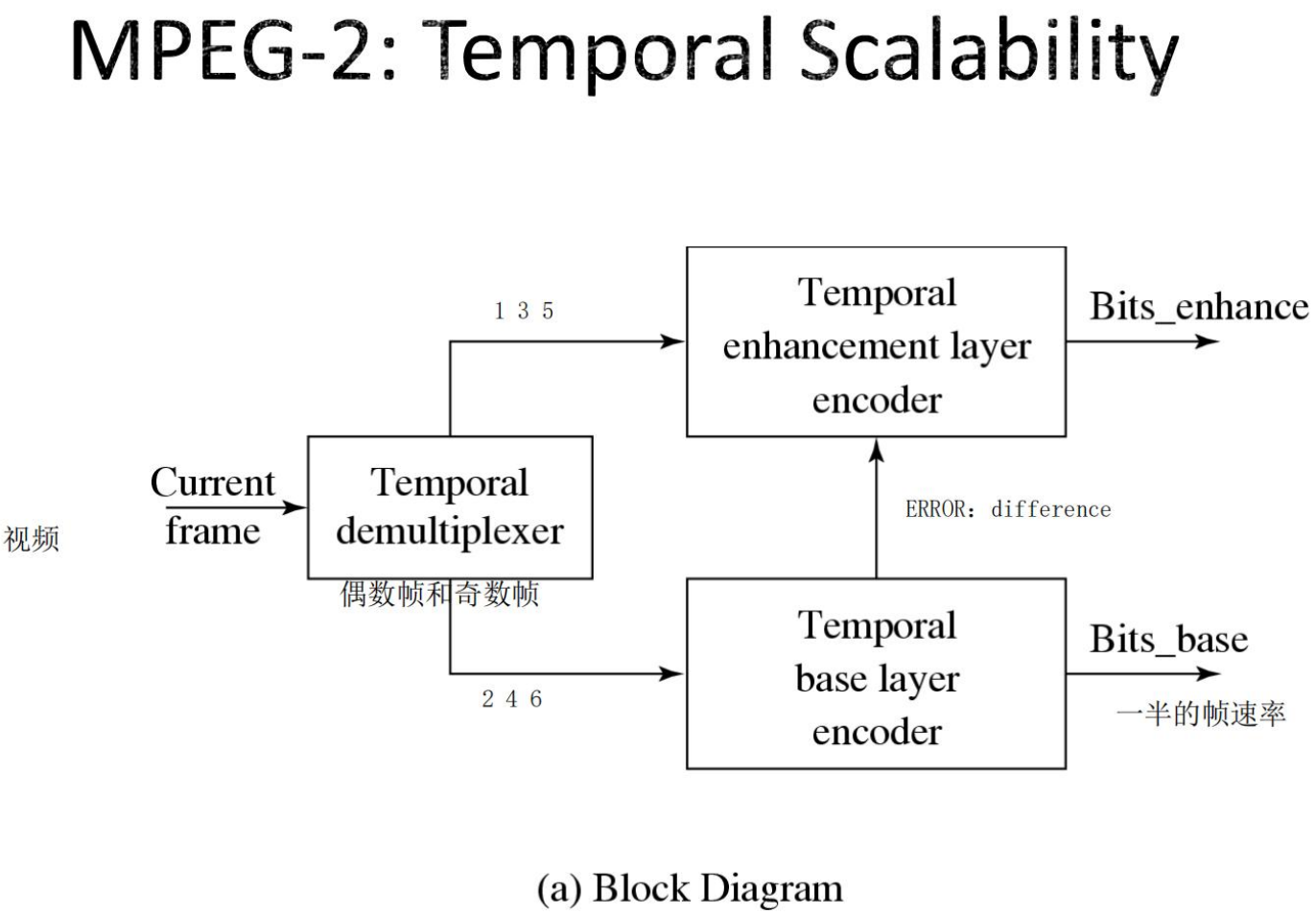
3.2.5.3.2.空间可扩展性

MPEG-2: Spatial Scalability 2:13:41

MPEG-2:空间可扩展性



3.2.5.3.3.时间可伸缩性



(a) Block Diagram

H.261: I-frame Coding

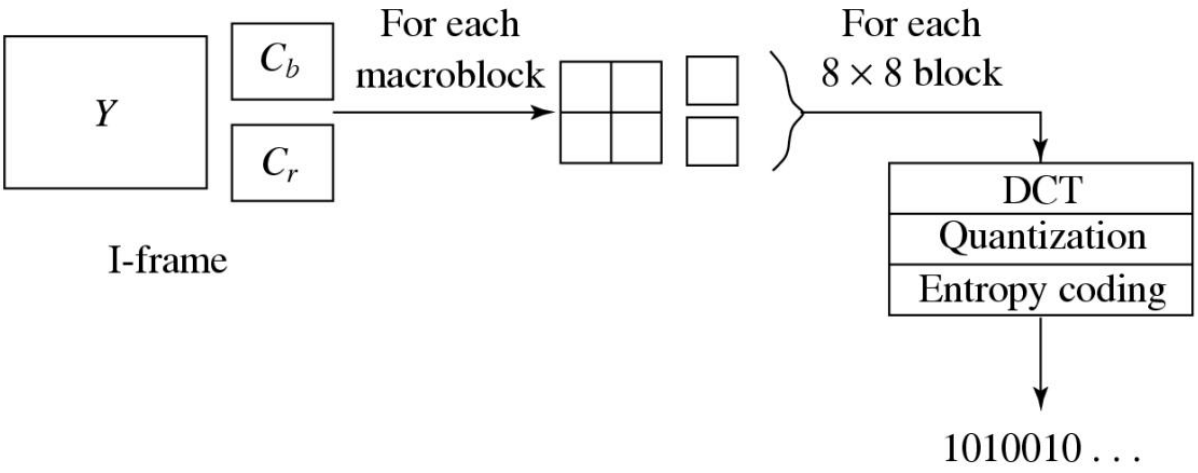
宏块包含4Y, 1Cb和1Cr的8 x 8像素块，用于4:2:0色度子采样

- A macroblock contains 4Y, 1Cb, and 1Cr of 8 x 8 pixel blocks for 4:2:0 chroma subsampling.
- Macroblocks are of size 16 x 16 pixels for the Y frame, and 8 x 8 for Cb and Cr frames.

