



## 交流图片

图像和视频编码课程

2014 年, 第 63-98 页

# 第3章-图像和视频的离散时间分析

大卫·R·布尔

显示更多 ▾

Outline | 共享 引用

<https://doi.org/10.1016/B978-0-12-405906-1.00003-9>

获取权利和内容

## 摘要

本章旨在概述离散时间信号和图像处理的重要原理，因为它们构成了几乎所有压缩算法的基础。它首先回顾了采样定理，并考虑二维（图像）和二维+时间（视频）采样的情况。接下来，它介绍了使用统计表示来描述信号的方法，重点是二阶统计，并说明了视频信号中存在的空间和时间相关性。过滤器和变换构成了大多数压缩系统的基本工作原理，并介绍了这些。接下来是讨论压缩期间由于滤波或转换后系数量化而引入的信号失真。预测算子在压缩的许多方面都很常用，并介绍了线性预测的基本概念，特别注意解码器漂移的缓解。这种架构很重要，因为它为所有当前的视频压缩系统提供了一个框架。最后，本章回顾了信息理论的基础，信息理论是当今所有符号编码方法的主题关键。通篇提供了示例。

上

下一

## 关键词

图像采样；图像统计；图像过滤；图像变换；量化；线性预测；信息理论

关于本章，首先要强调的是，它无意成为离散时间信号分析的课程。关于这个主题的许多优秀文本已经存在，涵盖1-D[1]和2-D[2]信号。我们假设读者熟悉采样、线性系统分析、数字滤波器、

变换、信息理论和统计信号处理的基础知识。本章的目的是概述这些重要主题，因为它们构成了本书中描述的许多压缩技术的基础。

我们首先回顾采样定理，并考虑二维（图像）和二维+时间（视频）采样的情况。这提供了代表我们图像或视频的离散时间样本，这些样本在压缩和编码的不同阶段进行处理。接下来，在第3.2节中，我们将介绍使用统计表示来描述信号的方法。我们专注于二阶统计，并使用这些来描述图像或视频信号中包含的冗余。正是这种冗余在压缩过程中被利用。

过滤器和变换是大多数压缩系统的基本工作马，这些在第3.3节中进行了介绍。大多数压缩系统中引入的信号失真可归因于滤波或转换后的系数量化；这一重要方面在第3.4节中涉及。预测算子在压缩的许多方面也很常用，从预测内部到运动估计。第3.5节介绍了线性预测的基本概念，我们特别注意了解码器漂移的缓解。这种架构很重要，因为它为所有当前的视频压缩系统提供了一个框架。最后，作为第7章的前身，我们回顾了信息理论的基础；这个主题是当今所有符号编码方法的关键。

### 3.1. 信号和图像采样

信号采样的方式对后续处理步骤有深远影响。特别是在图像或视频等感官信号的背景下，当在屏幕上查看时，样本必须提供基本（连续）自然场景的合适和现实的表示。因此，采样必须避免因别名而引入感知时空工件的方式进行。抽样过程解决了三个主要问题：

1. 需要哪些时空采样率来避免混叠并产生所需的信号保真度？
2. 样品应该如何在时空中分布？
3. 采样前需要哪些预处理，重建前需要哪些后处理？

这些问题受到一些因素的影响，包括：被采样信号的频率含量、观察者与当前观看条件相关的感知阈值、捕获和显示设备的特点以及与存储和传输相关的成本。本节试图回答这些问题。

#### 3.1.1。采样定理

##### 在一个维度

考虑一个连续的（让我们假设时域）信号 $x_a(t)$ 使用一系列δ函数采样 $\delta(t - nT)$ 产生：

$$x_s(t) = x_a(t)s(t) \quad (3.1)$$

其中：

$$s(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)$$

使用傅里叶变换，我们可以获得这些信号的频域表示，因此：

$$X_a(\omega) = \int_{-\infty}^{\infty} x_a(t)e^{-j\omega t} dt \quad (3.2)$$

和

$$S(\omega) = \frac{2\pi}{T} \sum_{k=-\infty}^{\infty} \delta(\omega - k\omega_s) \quad (3.3)$$

调用傅里叶变换的调制性质得到：

$$X_s(\omega) = \frac{1}{T} \sum_{k=-\infty}^{\infty} X_a(\omega - k\omega_s) \quad (3.4)$$

在哪里  $\omega_s$  表示连续频率变量。

因此，采样信号的频谱包括原始（连续时间）频谱的缩放复制版本。在此背景下，香农的采样定理捕获了采样要求，以避免别名[3]：

如果信号  $x_a(t)$  带状限制  $X_a(\omega) = 0$  为了  $|\omega| > \omega_m$ ，那么  $x_a(t)$  由其样品唯一确定，前提是  $\omega_s \geq 2\omega_m$ 。然后可以通过传递完全恢复原始信号  $x_s(t)$  通过理想的低通滤波器。

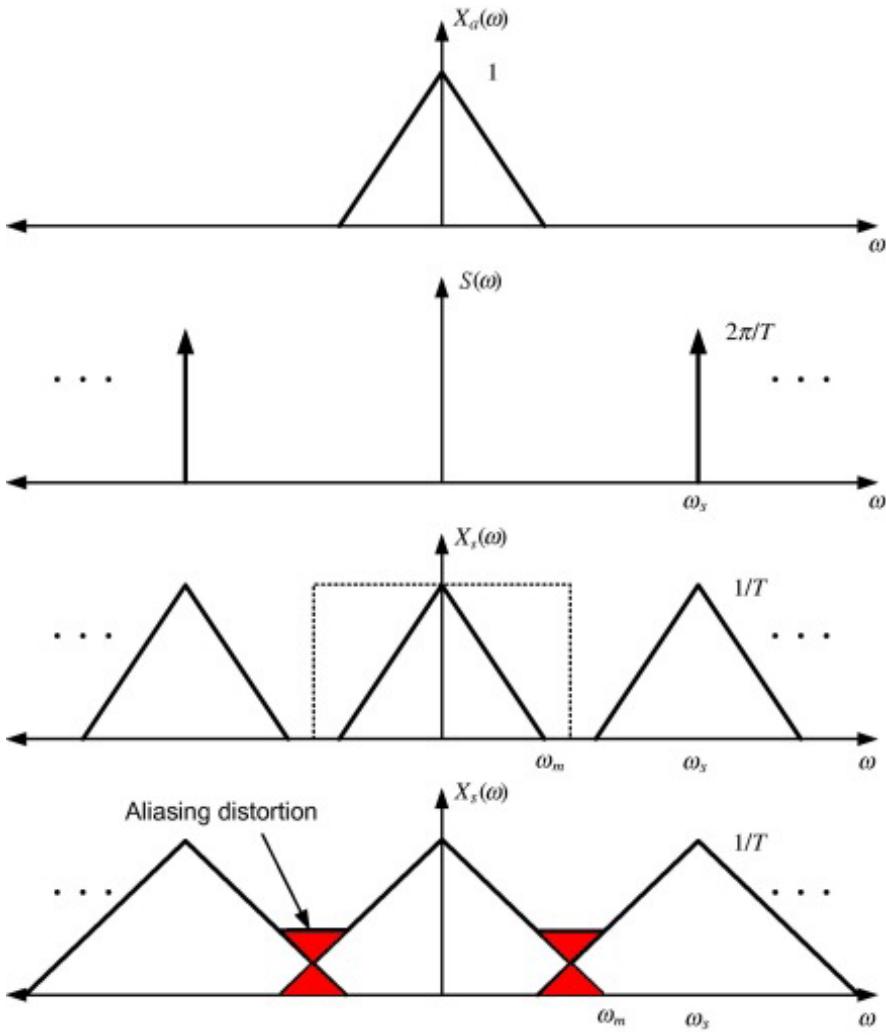
香农定理描述的采样过程的光谱特征如图3.1所示。倒数第二个子图特别注意底部的两个子图形，显示了使用理想重建过滤器可以完美重建的情况。相比之下，底部的子图显示了  $\omega_m$  超过频率  $\omega_s/2$  发生别名失真。频率  $\omega_s/2$  通常被称为奈奎斯特频率。

### 示例3.1—维信号混叠

考虑一个正弦信号  $x(t) = \cos(3\pi t/2)$  采样在  $\omega_s = \frac{4}{3}\omega_m$  与  $T = 1$ 。表明该采样违反了采样定理，并描述了别名在时域和频域中的影响。

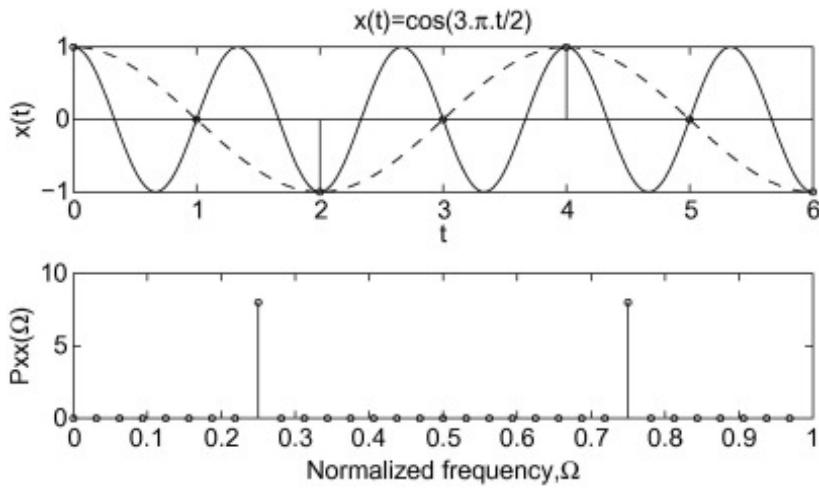
### 解决方案

结果如图3.2所示。正弦波  $x(t) = \cos(3\pi t/2)$  在  $\omega_s = \frac{4}{3}\omega_m$  与  $T = 1$ 。可以观察到，别名的频率为  $\frac{\omega_s}{4}$ （准确地说，这实际上是正弦波负频率分量的别名  $-\omega_m$ ）。这可以在信号域图（顶部）和频域图（底部）中观察到。在这里，按照离散时间系统的常规，我们绘制了归一化频率  $\Omega = \omega T$ ，其中采样由  $\Omega = 2\pi$ 。



[下载：下载全尺寸图像](#)

图3.1。采样和混叠的光谱特征。



[下载：下载全尺寸图像](#)

