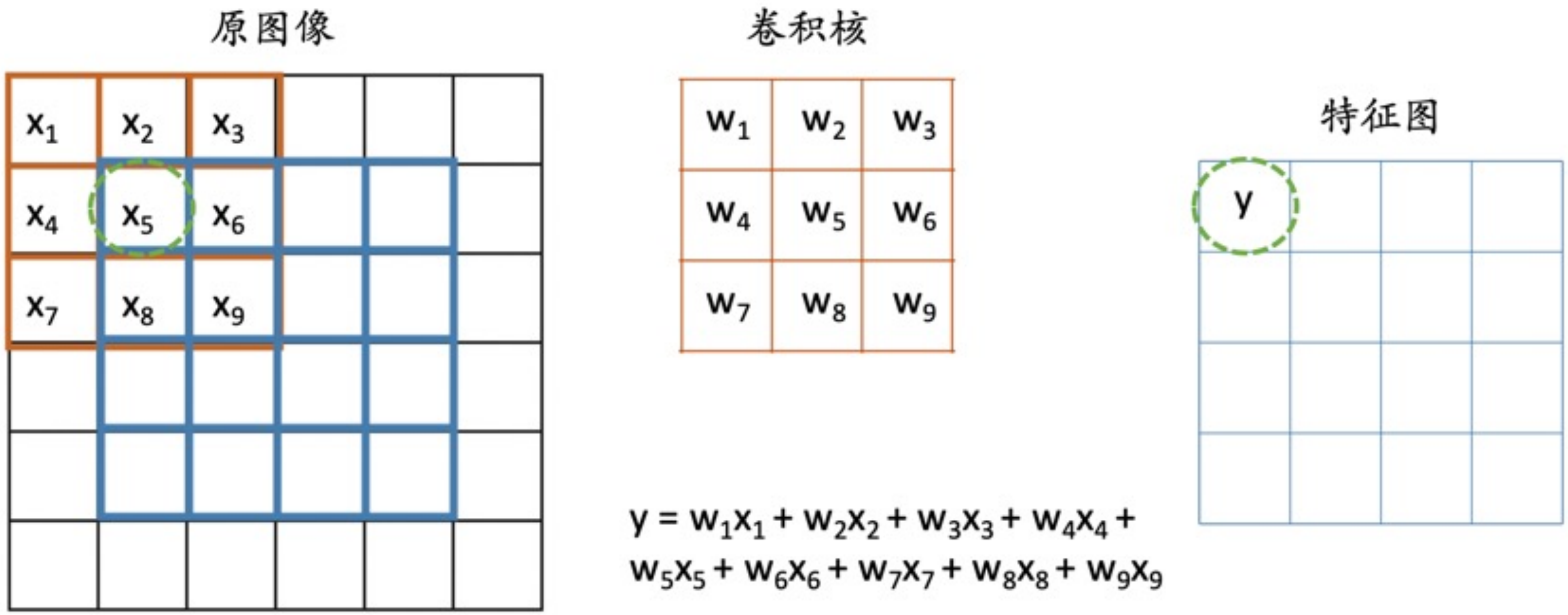
图像增强（2）

叶山 中国地质大学（北京）

yes@cugb.edu.cn

卷积和滤波

卷积操作 Convolution



$$y = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + w_6x_6 + w_7x_7 + w_8x_8 + w_9x_9$$

使用卷积核（convolution kernel,也称为模板、算子或掩膜）对原图像中的每个像素进行重新计算，形成“特征图”。具体来说，卷积核是一个矩阵，通常是一个较小的四方形网格结构，边长通常为基数个像素（如3x3的矩阵）。在进行卷积计算时，卷积核的中心会对准原图像的某个像素点（y）。然后，卷积核中的每个元素会与其覆盖的图像像素值一一相乘。这些乘积求和的结果，就是特征图中该位置(y)的新像素值。

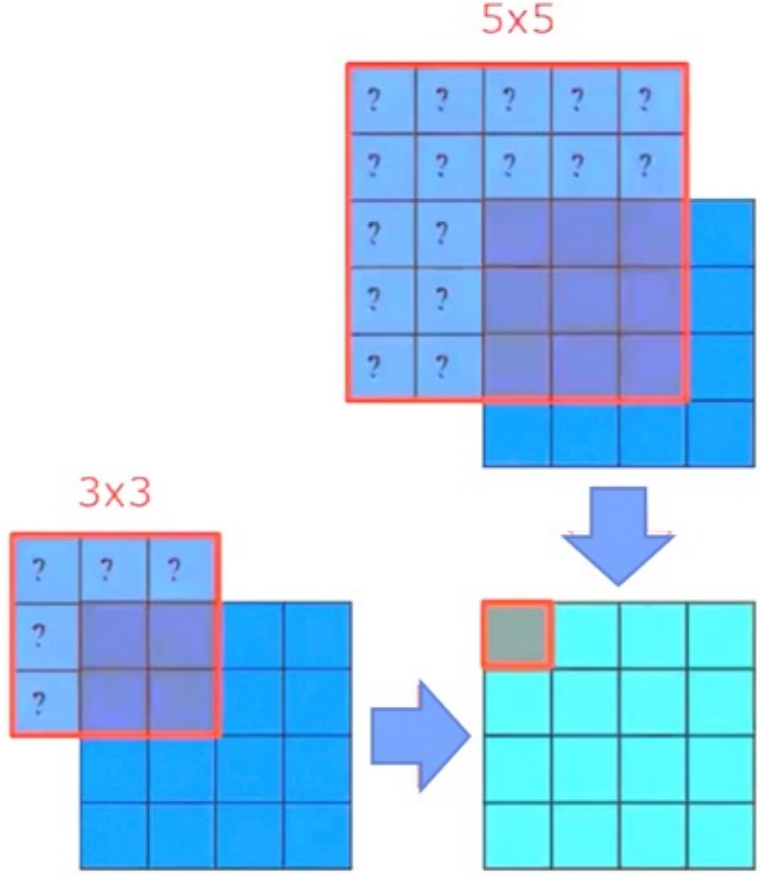
边缘填充 Padding

边缘填充（padding）

- 如果需要获得和原图同尺寸的输出图像，需要进行边缘填充
- 卷积核越大，补充越多

边缘填充类型

- 补零，也称零填充（zero padding）
- 边界复制（replication）
- 镜像（reflection）
- 块复制（wrap-around）



滤波 Filtering

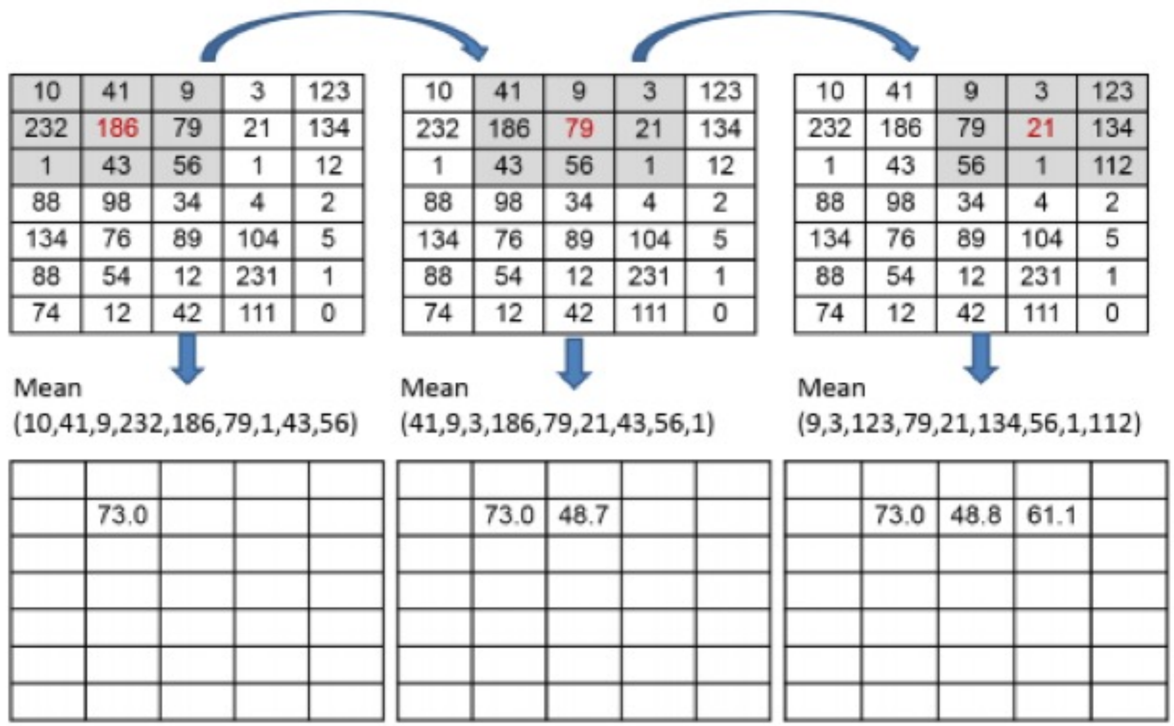
卷积操作在深度学习中多用于提取特征。除此之外，我们也可以用类似于卷积操作的手段，在预处理阶段对图像做降噪处理。

- 常见的滤波方法分均值滤波和中值滤波。

Salt and Pepper Noise
像胡椒和盐粒的噪声

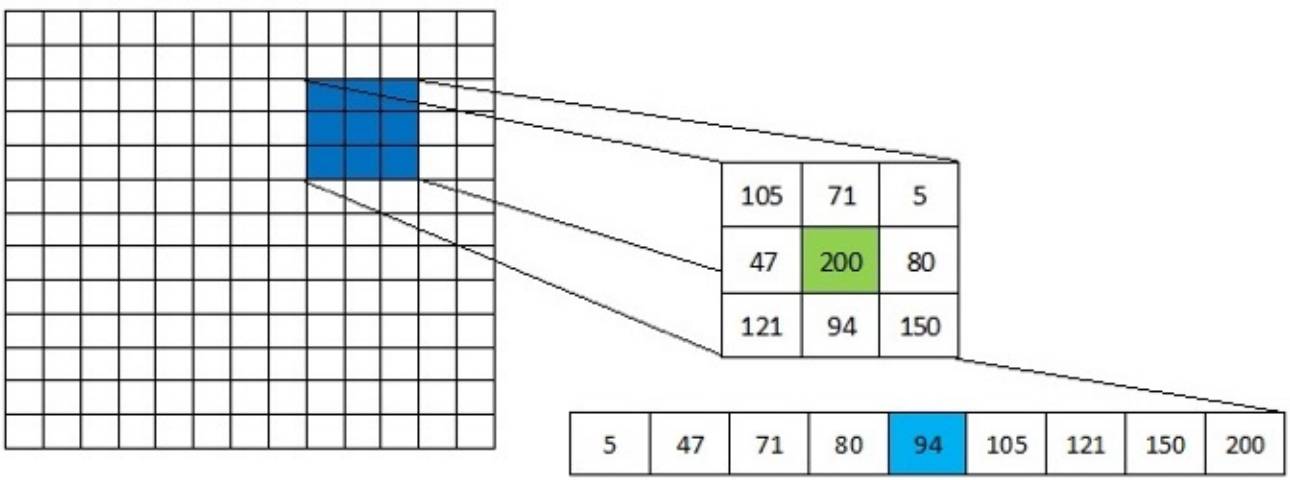
均值滤波 Mean Filtering

均值滤波是典型的线性滤波算法：在图像上对目标像素给一个模板（滤波器），该模板覆盖了目标像素及其周围的邻近像素。在滤波后的图像里，模板覆盖的全体像素的平均值，将被用来代替目标像素的原像素值。



中值滤波 Median Filtering

中值滤波是一种非线性滤波方法：将图像中某一点的像素值，用该点及其邻域中各点像素值的中位数代替，从而消除孤立的噪点。这种滤波方法对于滤除图像的“椒盐噪声”非常有效，且能改善图像的模糊程度。



Pop Quiz

1. RGB色彩空间中，大约有多少种颜色?

A. 1000 | B. 103万 | C. 1680万 | D. 3000万

2. 以下哪个光圈的开口最大?

A. f/4 | B. f/8 | C. f/11 | D. f/16

3. 以下哪个格式的图像为矢量文件?

A. TIFF | B. JPEG | C. SVG | D. PNG

4. 镜头分辨率的单位是?

A. lp/cm | B. lp/mm | C. dpi | D. ppi

5. 以下哪一项，卤素灯好于LED光源?