对于Y帧，macroblock的大小为16 x 16像素，对于badcrres为8 x 8像素。

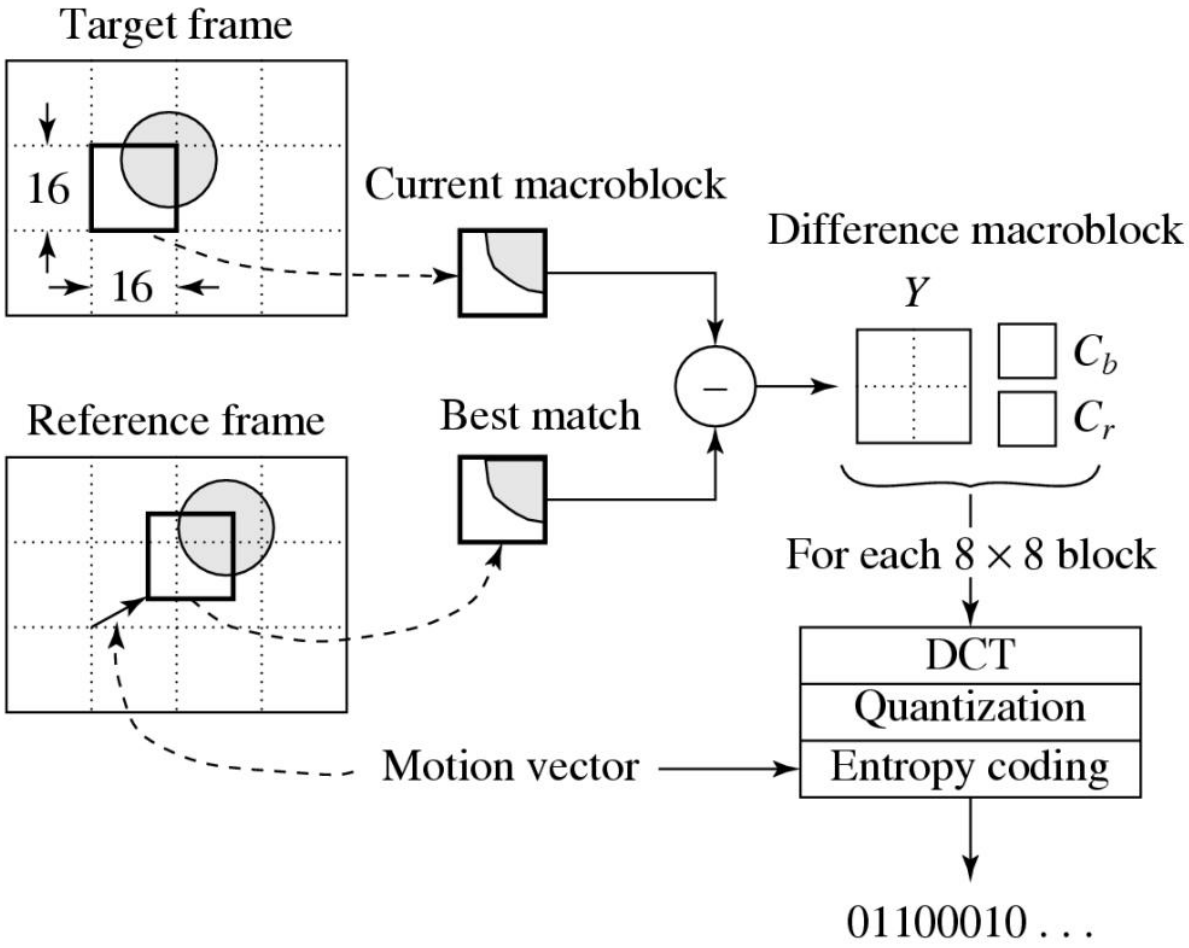
- For each 8 x 8 block, a DCT transform is applied, the DCT coefficients then go through quantization, zigzag scanning, and entropy coding.