图3.2。演示一维信号的别名。上：奈奎斯特频率以下的正弦采样。底部：显示光谱别名的傅里叶变换。

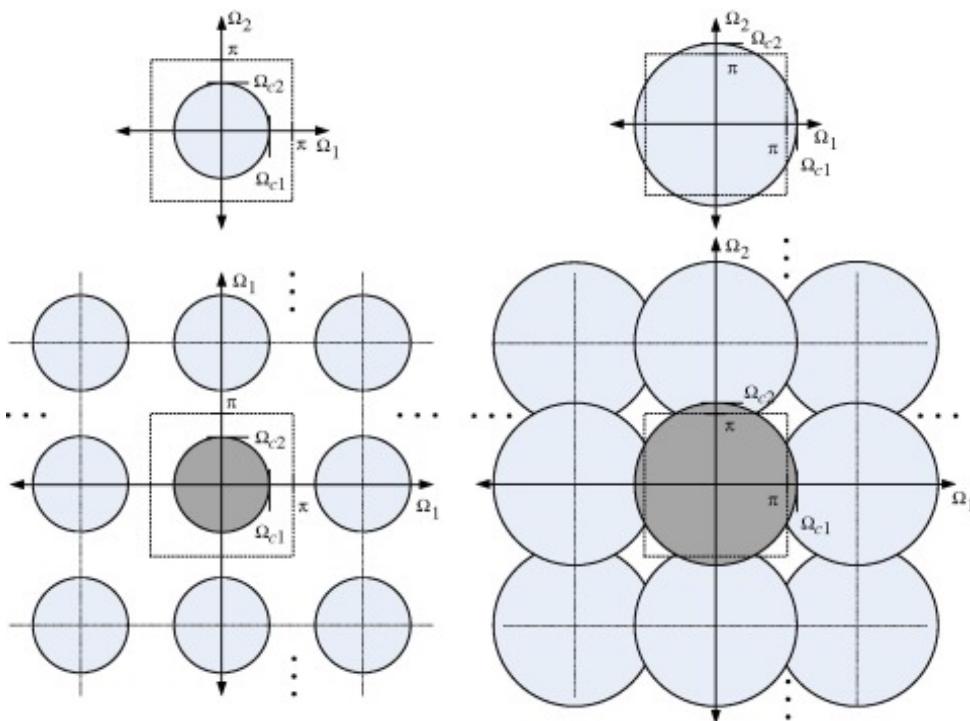
叶情节。

## 扩展至二维

假设二维矩形（正交）采样模式，我们可以将一维方法扩展到二维信号采样的情况。这给出了：

$$X_s(\omega_1, \omega_2) = \frac{1}{T_1 T_2} \sum_{k_1, k_2} X_a(\omega_1 - k_1 \omega_s, \omega_2 - k_2 \omega_s) \quad (3.5)$$

采样和混叠对二维光谱的影响如图3.3所示。在二维图像中，混叠误差将呈现垂直于高频或锐利边缘的方向的铃声。应该强调的是，虽然图3.3显示了圆形对称光谱，但实际上，该光谱可以根据图像中流行的水平和垂直频率或纹理进行倾斜。



[下载：下载全尺寸图像](#)

图3.3。采样和混叠的二维光谱特征。左：顶部——原始信号谱；底部——无别名的采样信号谱。右：顶部-原始信号谱；底部-采样信号光谱，由于亚奈奎斯特采样而具有别名。

## 扩展至三维

在时空采样的情况下，通常使用渐进式或交错式采样以固定帧速率采样（见第4章）。正如我们在第二章中所看到的，空间和时间维度的视觉敏感性大不相同。例如，传统的1080p/50 4:4:4格式将每秒捕获103,680,000个样本。然而，这包括每张图片的2,073,600个样本，但每秒只有50张图片。因此，空间域和时间域之间单位维度的样本数量存在很大差异。似乎时间域得到了原始交易，在某种程度上这是真的。我们将在第13章中回到这个。

如[第二章](#)所述，视觉敏感性取决于显示器的平均亮度。对于明亮的电视信号，通常根据时间[对比度灵敏度函数](#)和临界闪烁频率，认为50-70Hz的时间更新率就足够了，而对于在明亮环境中更仔细查看的计算机显示器，通常指定更高的更新率。

### 例3.2 我们需要多少个空间样本？

以16:9屏幕为例，宽度为1米，在 $3H$ （在哪里 $H$ 是屏幕的高度）。需要多少水平样本才能满足[人类视觉系统的对比灵敏度限制](#)？

### 解决方案

对于宽度为1米的16:9屏幕，请查看 $3H$ ，我们可以从简单的三角学中计算出所需的水平样本数量。屏幕高度为0.5625米，然后半视角是：

$$\theta = \tan^{-1}\left(\frac{0.5}{1.6875}\right)$$

假设空间分辨率限制为30 cpd（见[第2章](#)），每行产生1980个水平样本，这与实际使用的1920年非常接近。

虽然这种计算很有用，但随着屏幕尺寸的增加和观看距离的减少，将需要重新审视时空采样，但更多将在[第13章](#)之后进行。

### 3.1.2。多维采样格子

实值的采样格 $K$ -维空间，是从一组的[线性组合](#)可以形成的所有可能的向量的集合。 $K$ 线性独立基向量， $\mathbf{v}_k \in \mathcal{R}^K$ ,  $k = \{1, 2, 3, \dots, K\}$ 。因此，新的向量由加权组合组成如下：

$$\mathbf{w} = \sum_{k=1}^K c_k \mathbf{v}_k; \quad \forall c_k \in \mathcal{Z} \quad (3.6)$$

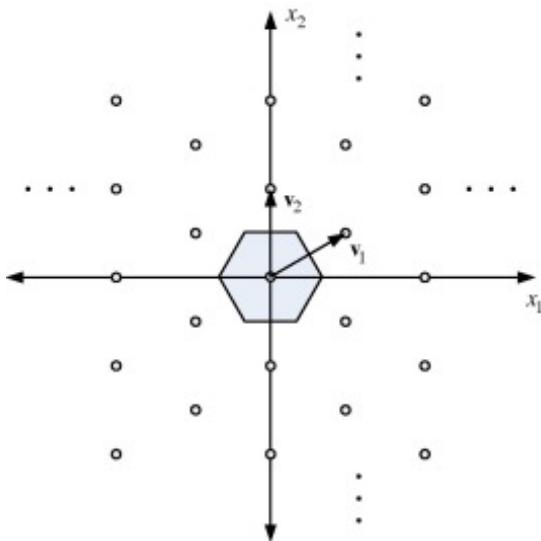
矩阵 $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K]$ 结合 $K$ 基向量称为[采样矩阵](#)或生成矩阵。例如，我们熟悉的矩形格的生成矩阵是：

$$\mathbf{v}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (3.7)$$

然而，我们可以探索更奇异的采样网格及其相关支持区域。例如，考虑生成矩阵：

$$\mathbf{v}_2 = \begin{bmatrix} \sqrt{3}/2 & 0 \\ 1/2 & 1 \end{bmatrix} \quad (3.8)$$

该生成矩阵的采样格如[图3.4](#)所示。可以看到，这代表一个六边形采样网格。晶格结构是生成和分析结构中不一定超立方体的采样结构的简单方便的工具。一维和二维采样的所有定理仍然适用，香农采样理论存在一种广义形式，该形式决定了采样网格的结构和密度取决于信号谱。与这种与图像和视频应用相关的更通用的多维采样方法相关的大量工作。读者参考杜布瓦[4]或王等人的作品，以了解这个主题的精彩介绍。



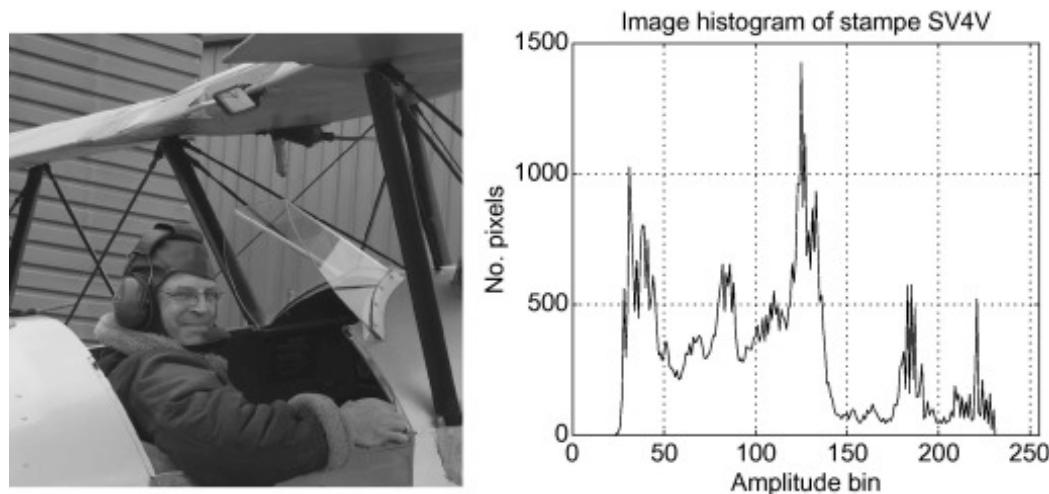
[下载：下载全尺寸图像](#)

图3.4。六边形采样格及其方程 (3.8) 定义的倒数。

## 3.2。图像统计

### 3.2.1. 直方图和分布

直方图是基于量化像素值的实际发生的概率密度函数的离散近似值。它可以提供有关图像中样本振幅分布、信号的动态范围与其表示的字长，以及图像曝光不足还是过度的重要信息。直方图示例  $256 \times 256$  图3.5给出了图像Stampe\_SV4。



[下载：下载全尺寸图像](#)

图3.5。示例图像直方图 $256 \times 256$ 图片Stampe\_SV4。

## 空间和子带分布

了解空间域和频域分布对压缩系统有重大益处。例如，第11章中描述的金字塔矢量量化方法利用了类拉普拉斯子带系数分布，而第13章概述的参数视频编码方法利用了双树复合小波系数的子带结构，以提供分段到相干纹理区域。

### 3.2.2。平均值

平均值和相关性从随机过程的平均值和矩的角度提供了部分统计特征。它们非常有用，因为它们提供了易于处理的数学分析，易于进行实验评估，并且非常适合随机过程线性操作的表征。

随机过程的平均值由以下方式给出：

$$\begin{aligned}\mu_n &= E\{x[n]\} \\ &= \int_{-\infty}^{\infty} x[n]f(x[n])dx[n]\end{aligned}\tag{3.9}$$

在哪里 $f(\cdot)$ 是过程的概率密度函数 (pdf) 和 $E\{\cdot\}$ 是期望运算符。在静止过程中，随机变量的pdf对所有都相同 $n$ 因此，平均值是恒定的。期望可以解释为通过多次重复实验获得的平均值。因此：

$$\mu_n = E\{x[n]\} = \lim_{N \rightarrow \infty} \left[ \frac{1}{N} \sum_{i=1}^N x_i[n] \right] \tag{3.10}$$

### 3.2.3。自然图像中的相关性

序列的自相关（或自协方差）表示其样本之间的线性统计依赖关系。它被定义为滞后于 $m$ 样本为：

$$R_{xx}[m] = E\{x[n]x[n+m]\} \tag{3.11}$$

在哪里 $x[n]$ 是一个静止的随机过程。特别是：

$$R_{xx}[0] = E\{|x[n]|^2\} \tag{3.12}$$

是信号的平均功率。在实践中，相关性是根据有限长度序列估计的 $x$ ，的 $N$ 样本如下：

$$R_{xx}[m] = \frac{1}{N} \sum_{n=0}^{N-|m|-1} x[n]x[n+m] \tag{3.13}$$

方程 (3.13) 产生有偏差的估计，因为大滞后的衰减。也可以使用无偏见的估计，这更吵，但在某些情况下更可取：

$$R_{xx}[m] = \frac{1}{N-|m|} \sum_{n=0}^{N-|m|-1} x[n]x[n+m] \tag{3.14}$$

自协方差的计算方式与自相关相同，但信号手段被移除。当自相关函数或自协方差函数被其最大值归一化时，它们通常分别称为自相关系数或自协方差系数。这些值在-1和+1之间。例如，自相关系数由<sup>1</sup>给出。