A. 寿命 | B. 响应速度 | C. 设计自由度 | D. 亮度

Gaussian滤波

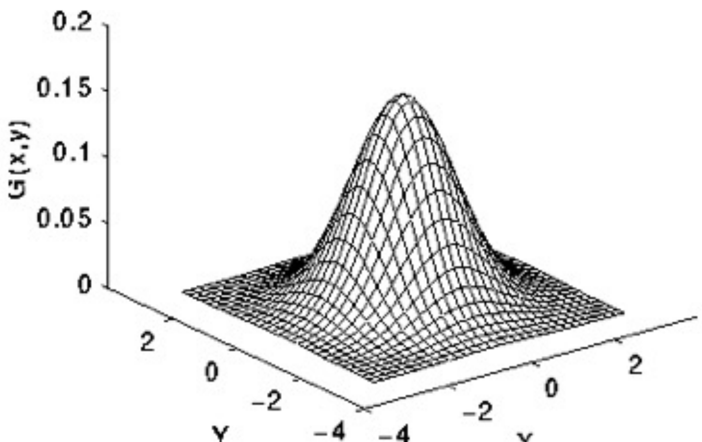
高斯滤波：用一个模板（卷积核/掩模/滤波器）扫描图像中的每一个像素，用模板框定邻域内像素的加权平均灰度值，去替代模板中心像素点的值，权值由模板决定。

高斯平滑滤波器对于抑制服从正态分布的噪声非常有效。它模拟人眼，关注中心区域（滤波器中间权重大，边缘权重小）。高斯滤波器的频带宽度由σ（标准差）决定，σ越大，整体平滑程度越好。通过调整σ，可以平衡对图像的噪声的抑制和图像的模糊化度。

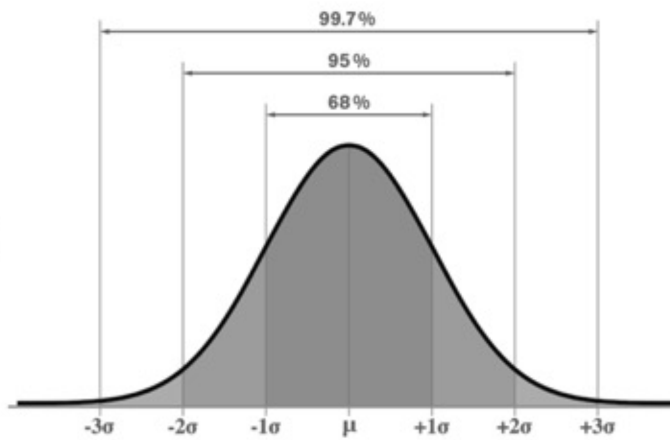
此外，高斯滤波也可以用于图像增强中的边缘检测，因为它强调了图像的长距离信息，这有利于边缘检测算子的计算。

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

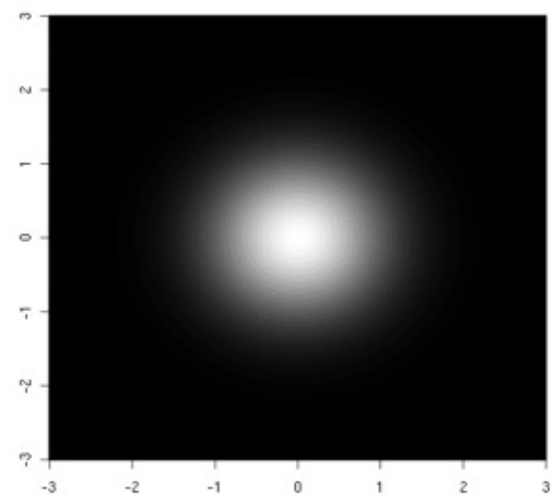
模板（滤波器）每个(x,y)位置的权重



高斯分布（正态分布）



增加维度



Gaussian滤波

1/16

1	2	1
2	4	2
1	2	1

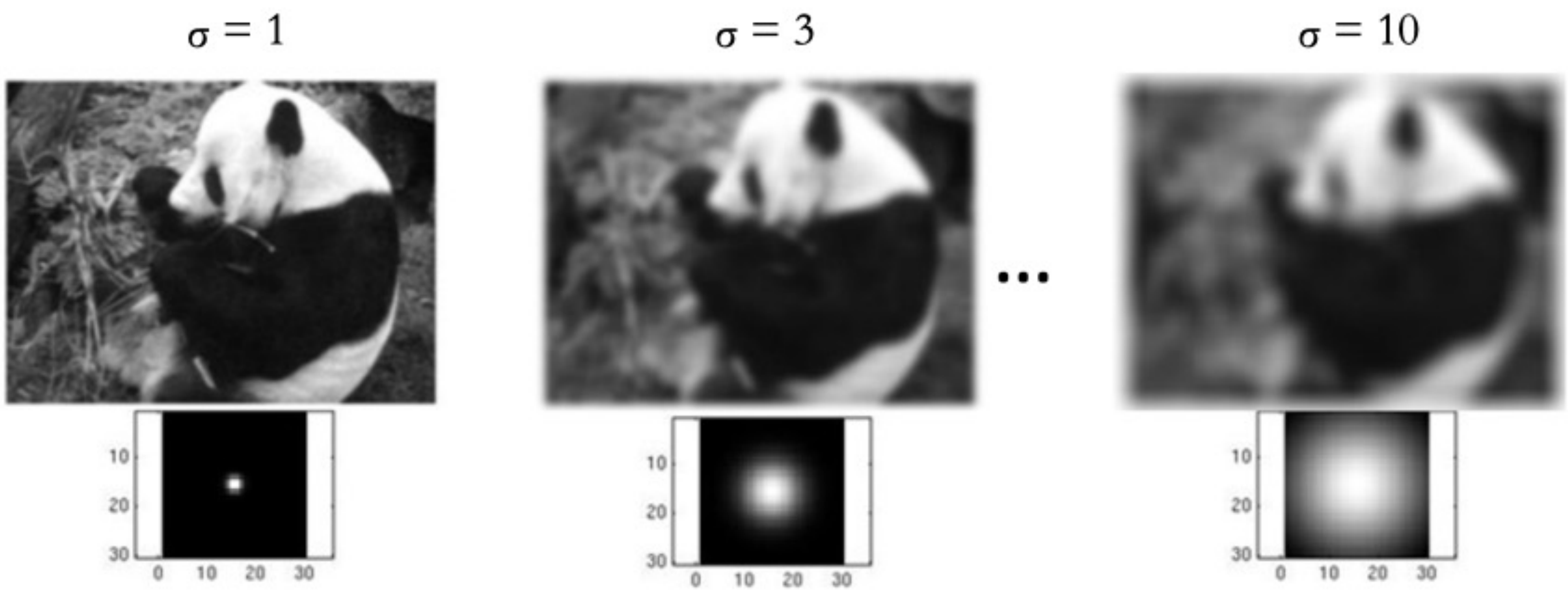
1/273

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1

1/1003

0	0	1	2	1	0	0
0	3	13	22	13	3	0
1	13	59	97	59	13	1
2	22	97	159	97	22	2
1	13	59	97	59	13	1
0	3	13	22	13	3	0
0	0	1	2	1	0	0

Gaussian滤波



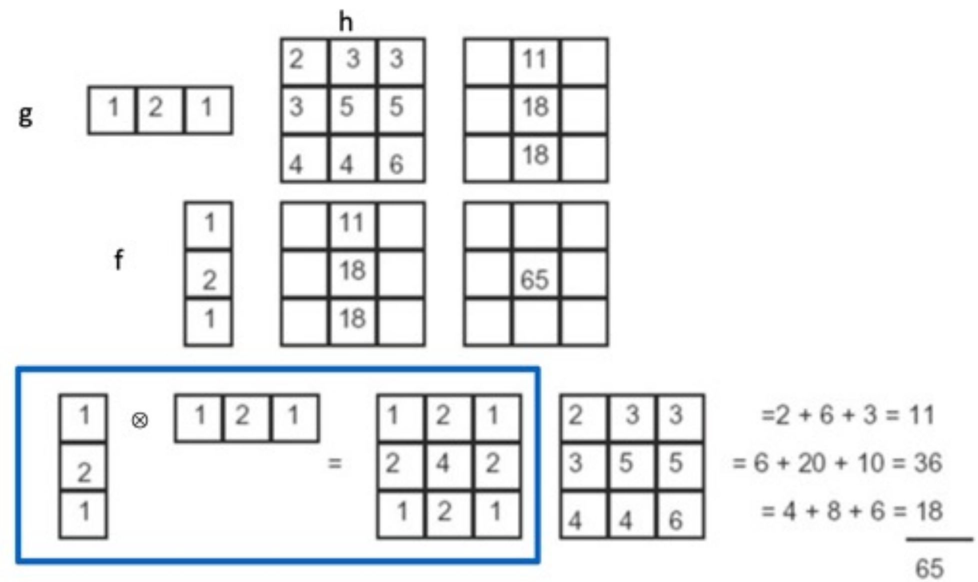
Gaussian滤波

分解特性：由于高斯卷积核的对称性，可将2D卷积核拆分为两个1D卷积核

- 列卷积
- 行卷积

目的：降低计算量

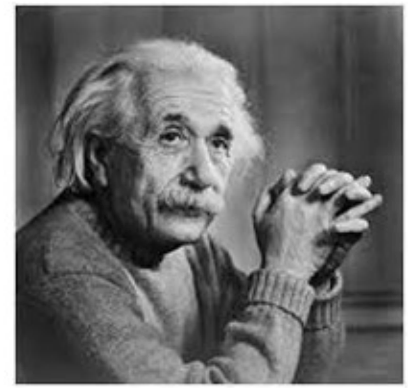
- 1个2D卷积：K^2次计算
- 2个1D卷积：2K次计算



Prewitt边缘检测器

- 水平梯度/垂直边缘

$$\text{均值平滑}\downarrow\overset{\text{水平梯度}\rightarrow}{\begin{bmatrix}-1&0&1\\-1&0&1\\-1&0&1\end{bmatrix}} = \underbrace{\begin{bmatrix}-1&0&1\end{bmatrix}}_{\text{水平梯度}} \otimes \begin{bmatrix}1\\1\\1\end{bmatrix}\text{均值平滑}$$



- 垂直梯度/水平边缘

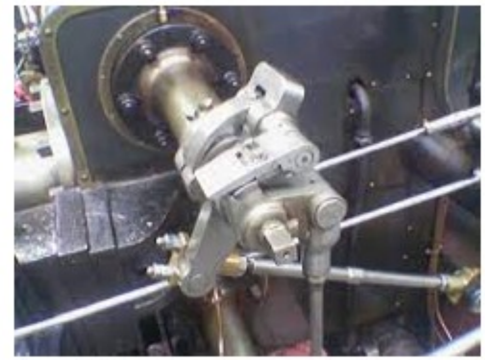
$$\text{垂直梯度}\downarrow\overset{\text{均值平滑}\rightarrow}{\begin{bmatrix}-1&-1&-1\\0&0&0\\1&1&1\end{bmatrix}} = \underbrace{\begin{bmatrix}1&1&1\end{bmatrix}}_{\text{均值平滑}} \otimes \begin{bmatrix}-1\\0\\1\end{bmatrix}\text{垂直梯度}$$

Sobel边缘检测器

- 水平梯度/垂直边缘

水平梯度 →

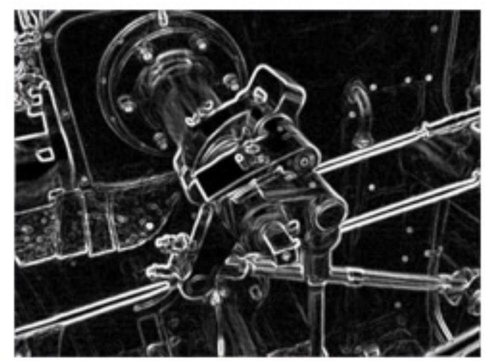
$$\begin{array}{c} \text{高斯平滑} \\ \downarrow \end{array} \begin{bmatrix}-1&0&1\\-2&0&2\\-1&0&1\end{bmatrix} = \underbrace{\begin{bmatrix}-1&0&1\end{bmatrix}}_{\text{水平梯度}} \otimes \begin{bmatrix}1\\2\\1\end{bmatrix} \text{高斯平滑}$$



- 垂直梯度/水平边缘

高斯平滑 →

$$\begin{array}{c} \text{垂直梯度} \\ \downarrow \end{array} \begin{bmatrix}-1&-2&-1\\0&0&0\\1&2&1\end{bmatrix} = \underbrace{\begin{bmatrix}1&2&1\end{bmatrix}}_{\text{高斯平滑}} \otimes \begin{bmatrix}-1\\0\\1\end{bmatrix} \text{垂直梯度}$$



Laplacian边缘检测器

拉普拉斯算子（Laplacian Operator）是N维欧几里德空间中的一个二阶微分算子，定义为梯度的散度。在图像边缘处理中，二阶微分的边缘定位能力更强，锐化效果更好。