对于每个8 x 8块，应用DCT变换，DCT系数然后经过量化，之字形扫描和熵编码。



I-frame Coding

H.261: P-frame Coding

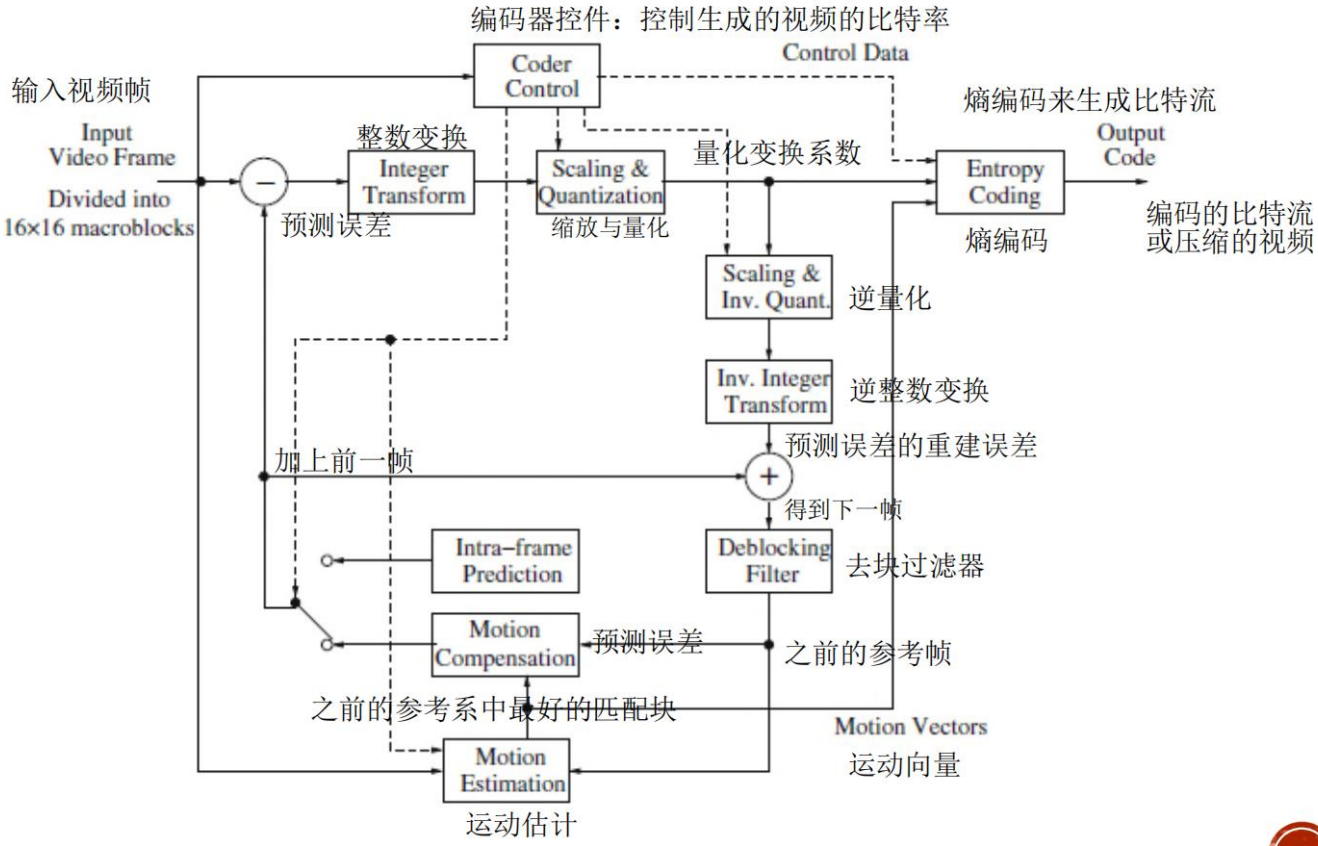


H.261 P-frame Coding Based on Motion Compensation

3.3.4.1 H264编码器

6427 week6
38:14 / 2:59:48

H.264: Encoder



4.1.5 众所周知的CNN架构

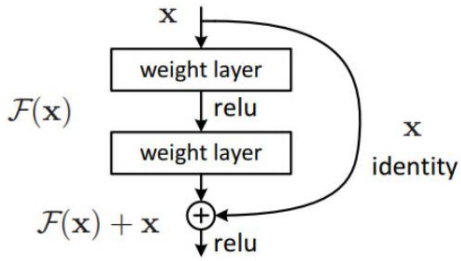
ResNet

消失梯度问题

继续向普通 CNN 添加许多不同的层，实际上性能在提高一段时间后就会开始下降

如果你将梯度堆叠在许多不同的层上，是的，你的早期层无法有效更新

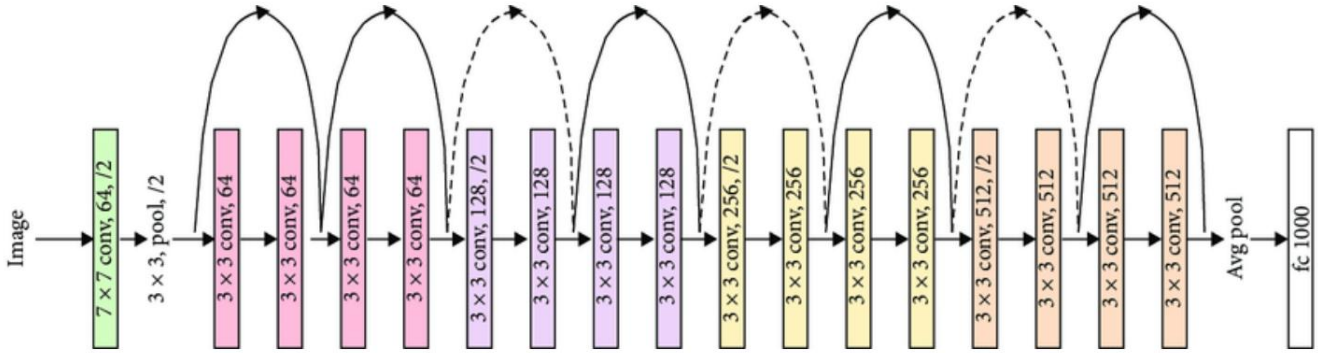