$$\rho_x(k) = \frac{r_x(k)}{r_x(0)} \quad (3.15)$$

自相关矩阵包括不同滞后的自相关值。广义平稳 (WSS) 实值信号的相关矩阵是信号向量与自身的外乘积的期望。因此：

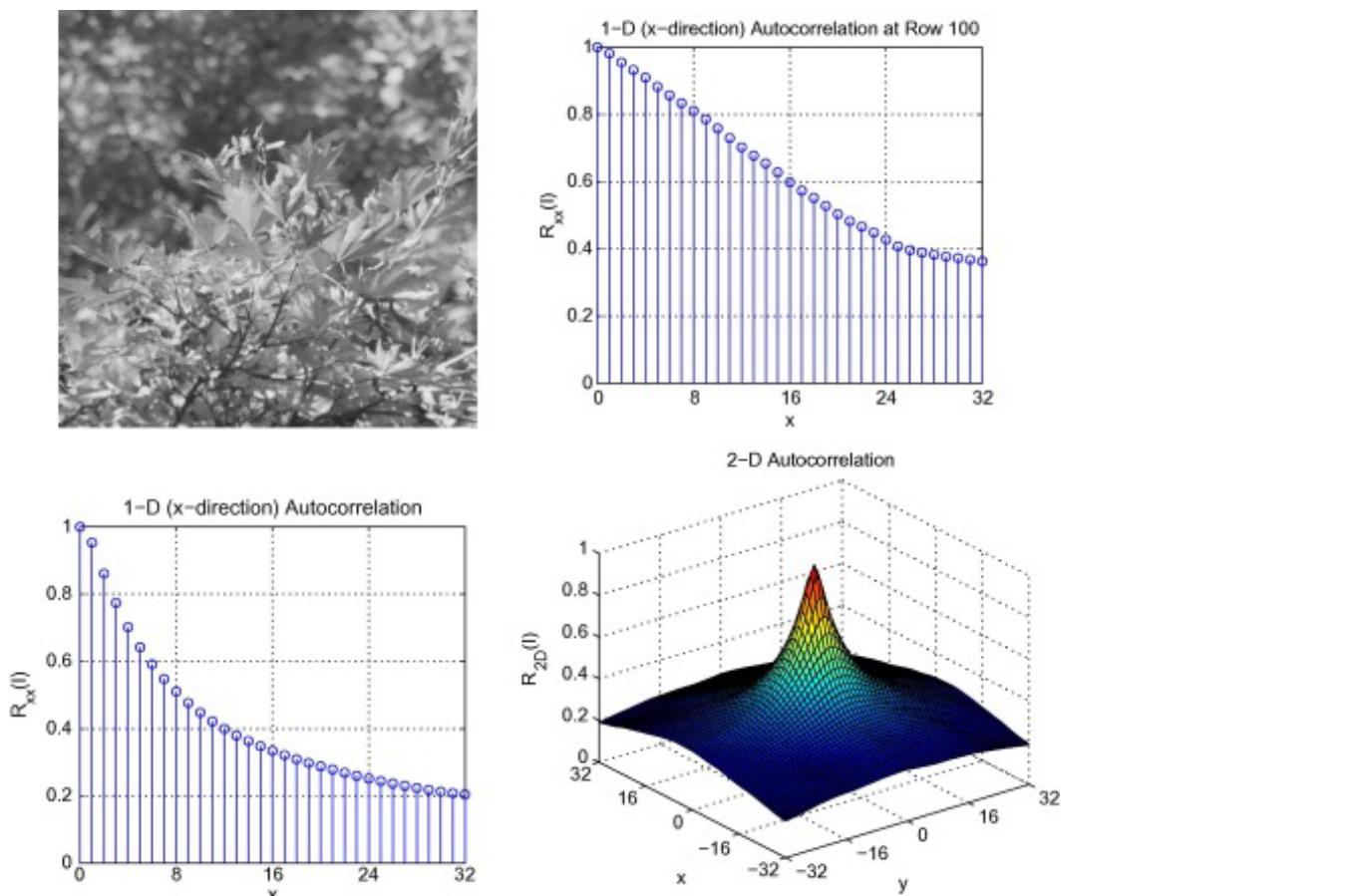
$$\mathbf{R}_x = E\{\mathbf{x}\mathbf{x}^T\} \quad (3.16)$$

$$\mathbf{R}_x = \begin{bmatrix} r_x[0] & r_x[1] & \cdots & r_x[M-1] \\ r_x[1] & r_x[0] & & r_x[M-2] \\ \vdots & & & \vdots \\ r_x[M-1] & r_x[M-2] & \cdots & r_x[0] \end{bmatrix} \quad (3.17)$$

相关矩阵是托普利茨和埃尔米特矩阵。对于实值信号，这意味着它是转置不变的。

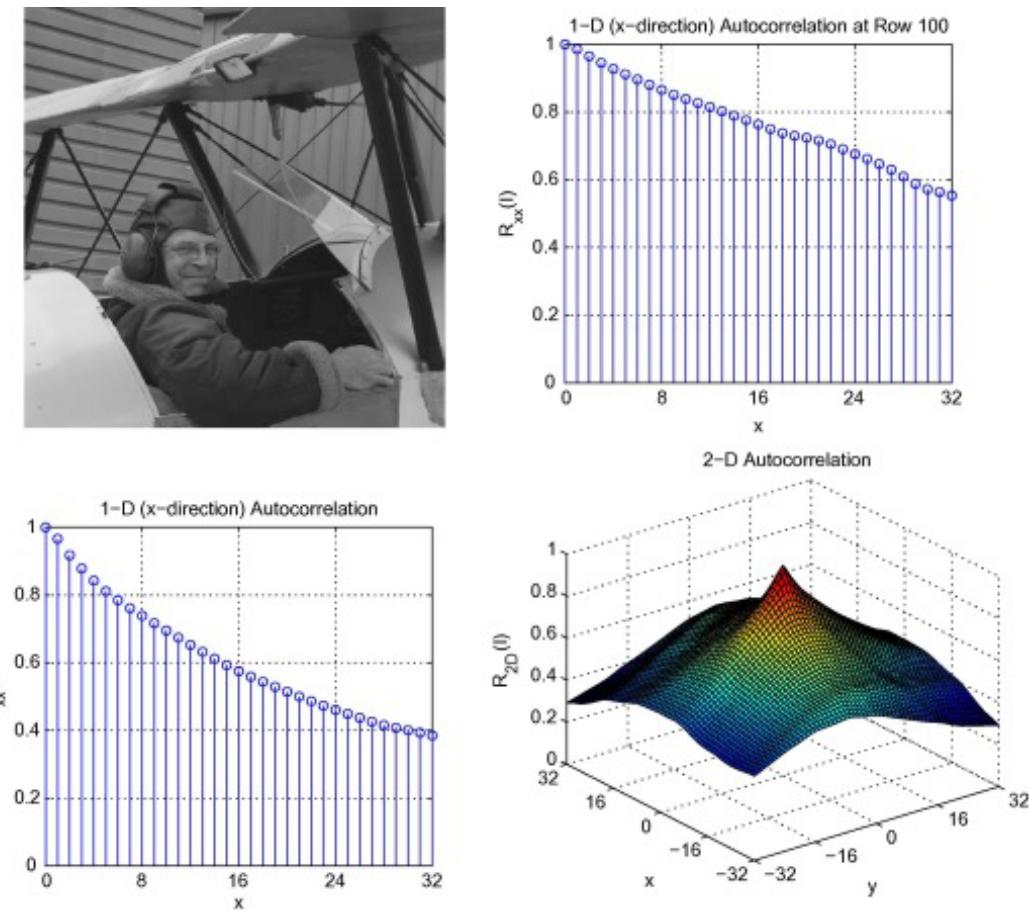
## 自然图像中的空间自相关

现在让我们考虑一些图像的相关性质。[图3.6](#), [图3.7](#)显示了两个的自相关性质 $512 \times 512$ 图像——宏 ?? 和 *Stampe\_SV4*。第一个具有高度纹理、空间活跃的前景，背景模糊，第二个高度结构化，有几个平原区域。除了原始图像外，每个图形包括 (i) 单行像素滞后高达32像素的自相关图（无偏零平均值），(ii) 基于图像中所有行平均数的类似图，以及 (iii) 二维版本。可以观察到，在这两种情况下，相关性都会随着滞后增加而逐渐下降，宏 ?? 图像的值下降速度比 *Stampe\_SV4* 图像的值下降得更快。这是因为后者的结构与前者更随机。[图3.8](#)显示了类似的一系列情节，但这次是为了 $256 \times 256$ 图像，仅显示最多滞后16的值。正如人们所料，自相关值以与 [图3.7](#)相似的方式下降，只是相应的滞后值减半。



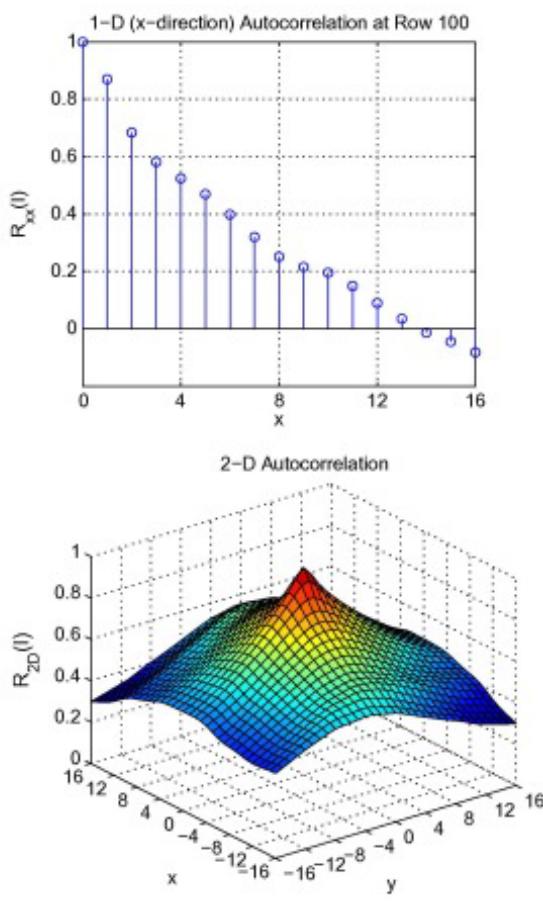
下载： [下载全尺寸图像](#)

图3.6。宏 ?? 图像的自相关图 ( $512 \times 512$ )。左上至右下：原始图像；第 100 行的自相关函数；整个图像的自相关函数；二维自相关表面。



[下载：下载全尺寸图像](#)

图3.7。Stampe\_SV4图像的自相关图 ( $512 \times 512$ )。左上至右下：原始图像；第 100 行的自相关函数；整个图像的自相关函数；二维自相关表面。



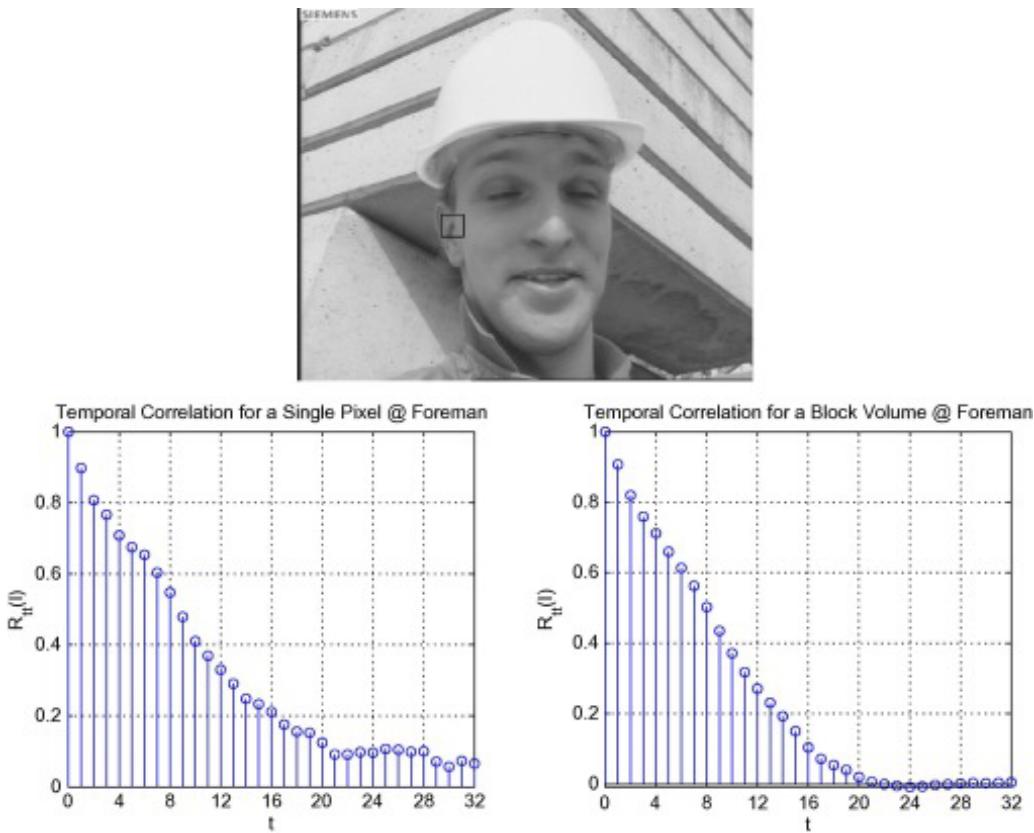
[下载：下载全尺寸图像](#)