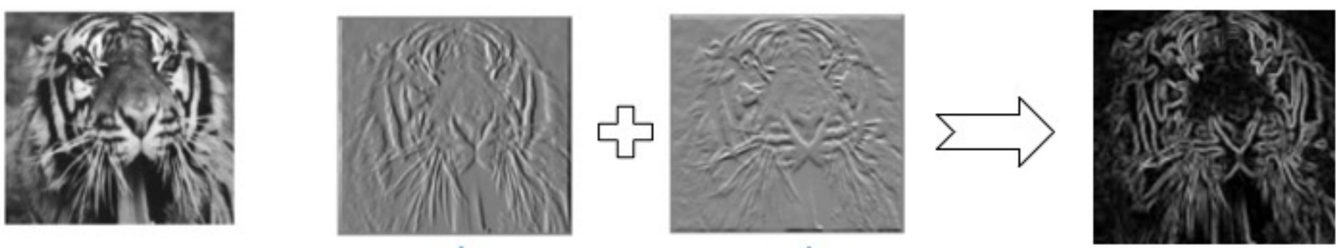
图像在x方向的二阶偏导数

$$\nabla^2 f_{(x,y)} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

图像在y方向的二阶偏导数

优势：快速。 其他常见的边缘检测器（如 Sobel）的计算成本更高，因为它们需要在两个方向上寻找梯度，然后对结果进行归一化。

缺点：输出中的噪声多，边缘定位不如Prewitt和Sobel等算法精准。



图像在x方向的二阶偏导数

图像在y方向的二阶偏导数

1	-2	1

1D

0	1	0
1	-4	1
0	1	0

2D 4邻域

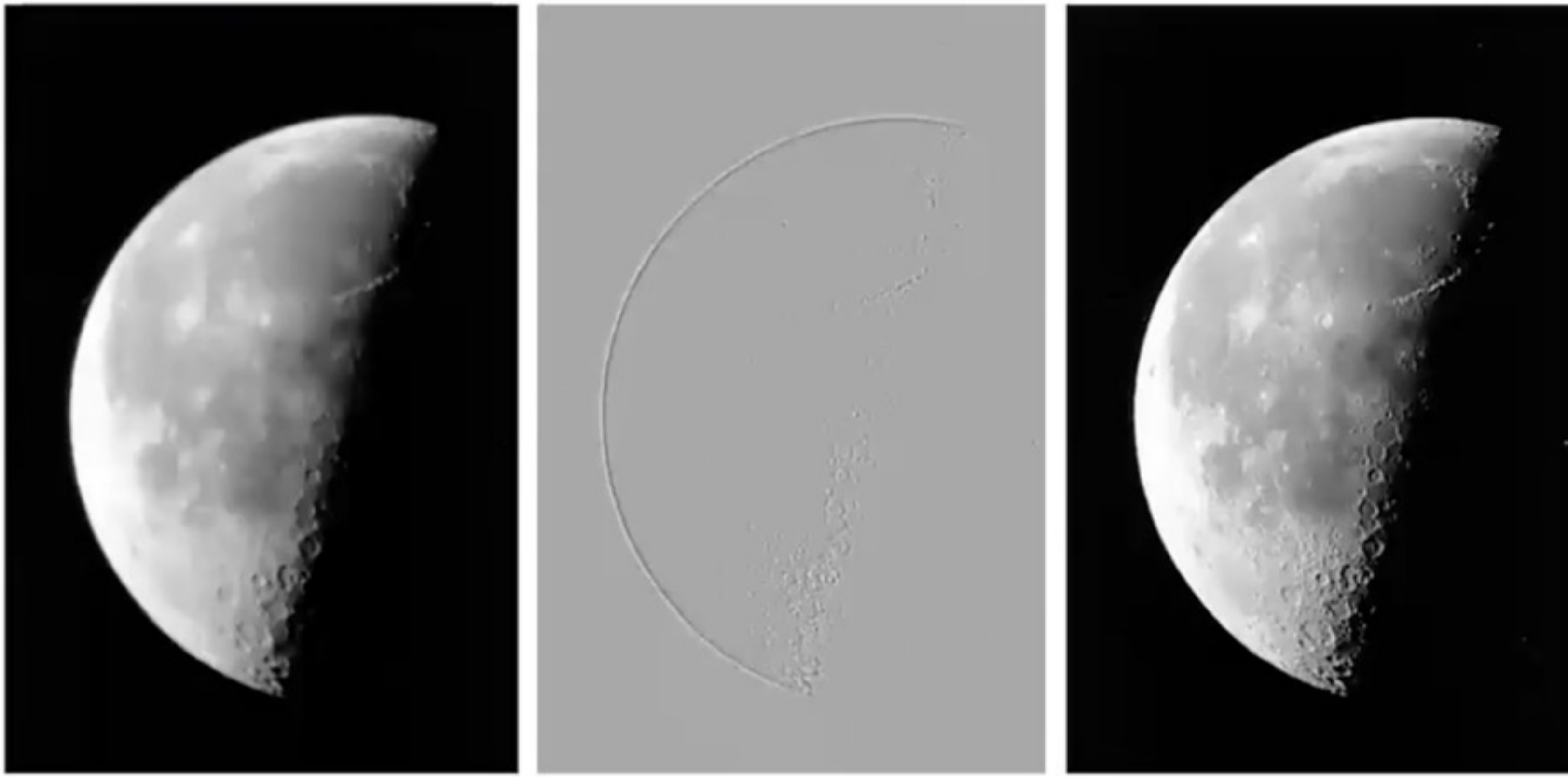
1	1	1
1	-8	1
1	1	1

2D 8邻域

0	0	1	0	0
0	1	2	1	0
1	2	-16	2	1
0	1	2	1	0
0	0	1	0	0

5x5

Laplacian边缘检测器



Laplacian滤波锐化

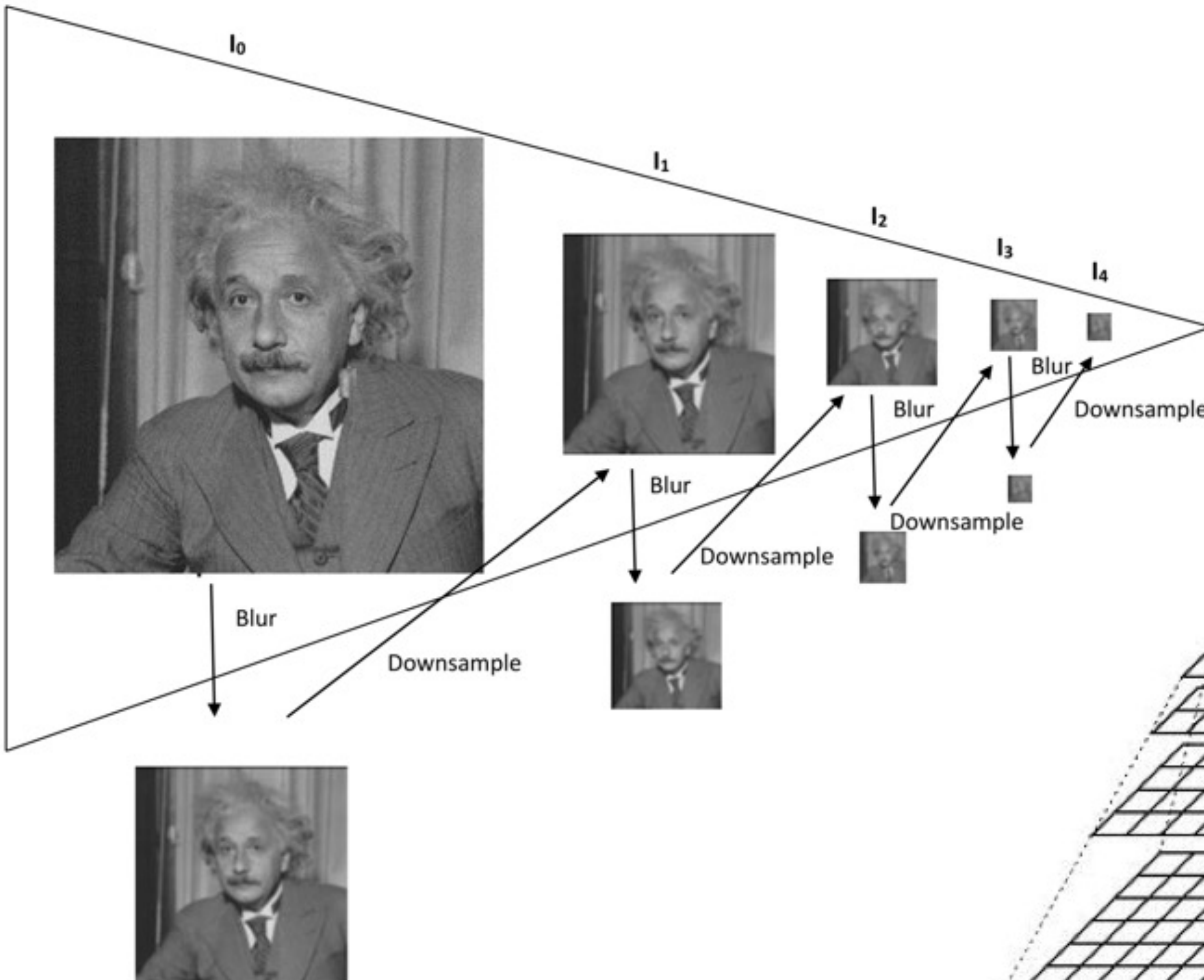
图像锐化

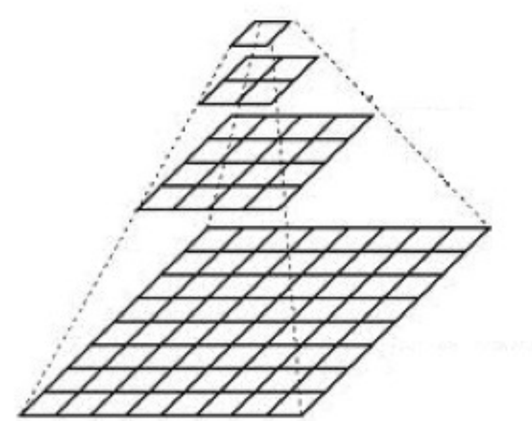
原图

锐化后的结果

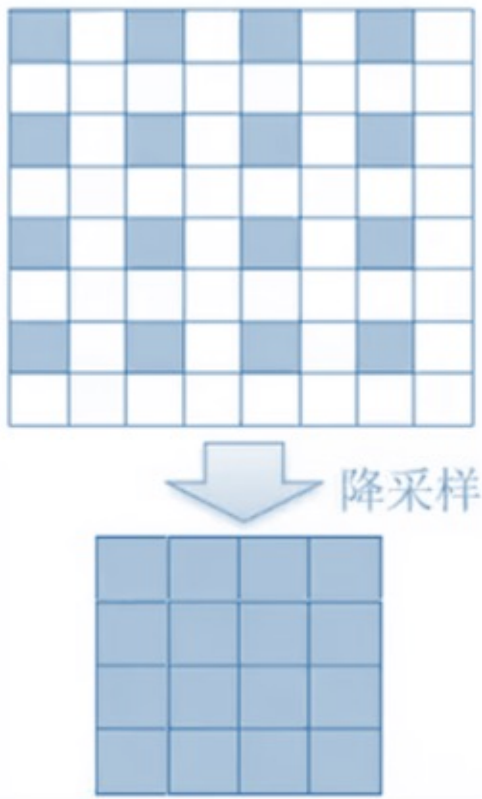
叠加

Gaussian金字塔





Gaussian金字塔



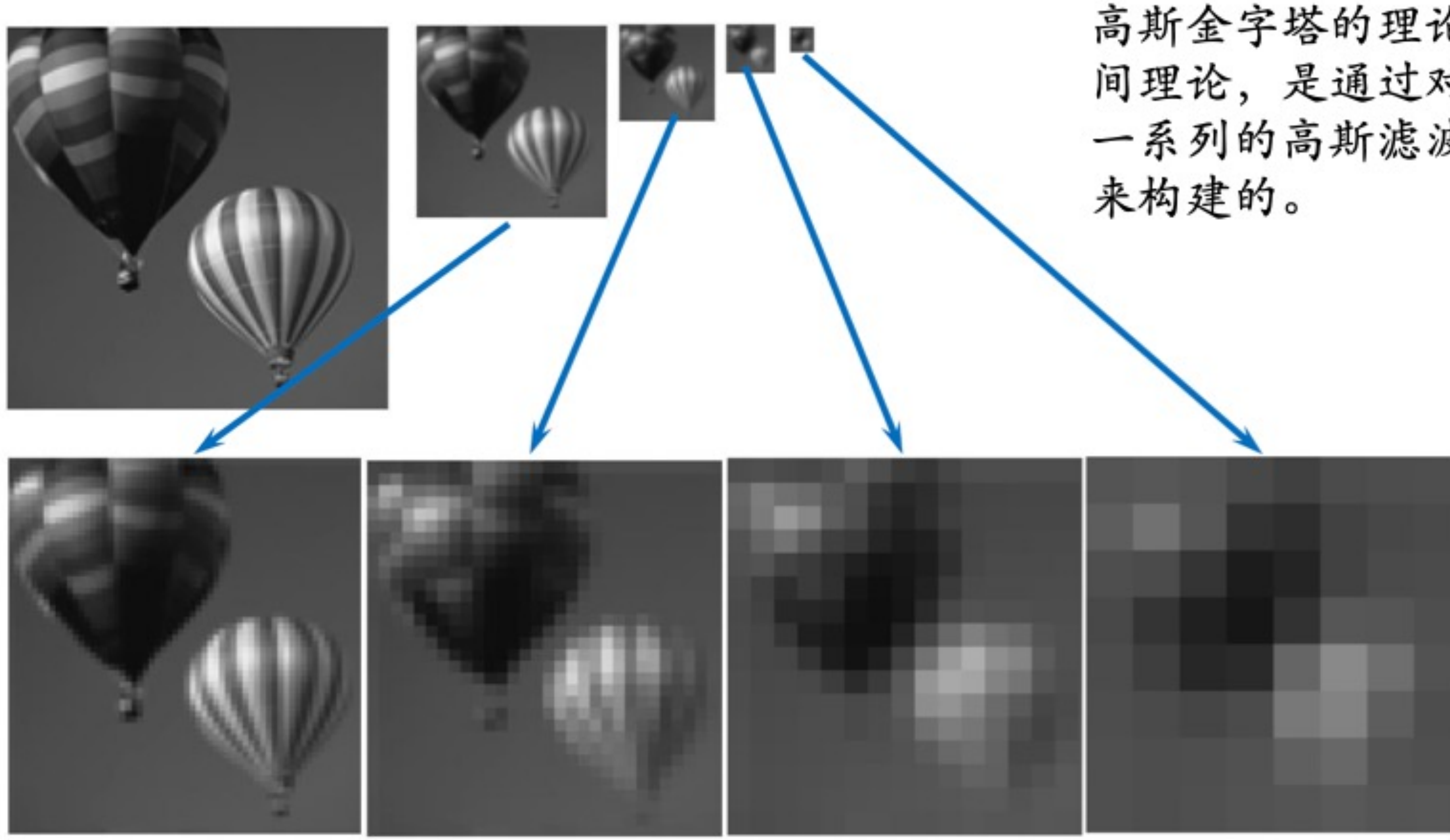
直接降采样（下采样）：会损失重要信息
高斯滤波+降采样：保留主要信息

直接降采样　　高斯+降采样

Gaussian金字塔

高斯金字塔的本质是信号的多尺度表达方式。它将同一信号或图片进行多次高斯模糊，并进行降采样，从而产生不同尺度下的多组信号或图片。

高斯金字塔的理论基础为尺度空间理论，是通过对原始图像进行一系列的高斯滤波和降采样操作来构建的。



Gaussian金字塔

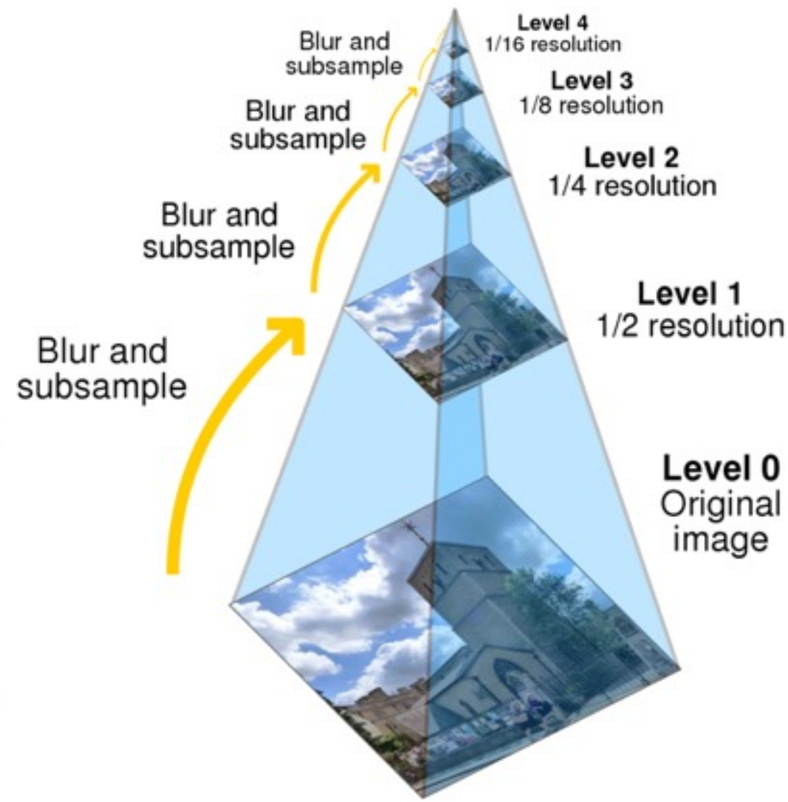
对原始图像进行高斯滤波，消除图像中的高频细节，得到一个平滑后的图像。

将平滑后的图像进行降采样，得到一个缩小后的图像。

这一过程可以重复多次，得到一系列不同尺度的图像。

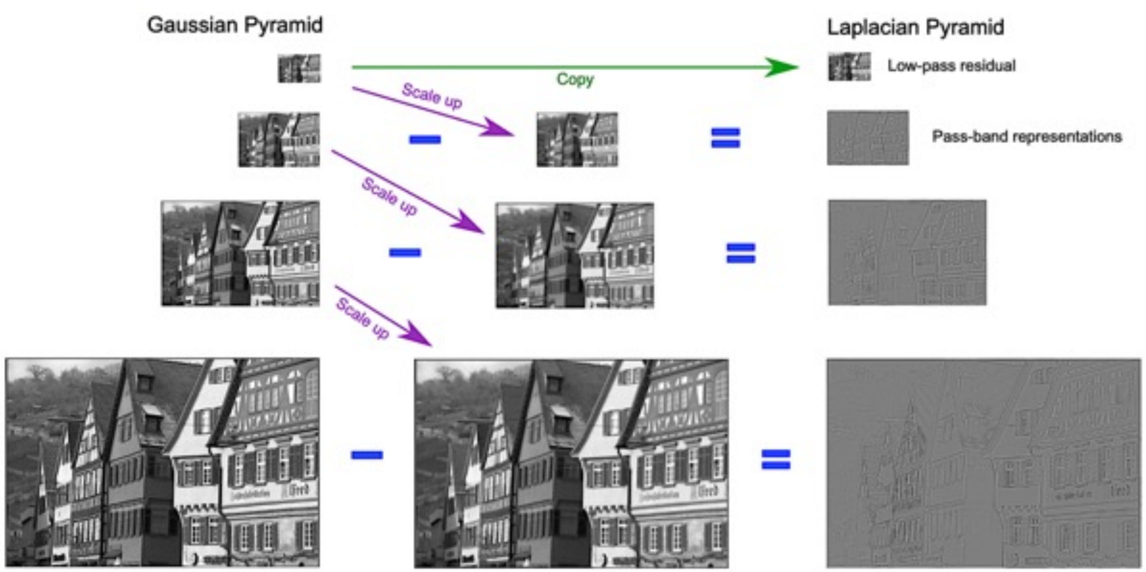
高斯金字塔的每一层都包含了原始图像的模糊版本，层与层之间的差分则表示了原始图像在不同尺度下的细节信息。