残差块



高速公路或跳过连接

高速公路连接来更有效地反向传播这个梯度

仍然可以有效地反向传播



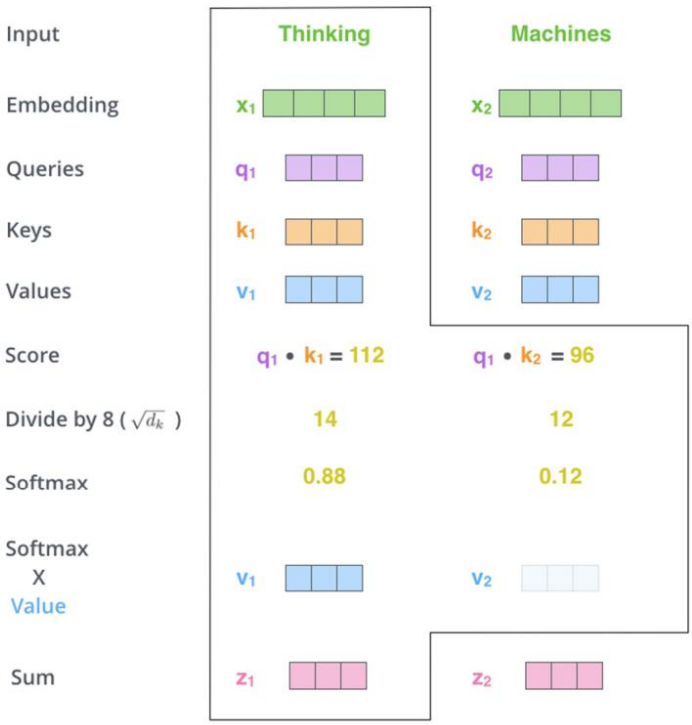
Scaled Dot-Product Attention 缩放的点积注意力

- Step 1: compute the correlation (dot product) between the query (q) and key (k) vectors.
- Step 2: correlation values from Step 1 are scaled and normalized using Softmax function.
- Step 3: multiplied output from Step 2 by corresponding value (v) vectors and sum them up.