图3.8。Stampe\_SV4图像的自相关图（ $256 \times 256$ ）。左上至右下：原始图像；第 100 行的自相关函数；整个图像的自相关函数；二维自相关表面。

利用这种样本间相关性是预测、转换和滤库压缩方法的基础。在每种情况下，变换或预测器的目的是修饰空间信息，使其更易于与频率相关的感知量化。正如我们从图表中所看到的，大滞后的自相关值变得越来越小，而小滞后的自相关值则很高。由于装饰变换和预测器在相关性较高的区域效果更好，因此它们通常适用于小区域（通常 $8 \times 8$ 块）而不是整个图像。安 $8 \times 8$ 传统上，块被选为计算效率和编码增益之间的良好妥协。然而，值得注意的是，图像分辨率将对此产生重大影响——因为低分辨率的图像将显示相同大小块的低相关性。

### 自然图像序列中的时间自相关

现在让我们考虑时间轴。工头序列的时间相关性示例如图3.9所示。这表明与上述空间情况相似。展示了两个情节。第一个是使用工头序列的300帧计算的单个像素的时间相关性。第二个是相同的计算，但这次平均超过 $16 \times 16$ 在顶部子图中指示的像素块。同样，我们可以观察到相邻的视频帧之间存在显著的时间相关性，实际上超过8帧，相关性仍然相对较高。这表明大多数图像序列存在显著的时间冗余。正如我们稍后将看到的，在转换之前，通过应用运动补偿预测，大多数压缩方法都会利用这一点。



[下载：下载全尺寸图像](#)

图3.9。工头的时间自相关图（30 fps）。右上至下：显示选定的样本框架 $16 \times 16$ 使用的块；单个像素的时间相关性； $16 \times 16$ 阻止。

### 3.3。过滤和转换

我们假设读者熟悉基本的线性系统理论、滤波器和变换。对于背景阅读，读者请参阅参考文献。[\[1\]](#), [\[2\]](#), [\[8\]](#), [\[9\]](#)。

#### 3.3.1。离散时间线性系统

我们在这本书中遇到的许多系统都表现出线性和移位不变性的性质。[2](#) 让我们简要描述这些性质。

#### 移位不变性

移位不变性系统是指输入信号自变量的移位导致输出信号的相应移位。那么如果一个系统对一个输入的响应 $x_0[n]$ 是 $y_0[n]$ 然后是对输入的响应 $x_0[n - n_0]$ 是 $y_0[n - n_0]$ 。

#### 线性

如果系统对输入的响应 $x_0[n]$ 是 $y_0[n]$ 以及对输入的响应 $x_1[n]$ 是 $y_1[n]$ 那么一个系统被称为线性，如果：

- 对  $x_0[n] + x_1[n]$  是  $y_0[n] + y_1[n]$  (加性)。
- 对  $ax_0[n]$  是  $ay_0[n]$ , 在哪里  $a$  是任何复常数 (缩放)。

推论是叠加原则, 该原则指出, 如果  $x[n] = ax_0[n] + bx_1[n]$  那么线性移位不变系统的输出将是  $y[n] = ay_0[n] + by_1[n]$ 。

### 3.3.2。卷积

如果线性时不变 (LTI) 系统对单位脉冲的响应,  $\delta[n]$ , 是  $h[n]$  然后  $h[n]$  被称为系统的脉冲响应。LTI系统的移位不变性、线性和筛选特性给出了:

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k] \quad (3.18)$$

或者:

$$y[n] = x[n] \star h[n]$$

这被称为 **卷积和**, 可用于确定离散时间 LTI 系统对任意输入序列的响应。

#### 示例3.3卷积

使用卷积方程 (3.18), 确定脉冲响应线性系统的响应  $h[n] = \{1, 2, 1\}$  输入  $x[n] = \{1, 2, 0, 1, -1, 2\}$ 。

#### 解决方案

卷积和体现了叠加原理。因此, 我们可以从每个输入样本的响应之和中形成系统整体输出, 如下所示:

$n$	0	1	2	3	4	5	6	7
$y_0[n]$	1	2	1					
$y_1[n]$		2	4	2				
$y_2[n]$			0	0	0			
$y_3[n]$				1	2	1		
$y_4[n]$					-1	-2	-1	
$y_5[n]$						2	4	2
$y[n]$	1	4	5	3	1	1	3	2

### 3.3.3。线性滤波器

离散时间线性系统的传递函数由以下方法给出：

$$H(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_M z^{-M}}{b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_N z^{-N}} \quad (3.19)$$

自  $x[n - k] \xrightarrow{Z} z^{-k} X(z)$ , 等效差分方程是 (与  $b_0$  缩放到统一) :

$$\begin{aligned} y[n] &= a_0 x[n] + a_1 x[n - 1] + a_2 x[n - 2] + \cdots \\ &\quad + a_M x[n - M] \\ &\quad - b_1 y[n - 1] - b_2 y[n - 2] - \cdots - b_N y[n - N] \end{aligned} \quad (3.20)$$

或者：

$$y[n] = \sum_{k=0}^M a_k x[n - k] - \sum_{k=1}^N b_k y[n - k] \quad (3.21)$$

在  $\{b_k\} = 0$  后，滤波器被描述为具有 **有限脉冲响应** (FIR)。在其他情况下，它被称为 **无限脉冲响应** (IIR)。**FIR过滤器** 在图像和视频处理中比IIR过滤器更常见。原因主要是因为它们的相位反应。我们在**第二章** 中看到，相位失真对**感知质量**有重要影响，设计具有**线性相位**特性的FIR滤波器很简单。这种特性只引入了滤波器响应的简单移位，没有相位失真。稍后我们将提供一些用于视频压缩的实用FIR滤波器的例子，但首先让我们更详细地检查一下滤波器频率响应。

### 扩展至二维

信号的二维卷积  $x[m, n]$  带有冲动的反应  $h[m, n]$  由以下方式给出：

$$y[m, n] = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} x[m - k, n - l] h[k, l] = x[m, n] \star \quad (3.22)$$

$h[m, n]$

二维**数字滤波器** 广泛用于图像和视频编码，例如与**采样率**变化相结合，在**滤波器组**的分析或合成阶段（**第6章**），用于子像素运动估计的**插值滤波器**（第8章和第9章，**第8 and 章第9章**）或更一般的**处理前**和**后处理操作**中。然而，出于灵活性和复杂性的原因，这些尽可能可以单独实施；即作为一维过滤阶段的级联。

### 可分离性

**可分离性** 是数字滤波器和变换的一个重要特性，在图像和视频处理中得到广泛利用。当（对于实值系数）：时，系统是可分离的：

$$\mathbf{h} = \mathbf{h}_1 \mathbf{h}_2^T \quad (3.23)$$

在哪里  $h_1[m]$  是一维脉冲响应，在信号行之间操作，并且  $h_2[n]$  是向下操作列的相应脉冲响应。我们现在可以将方程 (3.22) 重写如下：

$$y[m, n] = \sum_{k=-\infty}^{\infty} h_1[k] \sum_{l=-\infty}^{\infty} x[m - k, n - l] h_2[l] \quad (3.24)$$

如果我们定义：

$$y_2[m,n] = \sum_{l=-\infty}^{\infty} x[m,n-1]h_2[l]$$

然后我们可以将方程 (3.24) 重写为：

$$y[m,n] = \sum_{k=-\infty}^{\infty} y_2[m-k,n]h_1[k] \quad (3.25)$$

因此，这意味着如果方程 (3.23) 成立，那么我们可以通过首先过滤二维信号的列来实现二维滤波  $h_2[n]$ ，然后使用  $h_1[m]$ 。由于我们正在级联线性操作，因此在必要时可以互换这两个阶段。

应当指出，并非所有滤波器都是可分离的，可分离性确实对滤波器特性施加了一些方向限制。通常，如果过滤器的所有行和列都线性相关，则过滤器是可分离的。这个属性可以在示例3.4的解决方案中观察到。

### 示例3.4过滤器可分离性

考虑以下两个滤波脉冲响应：

$$\mathbf{h}_1 = [0.5 \ 0.3 \ 0.2]^T$$

$$\mathbf{h}_2 = [0.6 \ 0.1 \ 0.3]^T$$

形成等效二维数字滤波器。

### 解决方案

二维过滤器由以下方式给出：

$$\mathbf{h} = \mathbf{h}_1 \mathbf{h}_2^T = \begin{bmatrix} 0.30 & 0.05 & 0.15 \\ 0.18 & 0.03 & 0.09 \\ 0.12 & 0.02 & 0.06 \end{bmatrix}$$

为进一步练习，请满足自己二维过滤操作与使用可分离过滤器相同。

### 3.3.4。滤波频率响应

从我们对卷积和数字滤波的理解中，我们知道离散线性系统对采样正弦输入信号的响应将是输入信号的多个加权和移位版本。因此，输出也是相同频率的正弦信号，但振幅和相位不同。因此，我们可以将频率域中的滤波器视为振幅和相位的频率相关修饰符。如果我们用一个复指数的输入信号推广它：

$$x[n] = e^{j\Omega n}$$

然后从卷积方程 (3.18) 中，我们有：

$$y[n] = \sum_k h[k]e^{j\Omega(n-k)} = e^{j\Omega n} \sum_k h[k]e^{-j\Omega k}$$

因此：

$$y[n] = x[n]H(\Omega)$$

其中：

$$H(\Omega) = \sum_{k=-\infty}^{\infty} h[k]e^{-j\Omega k} \quad (3.26)$$

术语 $H(\Omega)$ 描述复指数在经过系统时所经历的振幅和相位变化，称为系统的频率响应。

### 3.3.5。实用过滤器示例

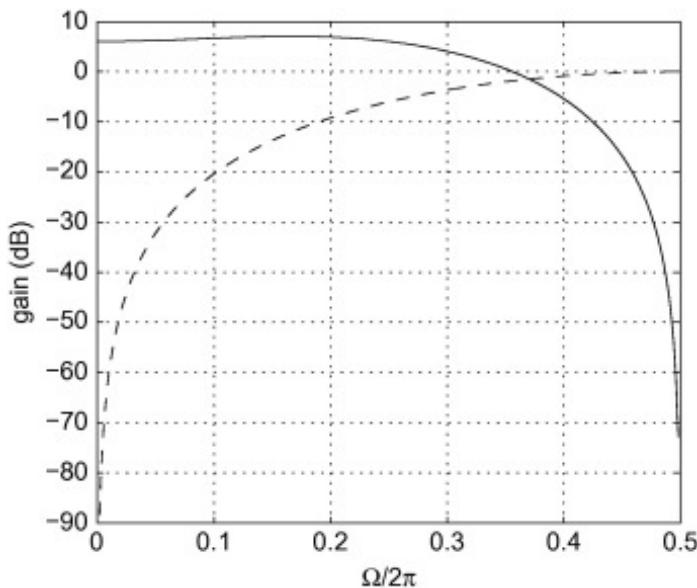
#### 勒格尔小波分析滤波器

在子采样和量化之前，用于将信号分析为低通和高通子带的一组常见的可分离滤波器是5/3 LeGall滤波器。低通滤波器和高通滤波器的脉冲响应为：

$$\mathbf{h}_0 = [-1, 2, 6, 2, -1]/4$$

$$\mathbf{g}_0 = [1, -2, 1]/4$$

这些过滤器的频率响应如图3.10所示。有趣的是，低通滤波器和高通滤波器不对称，它们之间似乎存在明显的光谱泄漏。然而，正如我们将在第6章中所看到的，这些过滤器确实性能良好，当组合到过滤器库中时，它们提供了完美重建的特性。