高斯滤波器在各个尺度上都能保持图像的平滑性和连续性，能够实现对图像的多尺度分析和处理。



Laplacian金字塔

拉普拉斯金字塔是高斯金字塔的延伸，用于重建高斯金字塔，其包含了一系列从高斯金字塔中派生的残差图像，这些图像表示了高斯金字塔相邻两层图像之间的差异。



对原始图像进行高斯平滑操作。

将平滑后的图像进行降采样，得到一个尺寸减小的图像。

对降采样后的图像进行上采样，恢复到原始图像的大小。

将上采样得到的图像与上一层的高斯金字塔图像进行差分（对应像素值相减）。

得到的结果是拉普拉斯金字塔的一层。

$L_k(I)=G_k(I)-u(G_{k+1}(I))$ 【L 拉普拉斯金字塔；G 高斯金字塔；k层数；u 上采样；I输入图像】

Laplacian金字塔

多尺度

- 拉普拉斯金字塔能够生成一系列不同尺度的残差图像，这些图像包含了原始图像在不同尺度下的细节信息。

信息丰富

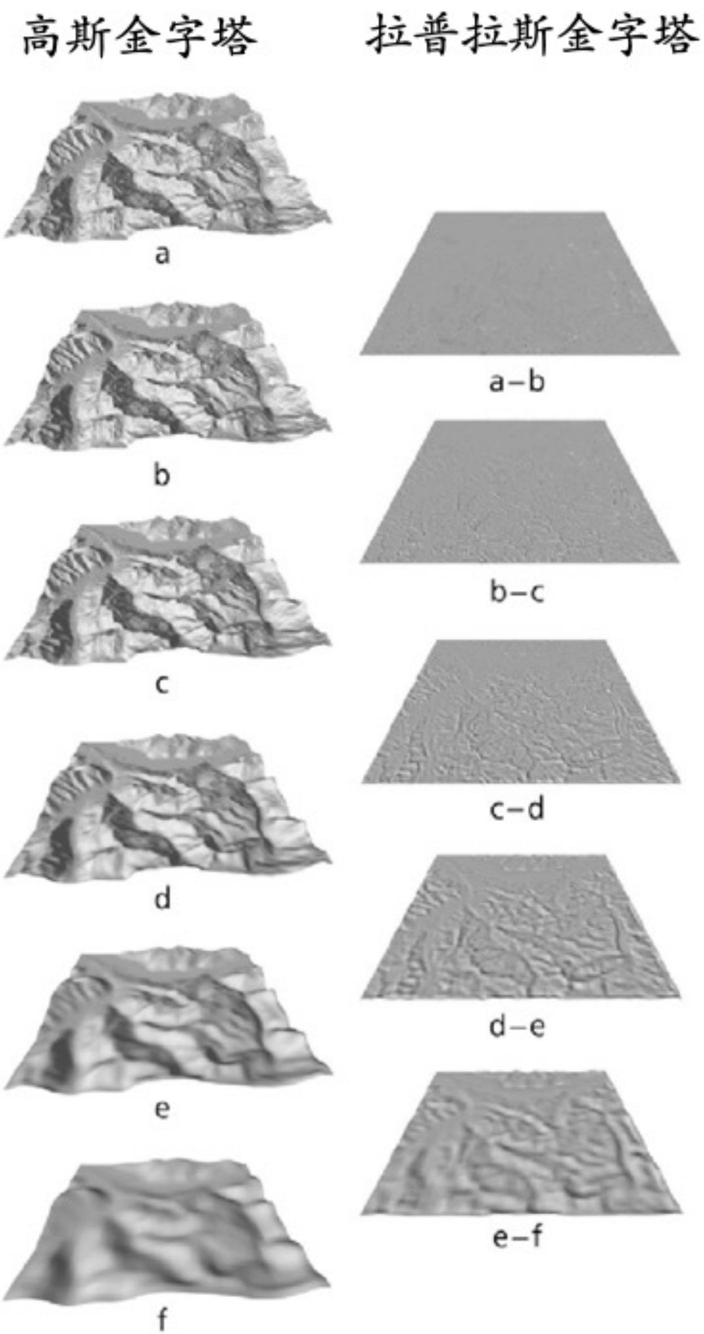
- 每一层拉普拉斯金字塔都包含了上一层的所有信息，使得图像在不同尺度上都具有丰富的信息被保存了下来。

应用广泛

- 拉普拉斯金字塔在图像处理中有多种应用，包括图像压缩（特别是无损压缩）、边缘检测、纹理分析等。

有可逆性

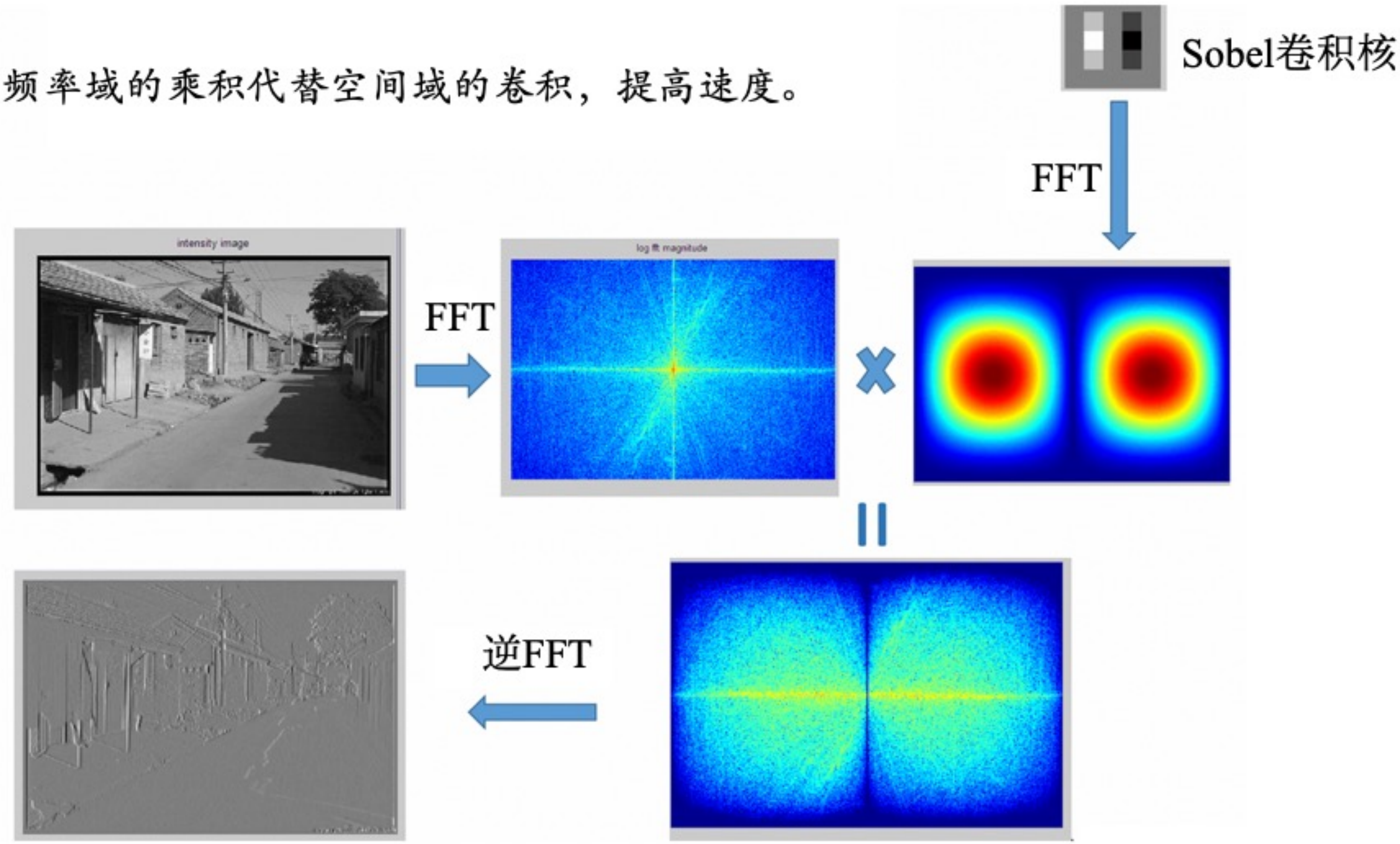
- 与高斯金字塔结合使用时，拉普拉斯金字塔可以用来重建原始图像，这是一个重要的特性，尤其在需要恢复原始数据的应用中。