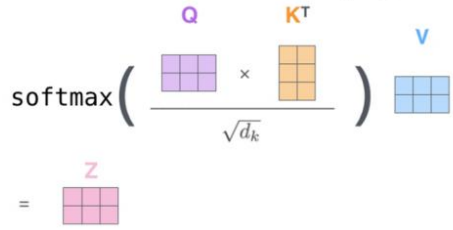
步骤1:计算查询(q)和键(k)向量之间的相关性(点积)。

步骤2:使用Softmax函数对步骤1的相关值进行缩放和归一化。

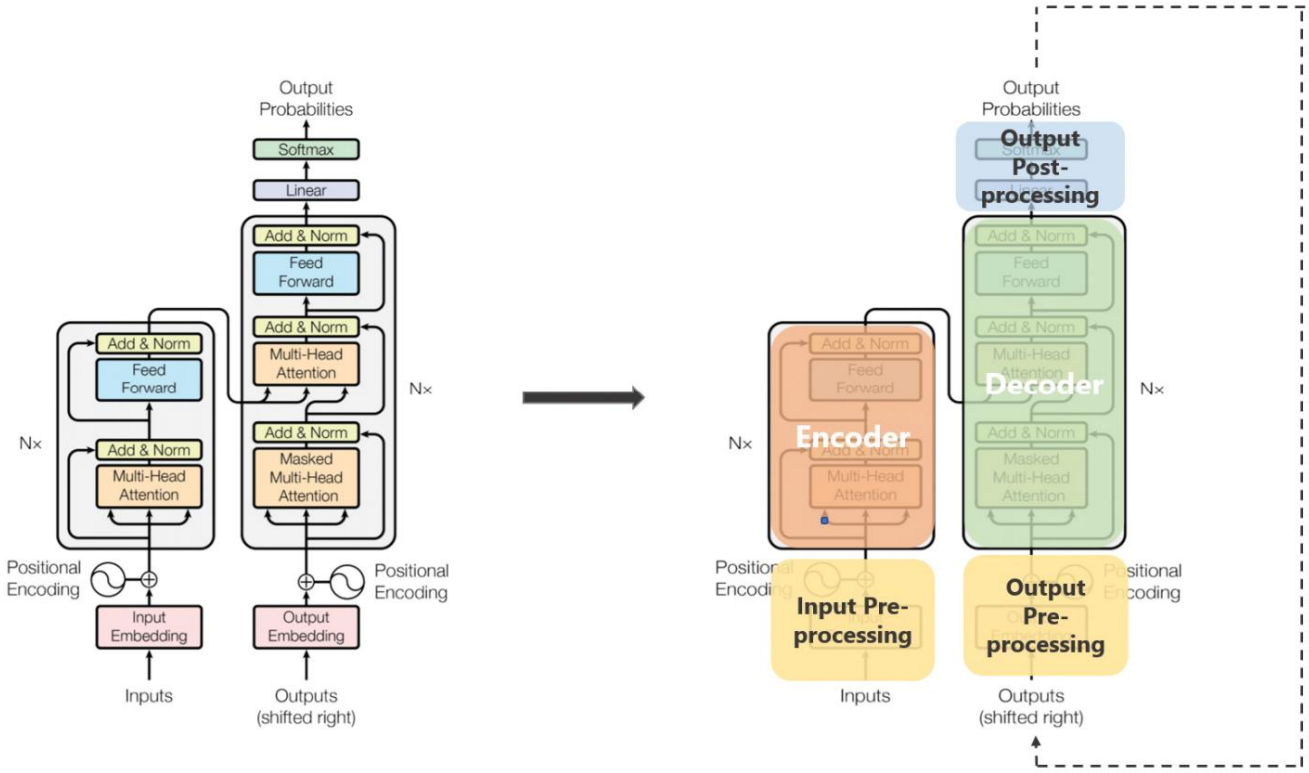
步骤3:将步骤2的输出与相应的值向量(v)相乘并求和。



$$A(Q,K,V) = softmax(\frac{QK^T}{\sqrt{d_k}})V$$



Transformer Architecture (1) 变压器结构



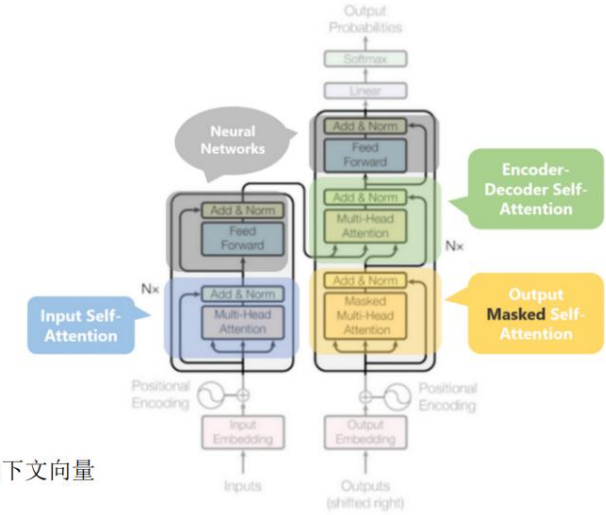
Source: Ria Kulshrestha, Transformers. https://towardsdatascience.com/transformers-89034557de14

4.3.2.5 自关注层概述

Self-Attention Layer Overview

自关注层概述

- Self-attention layer: $\mathbf{C} = \mathrm{Attn}(\mathbf{X}, \mathbf{X})$.
  - Inputs: $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m]$.
  - Parameters: $\mathbf{W}_Q$, $\mathbf{W}_K$, $\mathbf{W}_V$.



$\mathbf{c}_{:1}$ $\mathbf{c}_{:2}$ $\mathbf{c}_{:3}$ $\cdots$ $\mathbf{c}_{:m}$ 上下文向量

Self-Attention Layer (Parameters: $\mathbf{W}_Q$, $\mathbf{W}_K$, $\mathbf{W}_V$) QKV计算

$\mathbf{x}_1$ $\mathbf{x}_2$ $\mathbf{x}_3$ $\cdots$ $\mathbf{x}_m$ 向量

4.3.2.6 交叉注意/编码器-解码器自我注意