[下载：下载全尺寸图像](#)

图3.10。LeGall低通和高通分析过滤器的过滤器库响应。

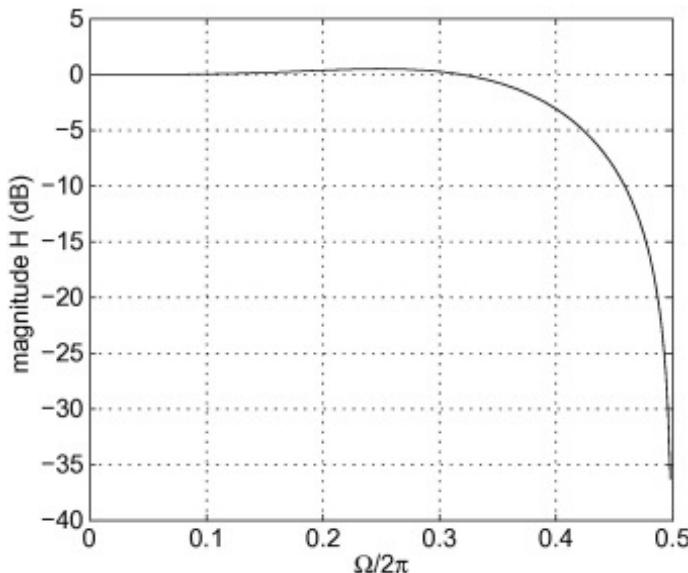
#### H.264/AVC半像素插值滤波器

第二个重要的过滤操作是插值，例如，在进行子像素运动估计之前，对图片的一个区域进行上采

样（见第8章和第9章第8 and 章第9章）。在目前部署的视频编码标准H.264/AVC中，采用了六次FIR过滤器：

$$\mathbf{h} = [1, -5, 20, 20, -5, 1]/32 \quad (3.27)$$

该滤波器的频率响应如图3.11所示。可以观察到，该特性与勒高低通分析滤波器没有区别。然而，在这种情况下，过滤器对原始整数像素分辨率样本没有影响，它只是用于生成插值。因此，滤波器特性是快速滚出的高截止频率和低通带波纹波之间的权衡。这种权衡在过滤器设计中很常见[1]。



[下载：下载全尺寸图像](#)

图3.11。H.264半像素插值滤波器的滤波器响应。

### 3.3.6。非线性滤波器

非线性滤波器不基于输入和通过系统函数输出之间的线性关系。相反，它们代表了更广泛的一类操作，没有任何显式频域传输函数。非线性滤波器在一系列图像处理和编码应用中，例如：去噪、边缘保持操作和某些形式的预测。例如，HEVC和H.264/AVC压缩标准中使用的解块过滤器是第9章中描述的内容自适应非线性过滤器。

一些运动矢量预测应用中使用的一类常见滤波器是秩阶滤波器，特别是中位数滤波器。下面我们简单介绍一下这些。

#### 秩序和中位数过滤器

秩阶滤波器是将输入样本进行排序并根据其在排名列表中的位置选择输出的运算符。以下是秩阶过滤器的例子：

$$y[m,n] = \min_{\mathcal{N}(m,n)} (x(i,j)) \quad (3.28)$$

$$y[m,n] = \max_{\mathcal{N}(m,n)} (x(i,j)) \quad (3.29)$$

在哪里  $\mathcal{N}$  代表邻里，与位置  $(m,n)$  应用该操作。最小运算符将衰减图像中的孤立峰或脊，而最大运算符将填充孤立的槽或孔。可以看出，重复使用这种类型的过滤器将持续侵蚀所描述的孤立区域。中位运算符将对所有  $N$  区域内的样本，输出位置<sup>3</sup>的排名值  $(N + 1)/2$  因此：

$$y[m,n] = \text{med}_{\mathcal{N}(m,n)} (x(i,j)) \quad (3.30)$$

中位算子在去除异常值（如盐和胡椒噪声）方面比线性滤波器更好。它将保持直边，但会圆角并扭曲纹理特征。功能扁平化的程度将取决于所使用的窗口大小。

我们当然可以将这些操作结合起来，生成如下非线性差分滤波器：

$$y[m,n] = \min_{\mathcal{N}(m,n)} (x(i,j)) - \max_{\mathcal{N}(m,n)} (x(i,j)) \quad (3.31)$$

产生图像的非线性梯度。

## 形态滤波器

上述侵蚀和膨胀操作可以推广和组合，形成一个强大的运算符工具箱，称为形态滤波器。这些基于形状的运算符支持区域侵蚀和膨胀以及测地线重建，并在分割[6]、对比度增强和低比特率视频编码的特征优先排序[7]中找到了应用。

### 3.3.7。线性变换和DFT

线性变换，特别是离散余弦变换（DCT），构成了大多数帧内图像和剩余编码的基础，详见第5章。我们在这里介绍了这些，重点是离散傅里叶变换（DFT）。傅里叶变换在信号处理应用程序中广泛使用，以基于信号域及其相关频域的离散表示进行光谱分析。参考文献提供了傅里叶变换的绝佳介绍。[8]。

虽然离散时间傅里叶级数和离散时间傅里叶变换分别为分析周期和非周期信号提供了基础，但后者产生连续频域函数，因此不适合数字实现。无处不在的离散傅里叶变换克服了这一点，并允许在信号及其相关频域进行离散时间处理。

## 离散傅里叶变换

让我们首先考虑DFT的一维版本。如果  $x[n]$  是在范围内定义的有限持续时间信号  $0 \leq n \leq N_1 - 1$ ，周期信号  $\tilde{x}[n]$  可以构造等于  $x[n]$  在一个周期内。让  $N \geq N_1$  让  $\tilde{x}[n]$  定期  $N$ 。如果我们计算这个新序列的离散时间傅里叶级数（DTFS）：

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} \tilde{x}[n] e^{-jk(2\pi/N)n}$$

放开  $X(k) = N \cdot c_k$ ，这成为DFT的分析方程：

$$X(k) = \sum_{n=0}^{N-1} x[n] e^{-jk(2\pi/N)n}; \quad k = 0, 1, 2, \dots, N - 1 \quad (3.32)$$

反向DFT由以下方式给出：

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{jk(2\pi/N)n}; \quad n = 0, 1, 2, \dots, N-1 \quad (3.33)$$

此NDFT的项表示连续函数的样本 $X(\Omega)$ , 在 $2\pi$ 间隔。这些可以很容易地显示为等于 $z$ -在 $z$ -飞机。

通常我们使用矩阵向量形式的DFT:

$$\mathbf{X} = \mathbf{W}\mathbf{x} \quad (3.34)$$

在哪里

$$W_N = e^{-j(2\pi/N)}$$

因此, 我们以矩阵形式:

$$\begin{aligned} & \begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ \vdots \\ X(N-1) \end{bmatrix} \\ &= \begin{bmatrix} W^0 & W^0 & W^0 & \cdots & W^0 \\ W^0 & W^1 & W^2 & \cdots & W^{N-1} \\ W^0 & W^2 & W^4 & & W^{2(N-1)} \\ \vdots & & & & \vdots \\ W^0 & W^{N-1} & W^{2(N-1)} & \cdots & W^{(N-1)(N-1)} \end{bmatrix} \\ & \begin{bmatrix} x[0] \\ x[1] \\ x[2] \\ \vdots \\ x[N-1] \end{bmatrix} \end{aligned} \quad (3.35)$$

同样, 逆DFT由以下方式给出:

$$\mathbf{x} = \frac{1}{N} \mathbf{W}^* \mathbf{X} \quad (3.36)$$

二维DFT

与第3.3.3节中的过滤器类似, 变换操作可以是可分离的, 这是应用二维变换的传统方式, 主要是因为相关的复杂性节省。我们在第五章中更详细地介绍了这一点。

DFT和压缩

试图利用DFT的能力来分析信号的频域分量似乎是合乎逻辑的。然而, 这有一些重大问题, 这意味着在实践中没有做到这一点。这些是:

1. 虽然图像总是实值, 但DFT系数是复值的, 因此我们将信息量增加一倍。

2. 对于短序列或小的二维区域，如典型的压缩区域，周期的基本假设产生的不连续性可能会对性能产生重大影响，因为它们在频域产生环状或光谱泄漏。
3. 在光谱分析应用程序中，环通常通过在扩展前对输入数据应用非矩形窗口函数（如汉明窗口）来解决。这些减少了光谱环路，但涂抹了锐利的频域边缘，扭曲了底层信号的特性。再说一遍，这正是我们不想在[压缩系统中](#)想要的！

[第5章](#)对DFT和更合适的变换（如DTC）进行了比较。

## 3.4。量化

### 3.4.1。量化的基本理论

量化器将输入值（或向量量化器上下文中的值）映射到输出值。它可以取连续振幅[离散时间信号](#)并产生量化的数字信号（字长有限），也可以取以前数字化的样本并进一步量化以缩小其动态范围。前者在A到D转换器等采集设备中很常见，后者在[压缩系统中](#)，例如，转换输入信号的系数可能会映射到不同的范围，小值为零。

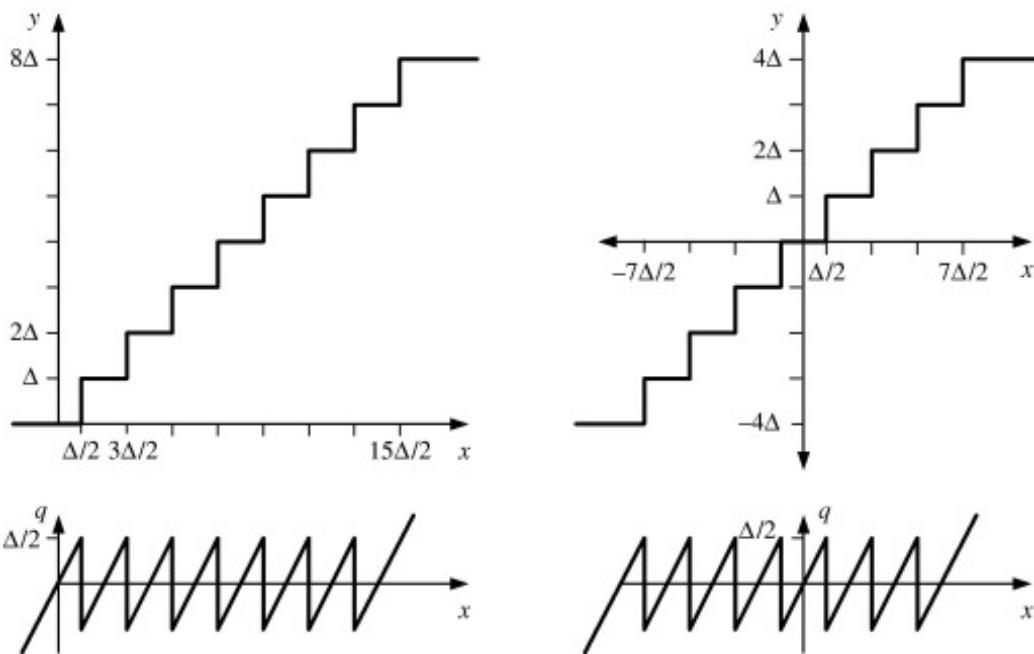
一旦对图像或视频信号进行采样，以产生一组具有给定字长的离散值，[压缩过程](#)通常涉及某种形式的信号分析（通常是转换）来装饰内容，然后进行[量化](#)，以尽可能保持[感知质量](#)的方式近似原始信号。忽略了数值[不精确](#)，量化是任何引入损耗的压缩过程的阶段。因此，了解量化对图像或视频信号感知质量的影响至关重要。