频率域

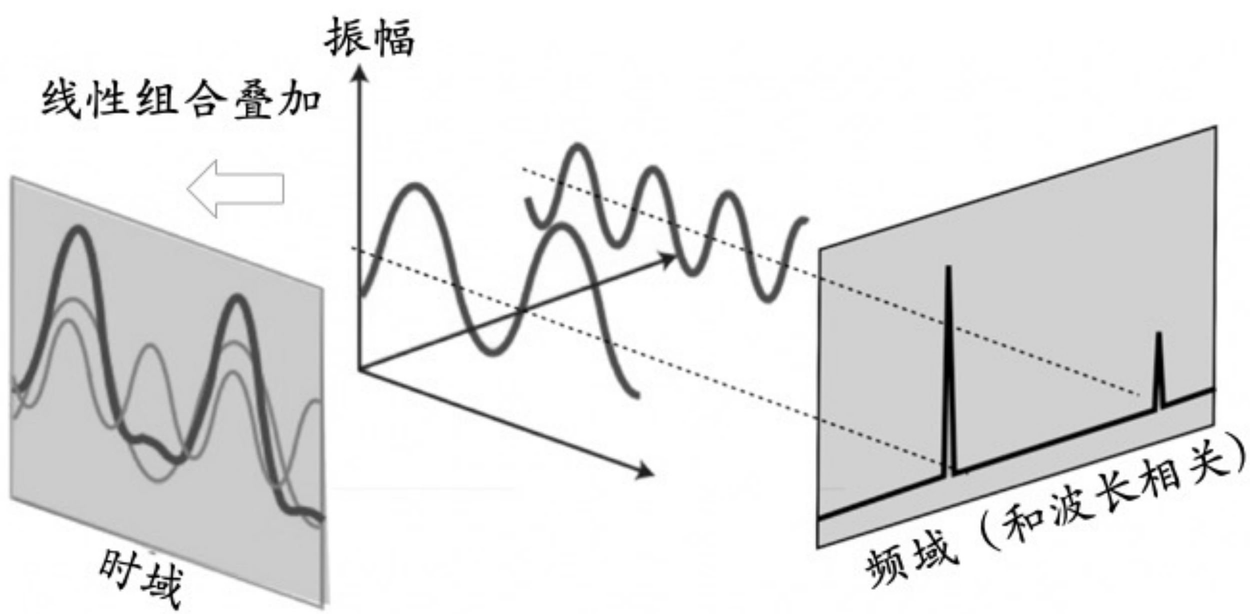
频率域滤波与转换

用频率域的乘积代替空间域的卷积，提高速度。

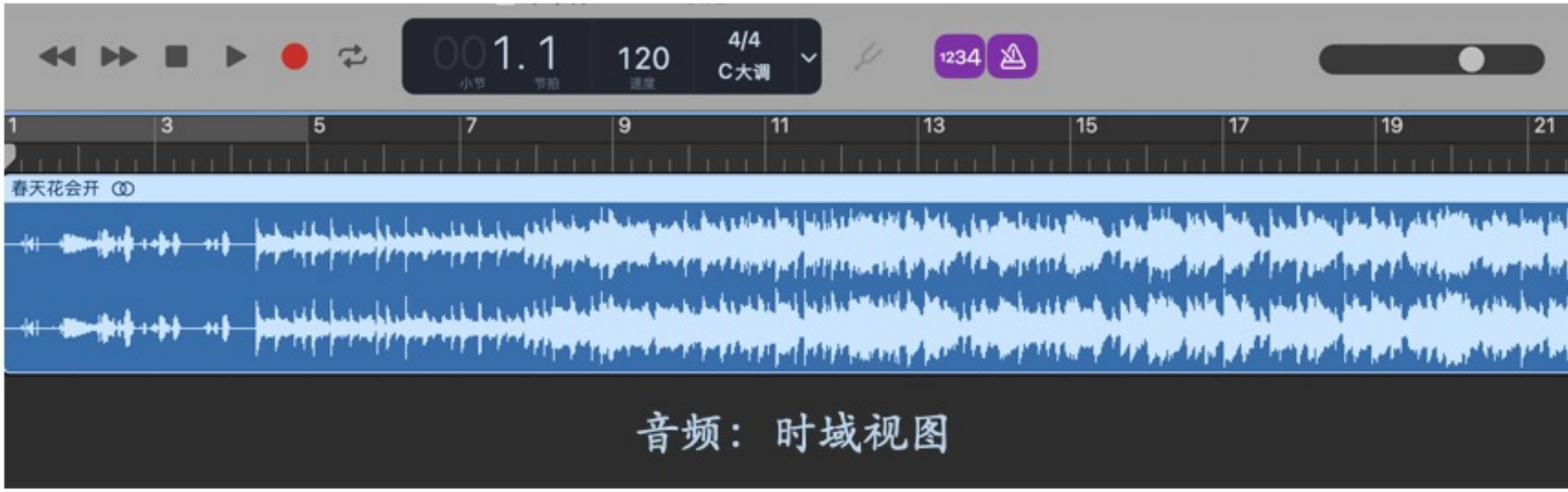


傅里叶变换

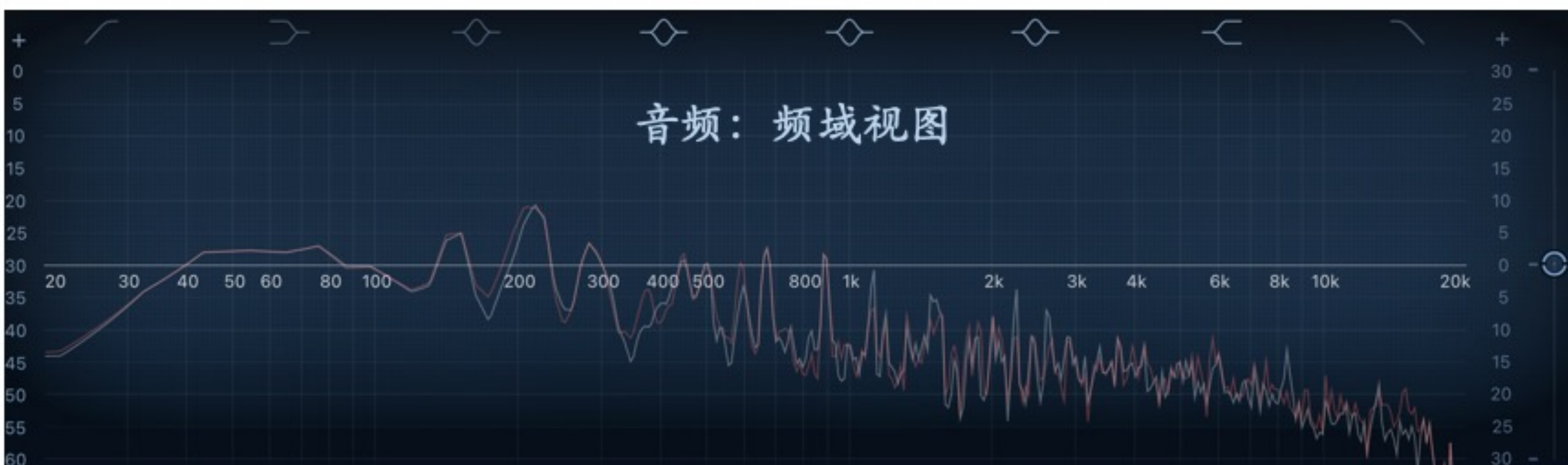
一个函数可以用许多个不同的周期函数（比如正弦/余弦函数等）的线性组合来逼近。这些组合系数在保有原函数的几乎全部信息的同时，还直接地反映了该函数的“频域特征”。



傅里叶变换



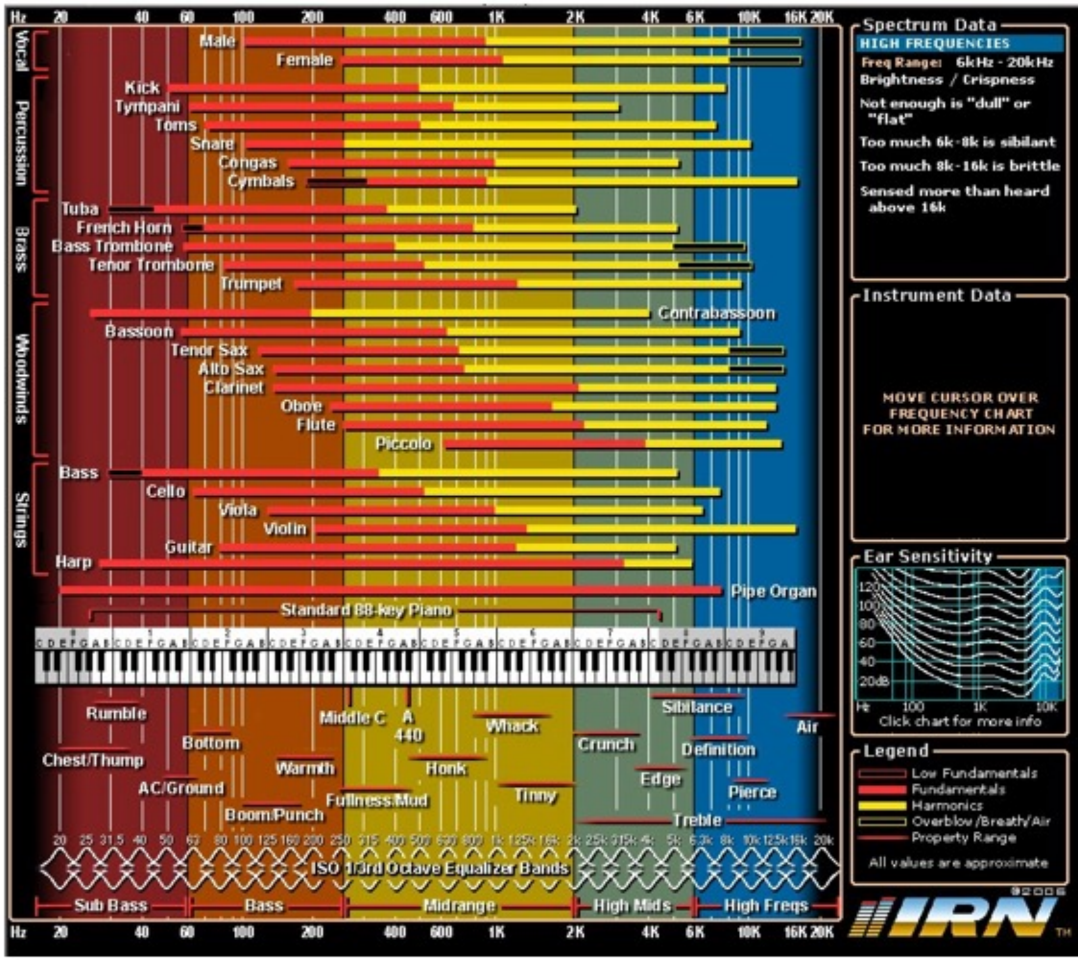
001.1
120
4/4
C大调
1234
1 3 5 7 9 11 13 15 17 19 21
春天花会开
音频：时域视图



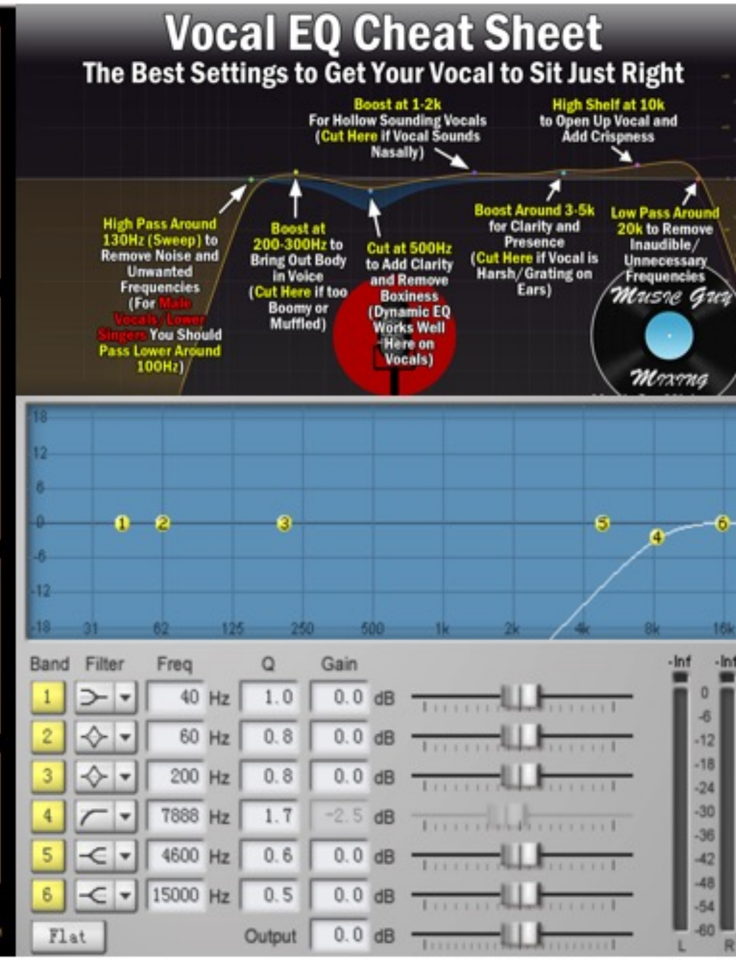
音频：频域视图
20 30 40 50 60 80 100 200 300 400 500 800 1k 2k 3k 4k 6k 8k 10k 20k
0 5 10 15 20 25 30 35 40 45 50 55 60
30 25 20 15 10 5 0 5 10 15 20 25 30

傅里叶变换

乐器



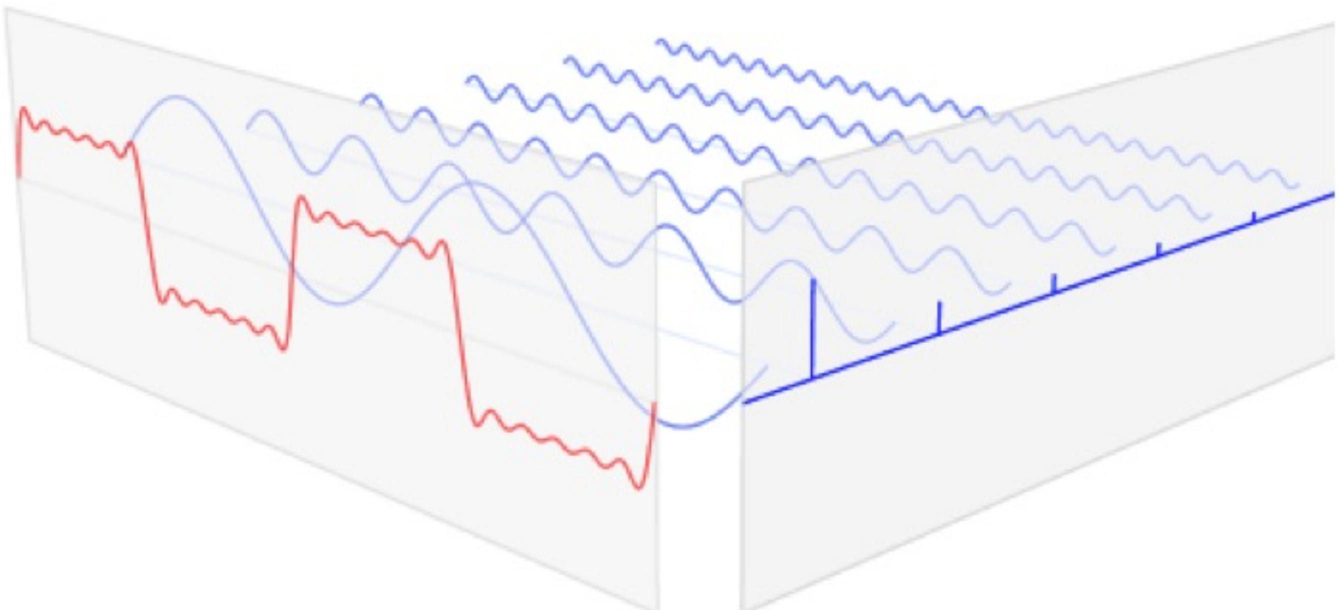
Hz 20 40 60 100 200 400 600 1K 2K 4K 6K 10K 16K 20K
Vocal
Male
Female
Percussion
Kick
Tympani
Toms
Snare
Congas
Cymbals
Brass
Tuba
French Horn
Bass Trombone
Tenor Trombone
Trumpet
Woodwinds
Contrabassoon
Bassoon
Tenor Sax
Alto Sax
Clarinet
Oboe
Flute
Piccolo
Strings
Bass
Cello
Viola
Violin
Guitar
Harp
Pipe Organ
Standard 88-key Piano
Rumble
Chest/Thump
Bottom
AC/Ground
Boom/Punch
Warmth
Middle C
A 440
Fullness/Mud
Honk
Whack
Tinny
Crunch
Sibilance
Definition
Edge
Air
Pierce
Treble
ISO 1/3rd Octave Equalizer Bands
Sub Bass
Bass
Midrange
High Mids
High Freqs
Spectrum Data
HIGH FREQUENCIES
Freq Range: 6kHz - 20kHz
Brightness / Crispness
Not enough is "dull" or "flat"
Too much 6k-8k is sibilant
Too much 8k-16k is brittle
Sensed more than heard above 16k
Instrument Data
MOVE CURSOR OVER FREQUENCY CHART FOR MORE INFORMATION
Ear Sensitivity
Click chart for more info
Legend
Low Fundamentals
Fundamentals
Harmonics
Overblow/Breath/Air
Property Range
All values are approximate
IRN



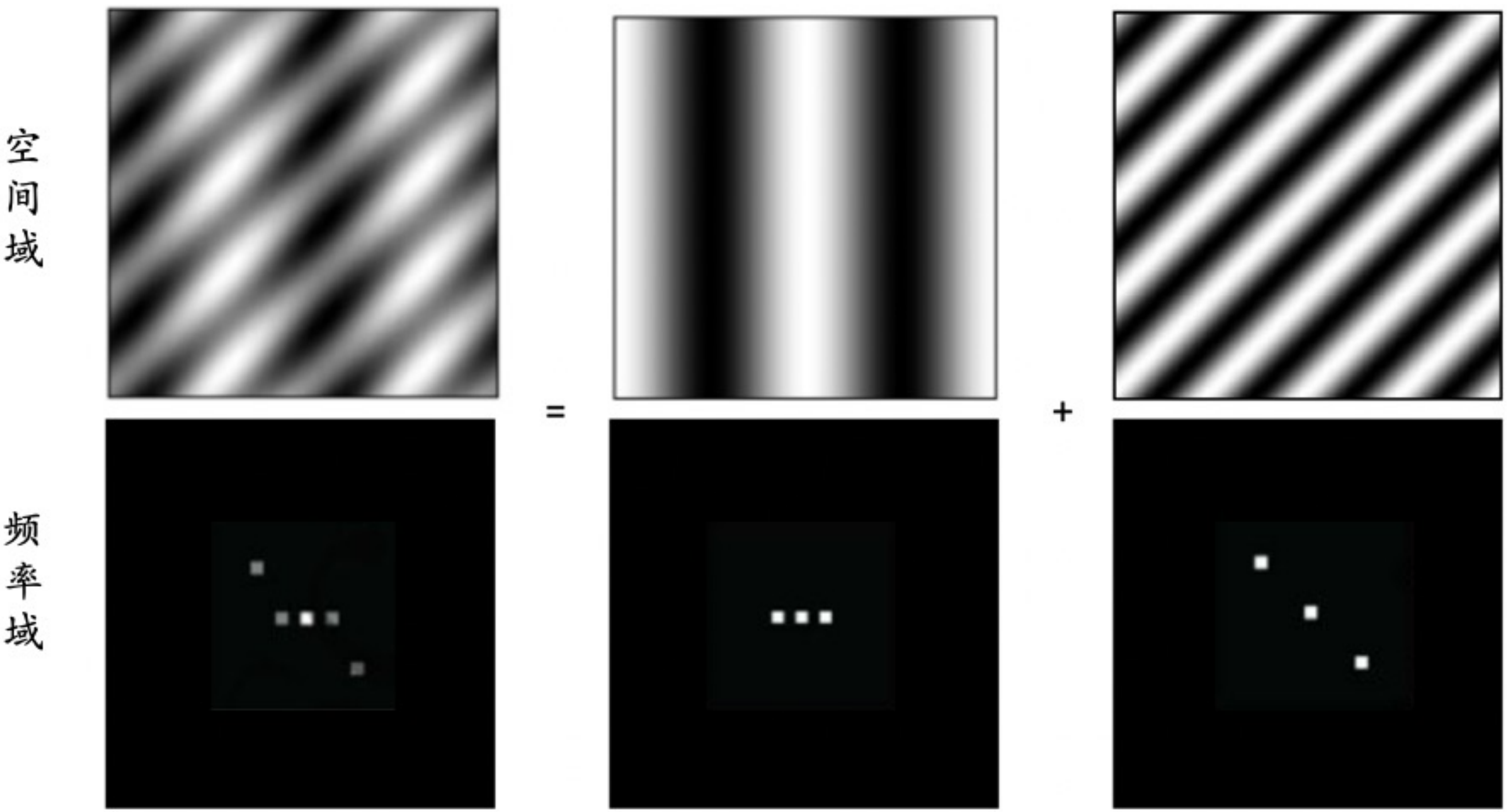
Vocal EQ Cheat Sheet
The Best Settings to Get Your Vocal to Sit Just Right
Boost at 1-2k For Hollow Sounding Vocals (Cut Here if Vocal Sounds Nasally)
High Shelf at 10k to Open Up Vocal and Add Crispness
High Pass Around 130Hz (Sweep) to Remove Noise and Unwanted Frequencies (For Male Vocals, Lower Singers You Should Pass Lower Around 100Hz)
Boost at 200-300Hz to Bring Out Body in Voice (Cut Here if too Boomy or Muffled)
Cut at 500Hz to Add Clarity and Remove Boxiness (Dynamic EQ Works Well Here on Vocals)
Boost Around 3-5k for Clarity and Presence (Cut Here if Vocal is Harsh/Grating on Ears)
Low Pass Around 20k to Remove Inaudible/Unnecessary Frequencies
Music Guy Mixing
Band Filter Freq Q Gain
1 40 Hz 1.0 0.0 dB
2 60 Hz 0.8 0.0 dB
3 200 Hz 0.8 0.0 dB
4 7888 Hz 1.7 -2.5 dB
5 4600 Hz 0.6 0.0 dB
6 15000 Hz 0.5 0.0 dB
Flat
Output 0.0 dB

频率域

傅里叶变换



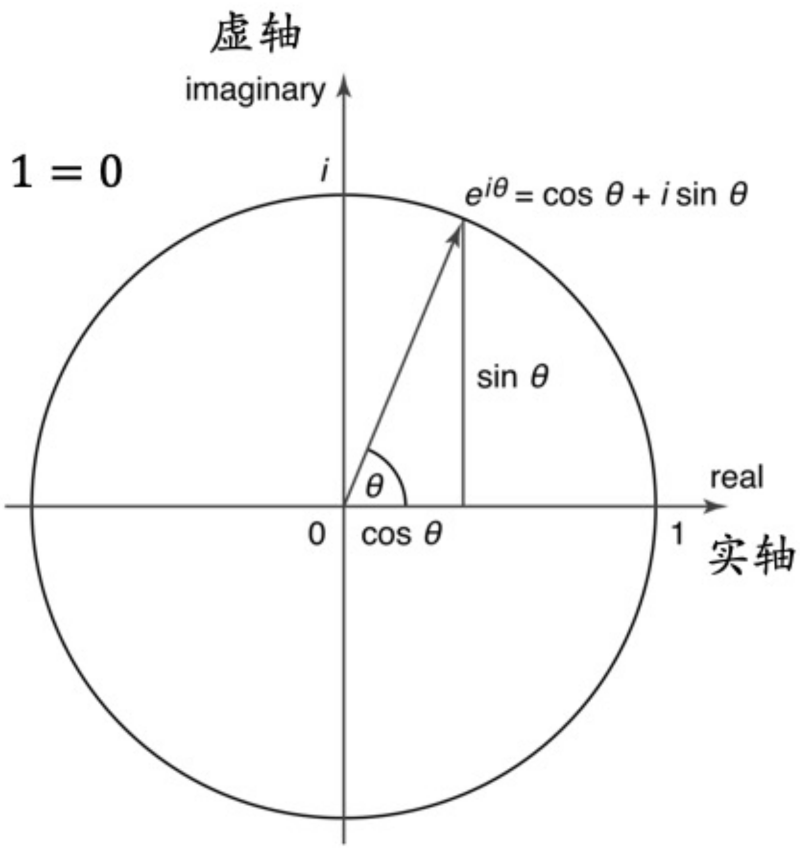
傅里叶变换



信号分解

傅里叶变换

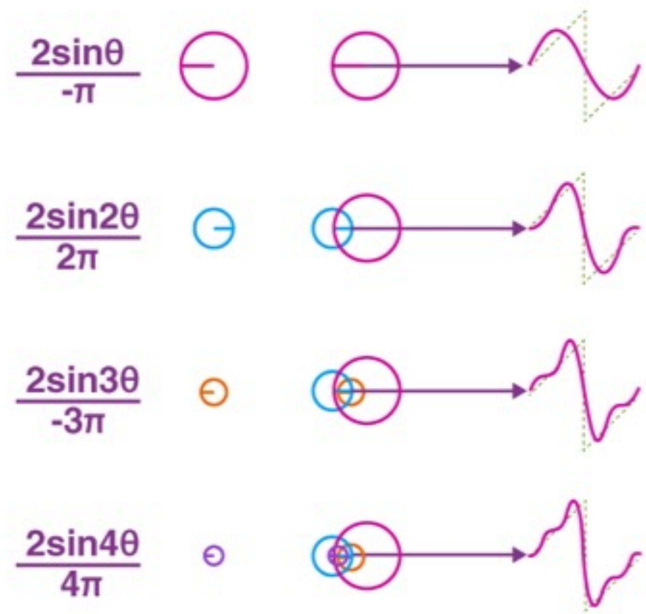
$$e^{\pi i} + 1 = 0$$



$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega$$

通过欧拉公式，我们可以将傅里叶变换中的正弦和余弦函数表示为指数函数，从而提供了更直观和简便的计算方法。

通过欧拉公式，我们可以将复杂的信号分解为简单的正弦波和余弦波成分，这些成分在频域中具有明确的物理意义，有助于我们更好地分析和处理信号。



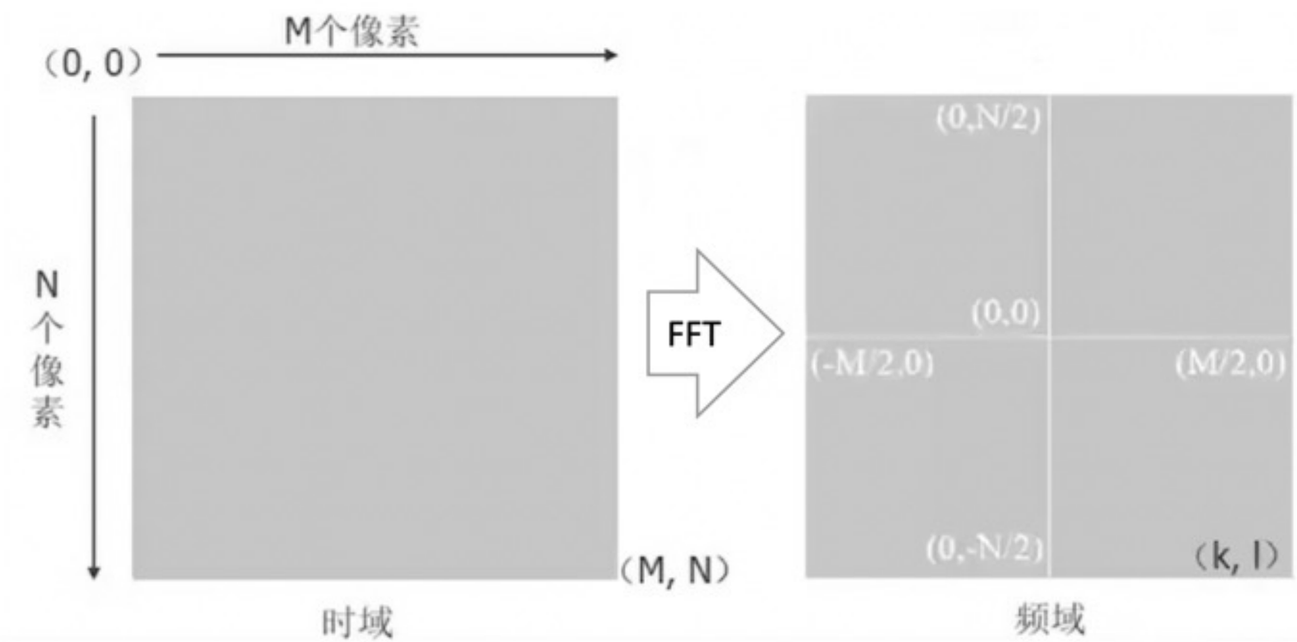
傅里叶变换

离散傅里叶变换

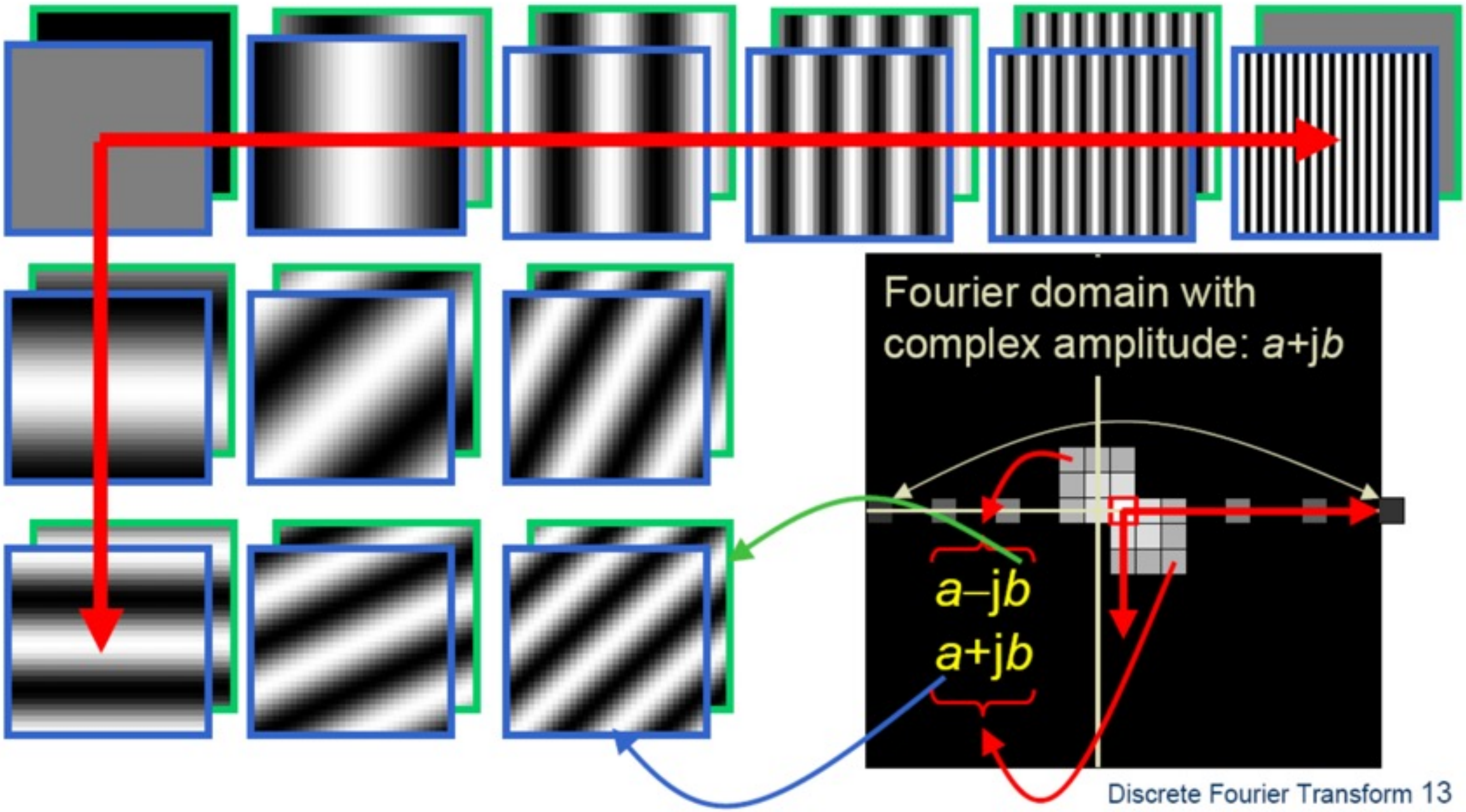
$$X(\omega) = \sum_{n=-\infty}^{\infty} x_n e^{-i\omega n}$$