### 统一量化

均匀量化器采用恒定步长 $\Delta$ ，在重建水平之间。如果量化器的输入是 $x$ ，这将映射到重建值， $y_i$ ，在哪里 $y_i$ 表示由步长参数决定的范围中心。我们定义了一个[量化错误](#)， $q[n] = x[n] - y[n]$ ，它被建模为零平均值，在范围内均匀分布的信号 $[-\Delta/2, \Delta/2]$ 。该信号的能量或方差由以下因素给出：

$$\sigma_q^2 = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} q^2 dq = \frac{\Delta^2}{12} \quad (3.37)$$

[图3.12](#)显示了两个常见的均匀量化器配置文件。左边是用于正值信号的常规中踏板量化器，右侧是类似的剖面，但用于双面输入信号。决策级别 $x$ 与相应的重建值一起指示 $y$ 。



[下载：下载全尺寸图像](#)

图3.12。常见的均匀量化器特性。

对于常规中踏板量化器，输入上的量化操作 $x$ ，产生重建水平 $y_i$ ，就决策级别而言 $x_i$ ，定义为：

$$\dot{x} = Q(x) = y_i; \quad x_{i-1} < x \leq x_i \quad (3.38)$$

这通常通过四舍五入操作来实现，例如：

$$\dot{x} = \text{NINT}\left(\frac{x}{\Delta}\right) \quad (3.39)$$

其中NINT是四舍五入到最近整数的操作。

为了重建量化信号，它通常必须重新缩放到其原始范围。这种缩放操作的定义是：

$$\tilde{x} = \Delta \dot{x} \quad (3.40)$$

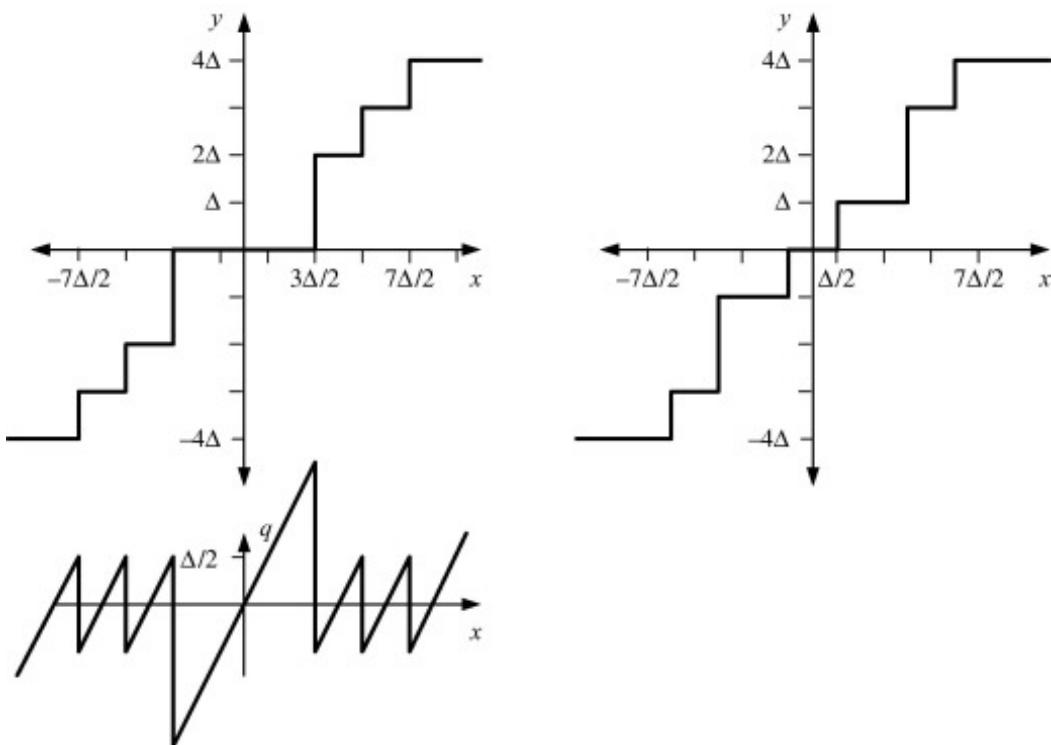
### 3.4.2。适应信号统计

在以下情况下，步长不均匀的量化器可能是有益的：

- 正在量化的信号没有显示统一的PDF格式。
- 已知低电平信号或其他信号范围包含过多的噪声。
- 人们希望将一系列低电平信号映射到零。
- 有必要将非线性振幅映射作为量化过程的一部分。

### 死区量化器

最常见的非均匀量化器形式是死区量化器。对于接近零的输入范围来说，这有更广泛的决策范围。在压缩期间，它的好处是确保噪声低电平信号不会被不必要的分配位。死区量化器的特性如图3.13（左）所示。



[下载：下载全尺寸图像](#)

图3.13。常见的非均匀量化器。左：中心死区。右：劳埃德·麦克斯量化器。

### 劳埃德麦克斯量化器

劳埃德·马克斯算法[10]是根据输入信号的流行PDF优化的非均匀量化器的设计方法。劳埃德·马克斯方法不是像统一PDF那样分配统一的步长，而是根据PDF曲线的等面积分区的平均值确定决策边界。劳埃德·马克斯量化器的特性如图3.13（右）所示。

### 3.4.3。HVS加权

在实践中，对于许多压缩应用来说，量化器步长是根据一些感知标准加权的，例如装饰变换操作产生的系数的**空间频率**。第五章对此进行了更详细的讨论。

### 3.4.4。矢量量化

**矢量量化** (VQ) 是一种将一组样本映射到单个**量化指数**的方法。这种方法已被应用于高度鲁棒的编解码器，例如第11章中讨论的基于金字塔向量量化 (PVQ) 的编解码器。我们在这里不会详细介绍，但有兴趣的读者请参阅参考文献。[9], [11]。

## 3.5。线性预测

### 3.5.1。基本前馈线性预测编码

线性预测的思想是利用随机数据序列中的统计关系，以便在编码前对其进行装饰。

用于线性预测的前馈分析滤波器如图3.14（上）所示，相应的合成滤波器如图3.14（下图）所示。暂时忽略量化块，分析滤波器的传递函数由以下方式给出：

$$\begin{aligned} E(z) &= X(z) - X_p(z) \\ &= X(z)(1 - P(z)) \end{aligned} \quad (3.41)$$

$$H(z) = \frac{E(z)}{X(z)} = 1 - P(z) \quad (3.42)$$

其中使用的符号如图所示。假设图3.14中的预测块是一个FIR过滤器，那么在样本域中，我们有：

$$e[n] = x[n] - \sum_{i=1}^N a_i x[n-i] \quad (3.43)$$

如果我们现在考虑解码器或合成过滤器：

$$Y(z) = E(z) + Y(z)P(z) \quad (3.44)$$

或者：

$$Y(z) = \frac{E(z)}{1-P(z)} = X(z) \quad (3.45)$$

#### 例 3.5 线性预测编码

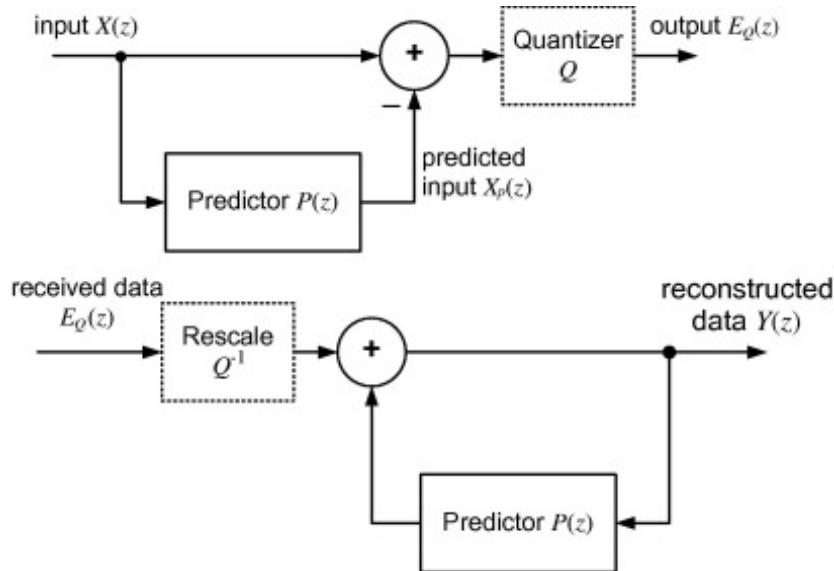
在何时考虑图3.14中的情况 $P(z) = z^{-1}$ 预测编码器的输入是 $x[n] = \{1, 2, 3, 2, 5, 4, 2, 4, 5, 6\}$ 。假设没有进行量化，请确认解码器输出与输入序列相同。

#### 解决方案

我们有 $e[n] = x[n] - x[n-1]$ 和 $y[n] = e[n] + y[n-1]$ 。因此，假设初始条件为零，并为每个输入示例工作：

$x$ [n]	1	2	3	2	5	4	2	5	6
$e$ [n]	1	1	1	-1	3	-1	-2	3	1
$y$ [n]	1	2	3	2	5	4	2	5	6

因此 $y[n] = x[n]$ 。



[下载：下载全尺寸图像](#)

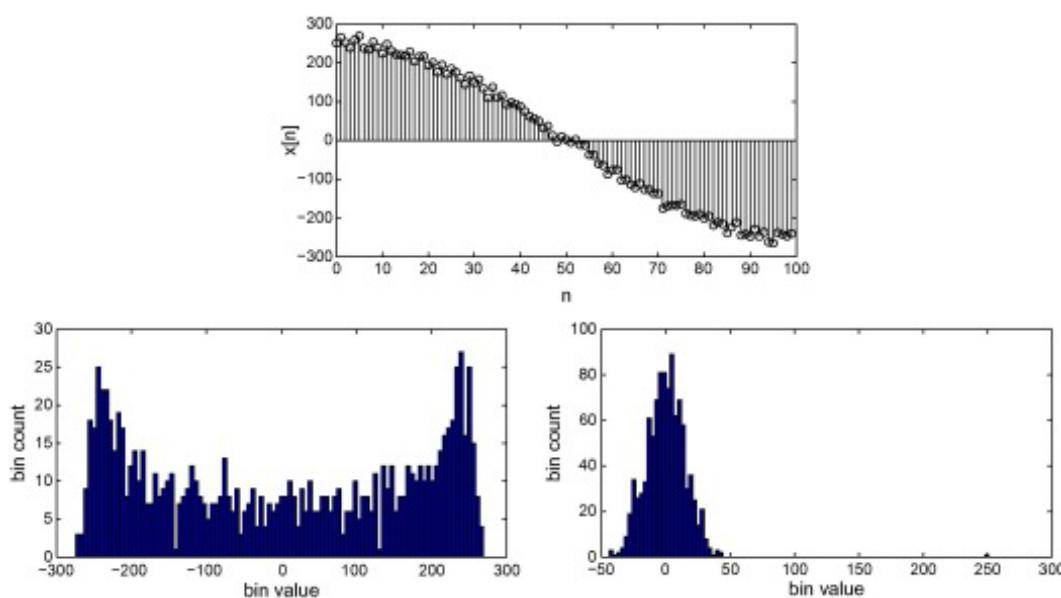
图3.14。前馈线性预测。顶部：编码器。底部：解码器。

### 预测器动态范围

预测滤波作为编码基础的好处是，虽然传输误差信号的动态范围可以比原始信号的两倍，但误差信号的方差比原始信号显著降低。图3.15证明了这一点，其中带有加性高斯噪声的正弦信号 ( $\sigma_v^2 = 0.05$ ):

$$x[n] = \cos(0.01\pi n) + v[n] \quad (3.46)$$

使用预测器过滤  $P(z) = z^{-1}$ 。



[下载：下载全尺寸图像](#)