二维离散傅里叶变换

$$F(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y) e^{-j\cdot 2\pi(\frac{ux}{M}+\frac{vy}{N})}$$

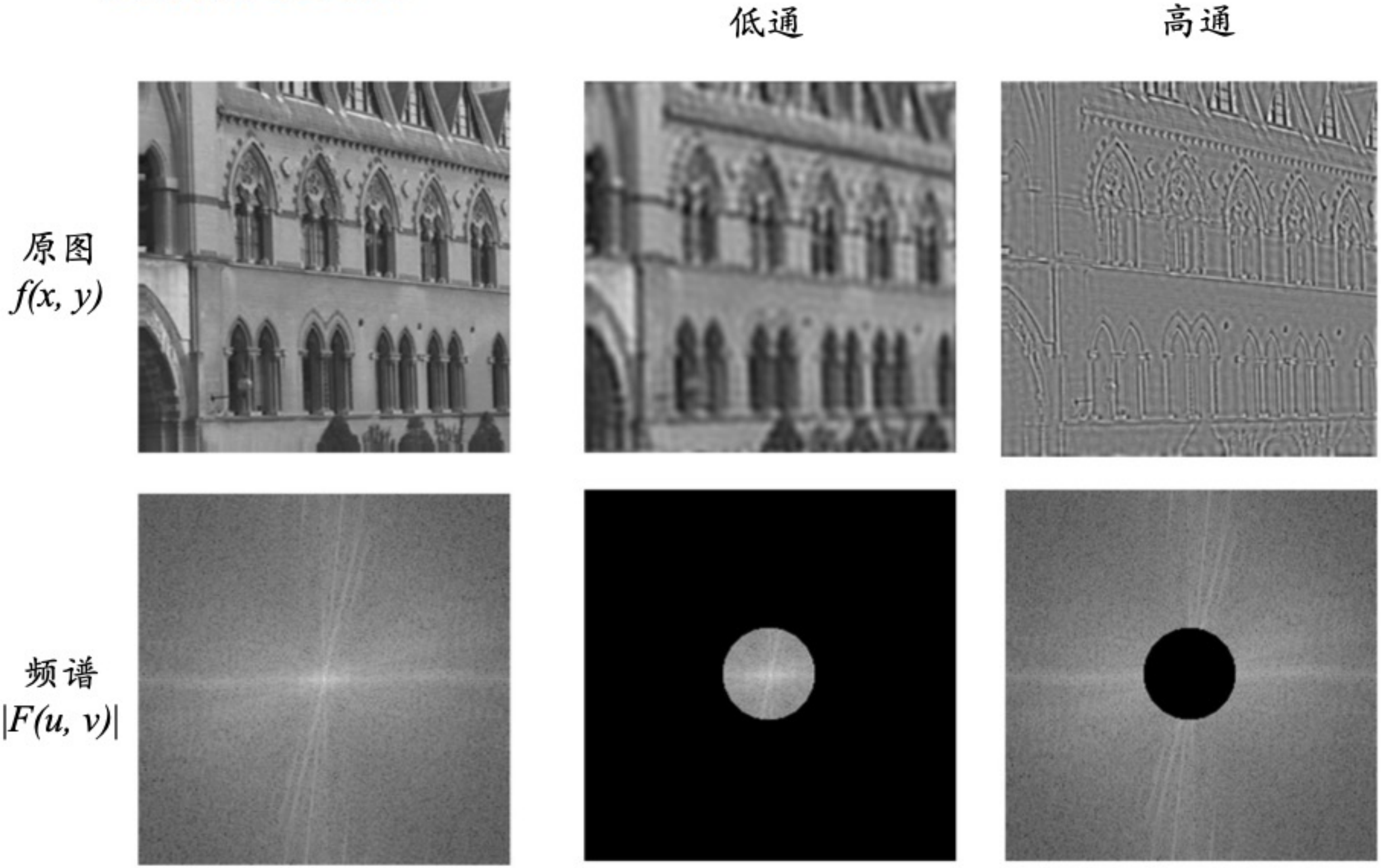


频谱滤波

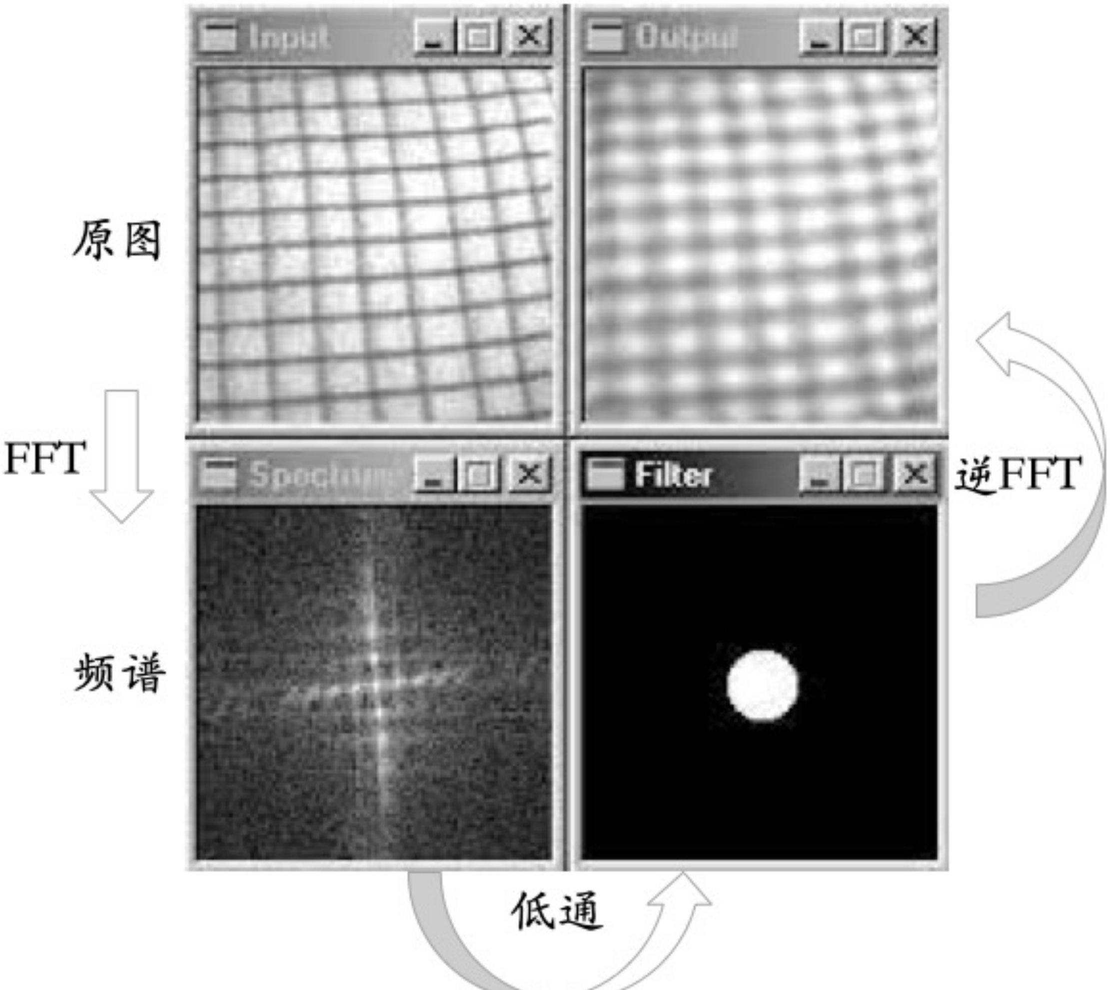


Fourier domain with
complex amplitude: a+jb
a–jb
a+jb
Discrete Fourier Transform 13

频谱滤波



频谱滤波



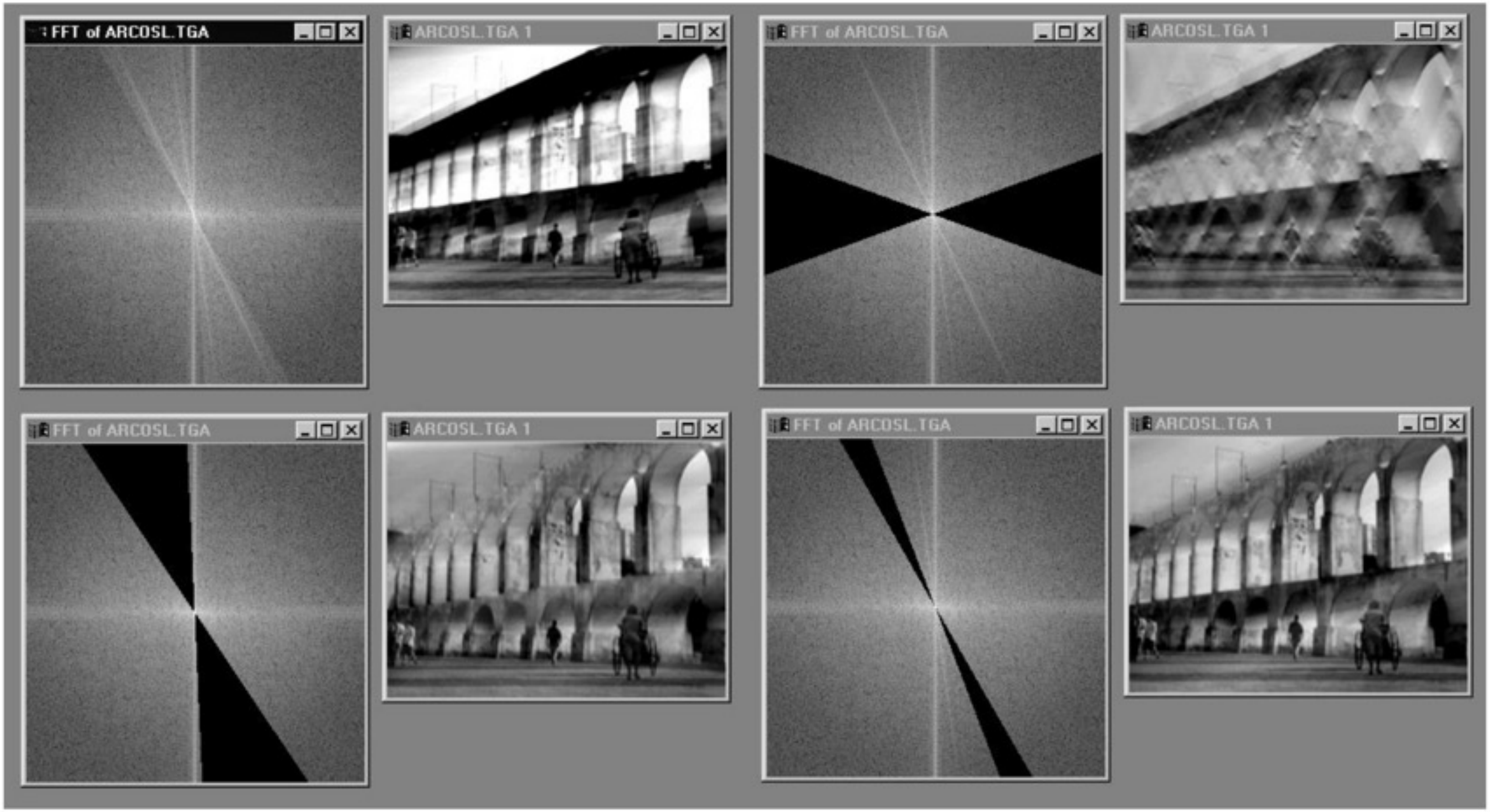
高通滤波：允许高频信号通过，而阻止或减弱低频信号，通常用于**提取图像的细节或边缘信息**。

- 原理：抑制低频成分（平滑区域）而保留高频成分（边缘和细节），突出显示图像中的快速变化部分。
- 应用：边缘检测、细节增强等。例如，在医学图像处理中，高通滤波可以帮助医生更清晰地看到病变组织的边缘。

低通滤波：减少或消除图像中的高频部分，保留低频信息，有助于**平滑图像并去除噪声**。

- 原理：衰减或消除图像中的高频成分（细节和噪声），保留低频成分（整体趋势），从而实现图像的平滑处理。
- 应用：图像去噪、平滑处理。例如，摄影师用低通滤波来减少照片上的颗粒感或消除摩尔纹等不需要的细节。