图3.15。预测信号动态范围。顶部：输入信号。左下角：分发1000个输入信号样本。右下角：分布1000个预测残差样本。

---

## 线性预测编码与量化

在引入量化的过程中，前馈架构就出现了问题。在这种情况下，我们有：

$$E_Q(z) = Q\{X(z) - X_p(z)\} \quad (3.47)$$

和

$$Y(z) = Q^{-1}\{E_Q(z)\} + Y(z)P(z) \quad (3.48)$$

或者：

$$Y(z) = \frac{Q^{-1}\{E_Q(z)\}}{1-P(z)} \neq X(z) \quad (3.49)$$

正如人们所料，量化过程在重建的输出中引入了错误。这是意料之中的，因为手术是有损的。然而，情况更糟，因为编码器和解码器之间有漂移的可能性。这在[示例3.6](#)中说明。

### 例 3.6 线性预测编码与量化

考虑[图3.14](#)中的案例，其中与以前一样， $P(z) = z^{-1}$ 预测编码器的输入是 $x[n] = \{1, 2, 3, 2, 5, 4, 2, 4, 5, 6\}$ 。假设进行量化，其中量化和缩放操作的形式是：

$$e_Q[n] = \text{rnd}\left(\frac{e[n]}{2}\right); \quad e_R[n] = 2e_Q[n]$$

四舍五入0.5到0，描述解码器相对于输入信号的输出。

### 解决方案

假设初始条件为0，并遵循与[示例3.5](#)相同的方法，但这次包括量化和缩放操作：

$x[n]$	1	2	3	2	5	4	2	5	6
$e[n]$	1	1	1	-1	3	-1	-2	3	1
$e_Q[n]$	0	0	0	0	1	0	-1	1	0
$e_R[n]$	0	0	0	0	2	0	-2	2	0
$y[n]$	0	0	0	0	2	2	0	2	2

因此 $y[n]$ 与 $x[n]$ 。此外，由于解码器的四舍五入误差积累，出现了显著的漂移。

### 3.5.2。反馈回路中用预测器进行线性预测

为了避免漂移问题，我们必须在编码器中复制解码操作的效果。这确保了在传输过程中没有数据损坏的情况下，编码器和解码器进程的数据完全相同，并一致运行。这个修改后的架构如图3.16所示。我们编码输出（不进行量化）的表达式现在由以下方式给出：

$$\begin{aligned} E(z) &= X(z) - X_p(z) \\ &= X(z) - \frac{E(z)P(z)}{1-P(z)} \end{aligned} \quad (3.50)$$

因此：

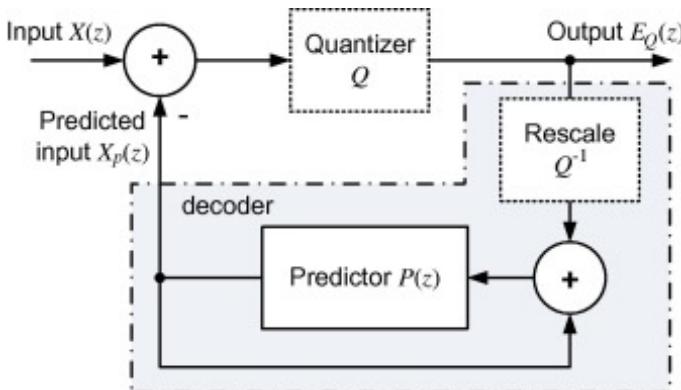
$$E(z) = X(z)(1 - P(z))$$

和以前一样，在解码器：

$$Y(z) = E(z) + Y(z)P(z) \quad (3.51)$$

或者：

$$Y(z) = \frac{E(z)}{1-P(z)} = X(z) \quad (3.52)$$



[下载：下载全尺寸图像](#)

图3.16。基于反馈的线性预测。

如果我们现在将量化纳入流程，反馈和前馈架构之间的区别将变得很明显。为了便于分析，我们将量化过程建模为加性量化噪声， $V(z)$ 带有差异 $\sigma_N^2$ 。然后，如图3.17所示：

$$\begin{aligned} E(z) &= X(z) - X_p(z) \\ &= X(z) - \frac{[E(z)+V(z)]P(z)}{1-P(z)} \end{aligned} \quad (3.53)$$

因此：

$$E(z) = X(z)(1 - P(z)) - V(z)P(z) \quad (3.54)$$

在解码器：

$$Y(z) = E(z) + Y(z)P(z) + V(z) \quad (3.55)$$

或者：

$$Y(z) = \frac{E(z)+V(z)}{1-P(z)} \quad (3.56)$$

用方程 (3.54) 代替  $E(z)$  在方程 (3.56) 中，直观地看到：

$$Y(z) = X(z) + V(z) \quad (3.57)$$

因此，我们不再在编码器和解码器之间有任何漂移，我们只是因为量化过程的加性噪声而出现不匹配。这是一个众所周知但非常重要的结果，因为这种预测架构在视频压缩中被广泛使用。

### 示例3.7 基于反馈的定量线性预测编码

考虑图3.16中的情况，其中与以前一样， $P(z) = z^{-1}$  预测编码器的输入是  $x[n] = \{1, 2, 3, 2, 5, 4, 2, 4, 5, 6\}$ 。假设进行量化，其中量化和缩放操作的形式是：

$$e_Q[n] = \text{rnd}\left(\frac{e[n]}{2}\right); \quad e_R[n] = 2e_Q[n]$$

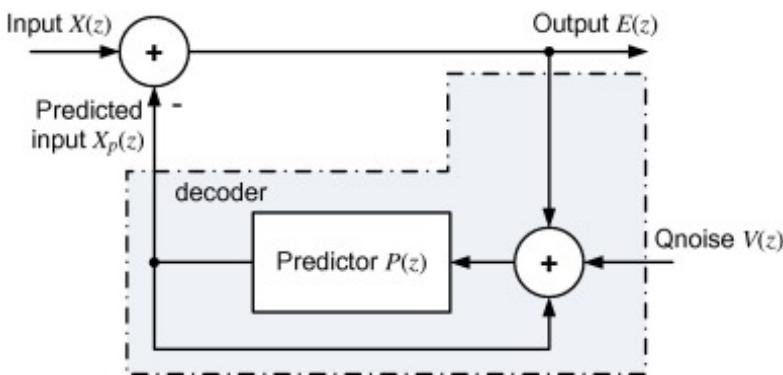
四舍五入0.5到0，描述解码器相对于输入信号的输出。

### 解决方案

假设初始条件为0，并遵循与示例3.6相同的方法，但这次基于图3.16中的架构：

$x[n]$	1	2	3	2	5	4	2	5	6
$e[n]$	1	2	1	0	3	-1	-2	3	2
$e_Q[n]$	0	1	0	0	1	1	-1	1	1
$e_R[n]$	0	2	0	0	2	2	-2	2	2
$y[n]$	0	2	2	2	4	4	2	4	6

不出所料， $y[n]$  与  $x[n]$  因为有损的量化过程。然而，与原始信号相比，预测信号的方差显著减少，编码器和解码器之间也没有漂移。



[下载：下载全尺寸图像](#)

图3.17。基于反馈的线性预测器，具有定量噪声建模。

### 3.5.3。维纳霍普夫方程和维纳滤波器

如果我们有一个脉冲响应 $\mathbf{h}$ 的线性FIR预测滤波器，我们希望滤波器系数被分配到值，迫使预测后的错误信号尽可能小。如果我们用二阶统计（即基于平均平方误差（MSE））来执行此优化，那么结果被称为维纳滤波器——一个最佳的非递归估计器。

如果我们有一个想要的序列（训练序列） $\mathbf{d}$ ，则：

$$e[n] = d[n] - y[n] \quad (3.58)$$

我们可以将成本函数定义为：

$$J = E\{|e[n]|^2\} \quad (3.59)$$

梯度在哪里，当梯度 $\nabla(J) = 0$ 关于所有滤波系数，那么滤波器被称为MSE意义上的最佳滤波器。可以证明，一个必要和充分的条件 $J$ 达到最小值是估计误差的相应值与当时进入估计过程的每个样本正交 $n$ 。这被称为正交原则。基于这一观察，可以直观地证明，最佳滤波系数可以从滤波器输入序列的自相关中计算出来 $r[k]$ 以及滤波器输入和所需响应的交叉相关性 $p[k]$ 滞后 $k$ 样品。这被称为维纳·霍普夫方程：

$$\sum_{i=0}^{\infty} h_o[i]r[i-k] = p[k] \quad (3.60)$$

或以矩阵向量形式：

$$\mathbf{h}_o = \mathbf{R}^{-1} \mathbf{p} \quad (3.61)$$

假设平稳，也可以证明误差表面恰好是滤波系数的二阶函数。因此，它将拥有与最佳滤波器权重对应的唯一最小点。有关更多详细信息，请参阅参考文献。[\[12\]](#)。

## 3.6。信息和熵

生成信息定量测量的概念由香农于1948年提出[13]。他使用熵（随机变量中不确定性的度量）来量化消息中包含的信息的价值。

了解信息理论的基础在图像和视频压缩中很重要，因为它被广泛用于转换和量化后的无损符号编码。通常，量化变换系数被扫描并编码运行长度。表示零运行和非零系数的符号使用可变长度的码字进行编码，以反映其信息内容的方式进行编码。这样就可以实现无损的比特率降低。

在研究符号统计之前，让我们先回顾一下信息理论的基础。如需更深入的报道，请参阅参考文献。[14]。

### 3.6.1。自我信息

信息用于减少事件或信号的不确定性。Shannon[13]定义了事件的自我信息*A*例如，它可以包括一组实验结果，例如：

$$i(A) = \log_2 \left( \frac{1}{P(A)} \right) = -\log_2(P(A)) \quad (3.62)$$

对数的使用是直观的，因为 $\log(1) = 0$ 和 $-\log(x)$ 随着 $x$ 从1减少到0。因此，如果事件的概率较低，则与之关联的信息较高。例如，如果 $P(A) = 0.5$ 然后 $-\log_2(0.5) = 1$ 如果 $P(A) = 0.25$ 然后 $-\log_2(0.25) = 2$ 。可以使用其他对数基数，但基数2是最常见的，因为它方便地以位为单位提供信息的度量。

### 独立事件

从两个独立事件获得的信息是每个事件信息的总和。这可以看作如下：

$$i(AB) = \log_2 \left( \frac{1}{P(AB)} \right)$$

如果*A*和*B*独立，然后：

$$P(AB) = P(A)P(B) \quad (3.63)$$

因此：

$$i(AB) = \log_2 \left( \frac{1}{P(A)P(B)} \right) = \log_2 \left( \frac{1}{P(A)} \right) + \log_2 \left( \frac{1}{P(B)} \right)$$

或者

$$i(AB) = i(A) + i(B) \quad (3.64)$$

### 3.6.2。熵

如果我们定义一个源，该源生成一系列随机和独立的事件，其结果来自字母表 $S = \{s_1, s_2, s_3, \dots, s_N\}$ ，则与事件相关的平均自我信息是：

$$H = \sum_{i=1}^N P(s_i)i(s_i) = -\sum_{i=1}^N P(s_i)\log_2(P(s_i))$$

数量*H*，平均自我信息，被称为源的熵，这告诉我们每个符号需要多少位来编码该源。Shannon

证明，没有编码方案可以以低于其熵的速度无损编码源。

## 熵和一阶熵

源的熵由以下方式给出：

$$H(S) = \lim_{N \rightarrow \infty} \frac{1}{N} G_N \quad (3.65)$$

在哪里

$$G_N = - \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_M=1}^N P(d_1 = s_{i_1}, d_2 = s_{i_2}, \dots, d_M = s_{i_M}) \\ \times \log P(d_1 = s_{i_1}, d_2 = s_{i_2}, \dots, d_M = s_{i_M})$$

和  $D = \{d_1, d_2, d_3, \dots, d_M\}$  和  $M$  是序列的长度。但是，如果每个符号都是独立且分布相同的 (*iid*)，则可以显示：

$$G_N = -M \sum_{i_1=1}^N P(d_1 = s_{i_1}) \log(P(d_1 = s_{i_1})) \quad (3.66)$$

然后熵是：

$$H(S) = - \sum_{i=1}^N P(s_i) \log_2(P(s_i)) \quad (3.67)$$

在大多数现实情况下，方程 (3.65)，(3.67) 给出的结果并不相同。因此，我们通过将方程 (3.67) 称为一阶熵来区分它们。

符号概率与该符号的自我信息之间的关系如图3.18所示。该图还说明了符号的概率与符号对源整体熵的贡献之间的关系。图3.18表明，比例尺两端概率的符号对源熵贡献很小，而范围中间的符号贡献值约为0.5。

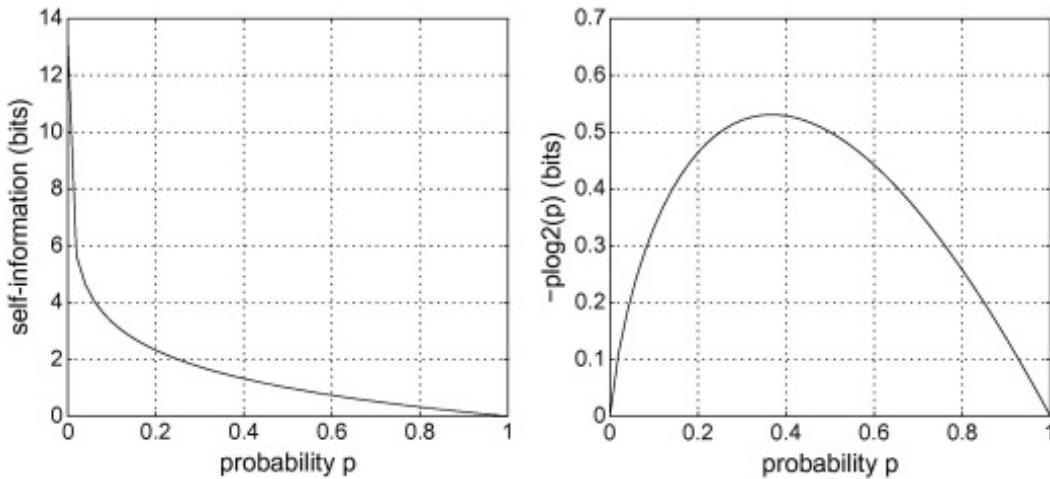
## 示例3.8熵

字母表  $S = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8\}$ 。如果这些符号在传输过程中出现的相应概率是  $P(S) = \{0.06, 0.23, 0.3, 0.15, 0.08, 0.06, 0.06, 0.06\}$ ，计算源的第一阶熵。

## 解决方案

一阶熵由以下方式给出：

$$H = - \sum_S P(s_i) \log_2(P(s_i)) \\ = -(0.06 \times \log_2 0.06 + 0.23 \times \log_2 0.23 + \cdots + 0.06 \times \log_2 0.06) \\ = 2.6849 \text{ bits / symbol}$$



[下载：下载全尺寸图像](#)

图3.18。自我信息和概率。左：单个事件的自我信息与概率的图。右：按概率加权的事件的自我信息图。

### 3.6.3. 符号和统计

在实践中，由于方程 (3.65) 无法计算，因此几乎总是使用一阶熵。因此，我们必须估计源的熵，这在很大程度上取决于源符号的统计。这一点在**示例3.9**中得到了强调。

#### 例3.9 源统计和熵

考虑符号的顺序： $D = \{1, 2, 3, 2, 1, 2, 3, 4, 5, 6, 5, 6, 7, 8, 9, 10\}$ 。计算此序列的第一阶熵。

#### 解决方案

这似乎是一个直截了当的问题。我们没有得到关于源统计的信息，因此让我们根据相对发生的频率来估计每个符号的概率。通过这种方式，我们有：

$$\begin{aligned} P(4) &= P(7) = P(8) = P(9) = P(10) = \frac{1}{16} \\ P(1) &= P(3) = P(5) = P(6) = \frac{2}{16} \\ P(2) &= \frac{3}{16} \end{aligned}$$

因此，一阶熵是

$$H = - \sum_S P(s_i) \log_2(P(s_i)) = 3.2028 \text{ bits}$$

因此，这说明，我们平均每个符号需要3.2位以上。

然而，如果我们现在仔细观察序列，我们可以观察到符号是相互关联的。那么，如果我们用一个简单的一阶线性预测器处理这个序列会发生什么？我们的剩余序列，

$$E = \{1, 1, 1, -1, -1, 1, 1, 1, 1, -1, 1, 1, 1, 1, 1\}.$$

因此，现在我们有了一套新的概率估计：

$$P(1) = \frac{13}{16}$$

$$P(-1) = \frac{3}{16}$$

在这种情况下，我们重新计算我们的熵为0.6962位/符号。那么，哪个是正确的？事实上，无论是对还是错，只是在后一种情况下，我们更好地了解了源的结构，并提供了对其熵的更好估计。

例如，我们可以通过计算其**量化级别的**直方图，并将每个**量化仓**的内容除以像素总数来估计图像或图像区域中出现像素值的概率。或者，我们可以使用一个或多个类似训练图像的直方图预计算它们。后者更常见，尽管我们在实践中很少使用实际像素值作为输入符号。我们更经常地使用一些基于量化空间或转换域数据的预测或运行长度编码的更有效的方法。稍后在第七章中了解更多。

### 3.7。摘要

本章回顾了信号处理和信息理论的一些初步主题，这些主题有助于将本书的其余部分置于上下文中。我们没有用一系列证明和定理给读者带来负担，而是专注于使用示例来帮助理解一些重要主题。我们检查了一些基本的统计措施，并用这些措施来描述图片冗余。过滤器和变换是大多数**压缩方法**的基础，因此引入了它们，并将在第5章和第6, 章第5章第, 6 and 章中进一步介绍 and 。我们还在同一章中重新审视了**量化**领域。

我们在这里花了一些时间在预测结构上，因为这些结构经常用于压缩，并且可以在空间和时间处理中为非预测方法提供可观的编码收益。预测内将在第9章和第12 and 章中涵盖，and 第8章和第9 and 章将涵盖运动预测，第, 8章和第9 and 章将涵盖。

最后，我们回顾了信息理论的基础，解释了其性能如何在很大程度上依赖于假设的源模型和上下文。这构成了大多数**压缩标准**中使用的符号编码的基础。我们在第七章中特别关注熵编码方法，并在第12章标准背景下跟进。

Recommended articles

Citing articles (0)

## 参考文献

[1] 美国。米特拉

**数字信号处理：一种基于计算机的方法**

(第四版)，麦格劳·希尔 (2011)

[谷歌学术](#)

[2] J. 伍兹

**多维信号、图像和视频处理和编码**

(第二版)，学术出版社 (2012年)

[谷歌学术](#)

[3] C.香农

**在噪音面前进行沟通**

无线电工程师学会学报, 37 (1) (1949年), p.10-21

- [4] E. 迪布瓦  
**时变图像的采样和重建及其在视频系统中的应用**  
IEEE会议记录, 73 (1985), p.502-522  
[在Scopus中查看记录](#) 谷歌学术
- [5] Y. 王, J. Ostermann, Y.-Q. 张  
**视频处理和通信**  
普伦蒂斯大厅 (2001)  
[谷歌学术](#)
- [6] R. 奥卡拉汉, D. 公牛  
**形态-光谱非监督图像分割联合**  
IEEE图像处理交易, 14 (1) (2005年), 第49-62页  
[CrossRef](#) [在Scopus中查看记录](#) 谷歌学术
- [7] P. 切雷皮埃斯基, D. 公牛  
**基于形态分割的增强型帧间编码**  
IEE Proceedings — Vision, Image and Signal Processing, 144 (4) (1997), pp.220-226  
[CrossRef](#) 谷歌学术
- [8] A. 奥本海姆, A. 威尔斯基, H. 纳瓦布  
**信号和系统**  
(第二版), 普伦蒂斯大厅 (1999年)  
A. 内拉瓦利, B. Haskell, 数字图片: 表示、压缩和标准, 第二版, 全体会议出版社, 1995年。  
[谷歌学术](#)
- [9] J.-R. 欧姆  
**多媒体通信技术**  
斯普林格 (2004)  
[谷歌学术](#)
- [10] 美国。劳埃德, PCM中的最小二乘量化 (1957年), 在IEEE跨通信30 (1982) 中转载[129-137](#)。  
[谷歌学术](#)
- [11] Y. 林德, A. 布佐, R. 灰色  
**一种矢量量化器设计的算法**  
IEEE通信交易, 28 (1980), 第84-95页  
[在Scopus中查看记录](#) 谷歌学术
- [12] M. 海耶斯  
[谷歌学术](#)

## 一种通信的数学理论

贝尔系统技术期刊, 27 (3) (1948年), 第379-423

[CrossRef](#) [在Scopus中查看记录](#) [谷歌学术](#)

[14] T. 封面, J. 托马斯

信息理论的要素

威利 (2006)

[谷歌学术](#)

<sup>1</sup> 注意: 我们删除了双下标, 因为它是多余的, 如果没有混淆结果, 我们将使用小写 $r$ 表示自相关。

<sup>2</sup> 移位不变性的线性系统通常被称为线性时不变性 (LTI)。虽然自变量并不总是时间, 但为了方便起见, 我们将在这里使用该术语。

<sup>3</sup> 注意: 中位过滤器通常由奇数次的水龙头定义。在样本数量为偶数的情况下, 通常取中间两项的平均数。

版权所有©2014爱思唯尔有限公司。保留一切权利。



[关于ScienceDirect](#)

[远程访问](#)

[购物车](#)

[广告](#)

[联系和支持](#)

[条款和条件](#)

[隐私政策](#)

 RELX™

我们使用cookie来帮助提供和增强我们的服务, 并定制内容和广告。继续即表示您同意[使用cookie](#)。

版权所有©2021爱思唯尔B.V.或其许可方或贡献者。ScienceDirect®是爱思唯尔B.V.的注册商标。

ScienceDirect®是爱思唯尔B.V.的注册商标。