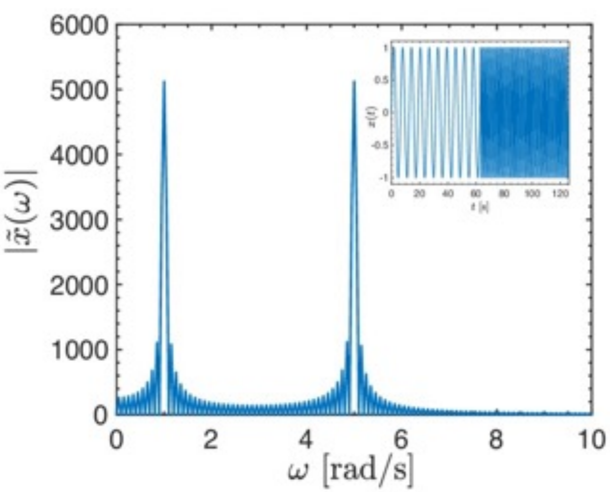
频谱滤波

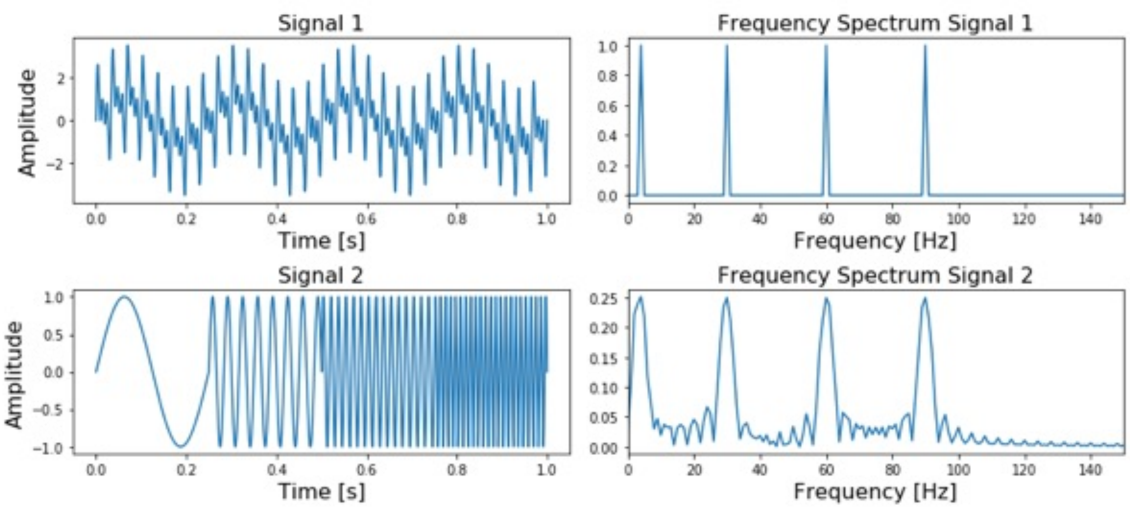


方向位滤波：取消某个方向范围内的细节，但保留范围外其他方向的细节。

短时傅里叶变换

- 傅里叶变换的缺陷：需假设信号平稳，但现实中的信号大部分为非平稳信号，因此缺乏时间和频域的定位功能（频率会突变）。

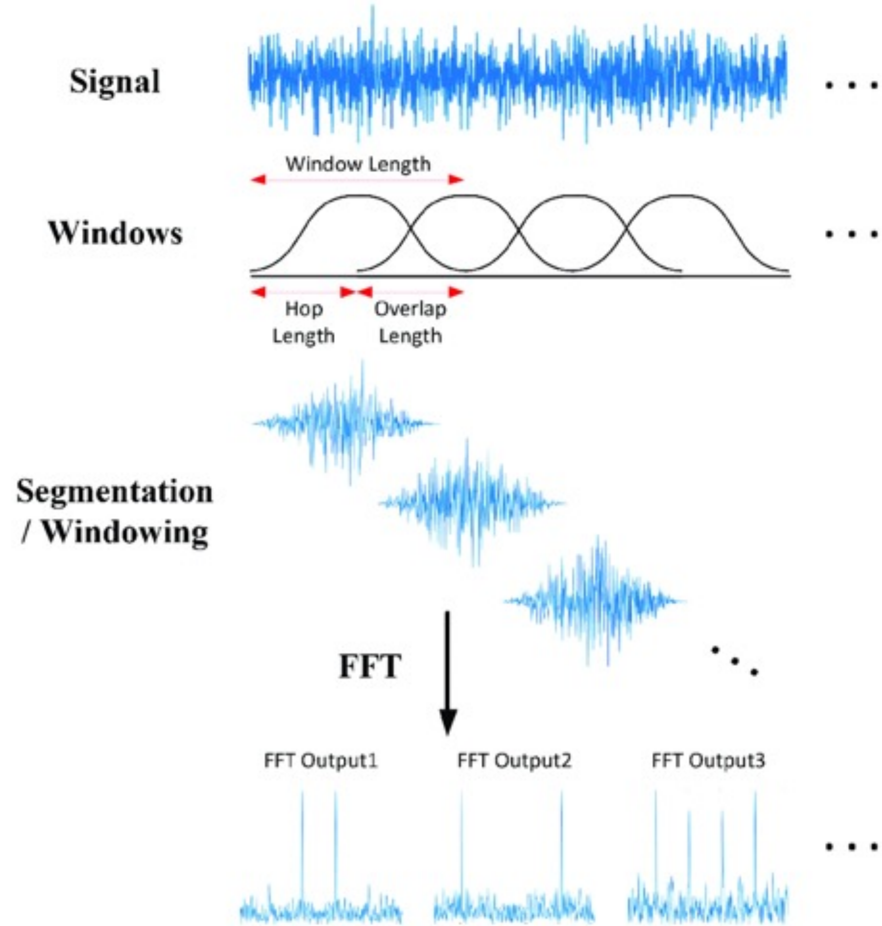




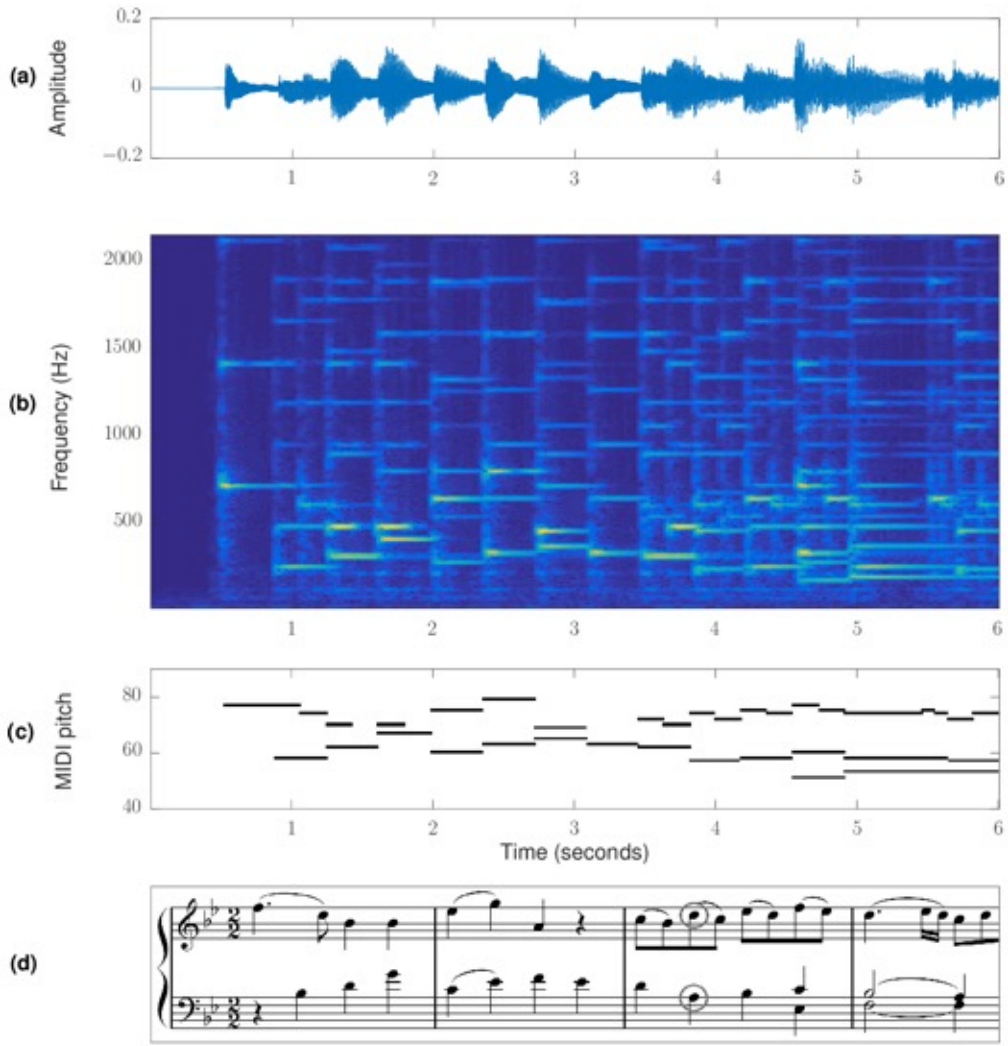
短时傅里叶变换

STFT：选择一个时频局部化的窗口函数g(t)并假设它在一个短时间间隔内是平稳的。移动窗口函数，使原信号f(t)和g(t)的乘积在不同的时刻（有限时间段）内是平稳信号，从而计算出各个不同时刻的频谱。

- 窗口函数：除在给定区间之外，取值均为0的实函数。比如：在给定区间内为常数，而在区间外为0。



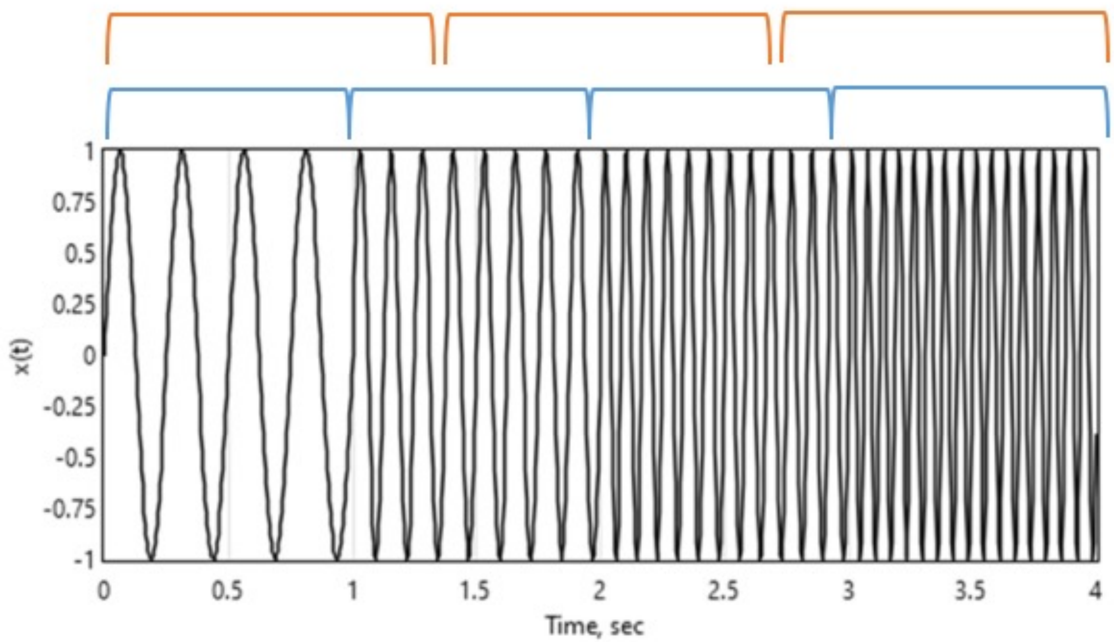
短时傅里叶变换



类比乐谱：每一个音符就是一个窗口。在这个窗口的时长范围内，乐器产生的频率是一定的（有一个固定音高）。

短时傅里叶变换

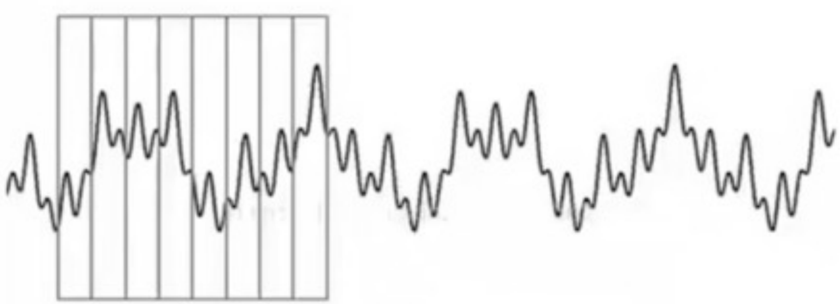
- 优势：获得频域信息的同时也获得时域信息。
- 缺点：窗口大小难以把握。



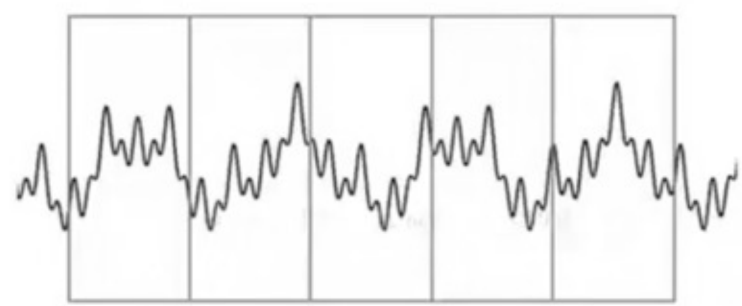
短时傅里叶变换

短时傅里叶变换中，窄窗口的时间分辨率高，但频率分辨率低；宽窗口的时间分辨率低，但是频率分辨率高。

对于时变的非稳定信号，高频适合小窗口，低频适合大窗口。但短时傅里叶变换的窗口是固定的，因此窗口的选择困难是它的缺点，即时间和频率的分辨率不能同时达到最优（海森堡测不准原理）。



时间分辨率高
频率分辨率低



时间分辨率低
频率分辨率高

小波变换

$$F(\omega)=\int_{-\infty}^{\infty} f(t)e^{i\omega t}dt$$

小波变换（Wavelet Transform，WT）继承和发展了短时傅立叶变换的局部化思想，同时克服了窗口大小不随频率变化的缺点。它用有限长或快速衰减的振荡波形（小波）来拟合信号。该波形被缩放和平移以匹配输入的信号。

小波变换提供了一个宽度随频率而改变的时间窗口，从而对时间（空间）频率进行局部化分析。频率增高时，时间窗口的宽度变窄。通过伸缩平移运算，小波变换可以对信号进行多尺度细化，最终达到高频处时间细分、低频处频率细分的效果，能自动适应时频信号分析的要求，从而聚焦到信号的任意细节。

连续小波变换

小波函数　平移参数

$$CWT(a,\tau)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t)\cdot\psi\left(\frac{t-\tau}{a}\right)dt$$

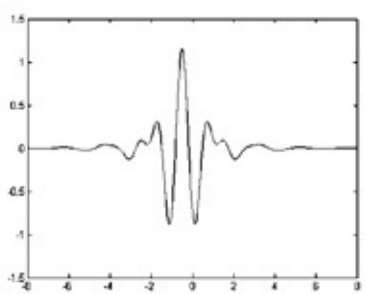
缩放尺度

离散小波变换

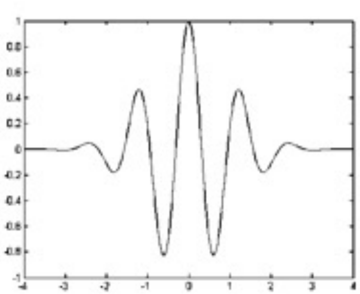
平移参数

$$DWT(j,k)=\int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{2^j}}\psi\left(\frac{t-k2^j}{2^j}\right)dt$$

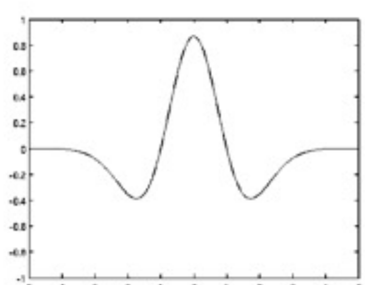
缩放尺度



Meyer小波



Morlet小波



Mexican Hat小波

Haar小波

Haar小波是最早、最简单的小波。

- 它是一种正交小波，由一系列矩形波（也称为方波）构成，具有紧支撑性和正交性，便于数学分析和计算。
- 进行Haar变换时，因为Haar小波函数的正交性，可求出信号在不同频率的Haar函数下所占有的比例。

$$\psi(t) = \begin{cases} 1 & 0 \le t < \frac{1}{2} \\ -1 & \frac{1}{2} \le t < 1 \\ 0 & \text{其他情况} \end{cases}$$