

4.1 短时傅立叶变换--概述



4.2.1 短时傅立叶变换--定义

- 定义：短时傅立叶变换也叫**短时谱**（加窗的方式）

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)e^{-j\omega m}$$

- 短时谱的特点：

- 1)时变性：既是角频率 ω 的函数又是时间 n 的函数
- 2)周期性：是有关 ω 的周期函数，周期为 2π


短时傅立叶变换主要用于语音分析合成系统，由其逆变换能够精确地恢复语音波形；

4.2.1 短时傅立叶变换--定义

- 短时傅里叶变换是窗选语音信号的原则傅里叶变换。下标 n 区别于原则的傅里叶变换。 $w(n-m)$ 是窗口函数序列。不同的窗口函数序列，将得到不同的傅里叶变换的成果。
- 短时傅里叶变换有两个自变量： n 和 ω ，所以它既是有关时间 n 的离散函数，又是有关角频率 ω 的连续函数。
- 与离散傅里叶变换和连续傅里叶变换的关系一样，若令 $\omega = 2\pi k/N$ ，则得离散的短时傅里叶变换，它实际上是在频域的取样。

$$X_n(e^{j\frac{2k\pi}{N}}) = X_n(k) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)e^{-j\frac{2k\pi m}{N}} \quad 0 \leq k \leq N-1$$

4.2.1 短时傅立叶变换--定义

- 这两个公式都有两种解释：
 - ① 当 n 固定不变时，它们是序列 $w(n-m)x(m)$ ($-\infty < m < \infty$)的原则傅里叶变换或原则的离散傅里叶变换。此时与原则傅里叶变换具有相同的性质，而 $X_n(k)$ 与原则的离散傅里叶变换具有相同的特征。
 - ② 当 ω 或 k 固定时，和 $X_n(k)$ 看做是时间 n 的函数。它们是信号序列和窗口函数序列的卷积，此时窗口的作用相当于一种滤波器。 

4.2.1 短时傅立叶变换--定义

- 频率辨别率 Δf 、取样周期 T 、加窗宽度 N 三者关系:

$$\Delta f = \frac{1}{NT}$$

- 窗形状对短时傅立叶变换的影响
 - 矩形窗——主瓣窄，衰减慢；
 - 汉明窗——主瓣宽，衰减快；
- 窗宽对短时频谱的影响
 - 窗宽长——频率辨别率高，能看到频谱快变化；
 - 窗宽短——频率辨别率低，看不到频谱的快变化；

4.2.2 短时傅立叶变换--原则傅里叶变换的解释

- 短时傅里叶变换可写为 

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} [x(m)w(n-m)]e^{-j\omega m}$$



- 当n取不同值时窗w(n-m)沿着x(m)序列滑动，所以w(n-m)是一种“滑动的”窗口。
- 因为窗口是有限长度的，满足绝对可和条件，所以这个变换是存在的。与序列的傅里叶变换相同，短时傅里叶变换伴随 ω 作周期变化，周期为 2π 。

4.2.2 短时傅立叶变换--原则傅里叶变换的解释

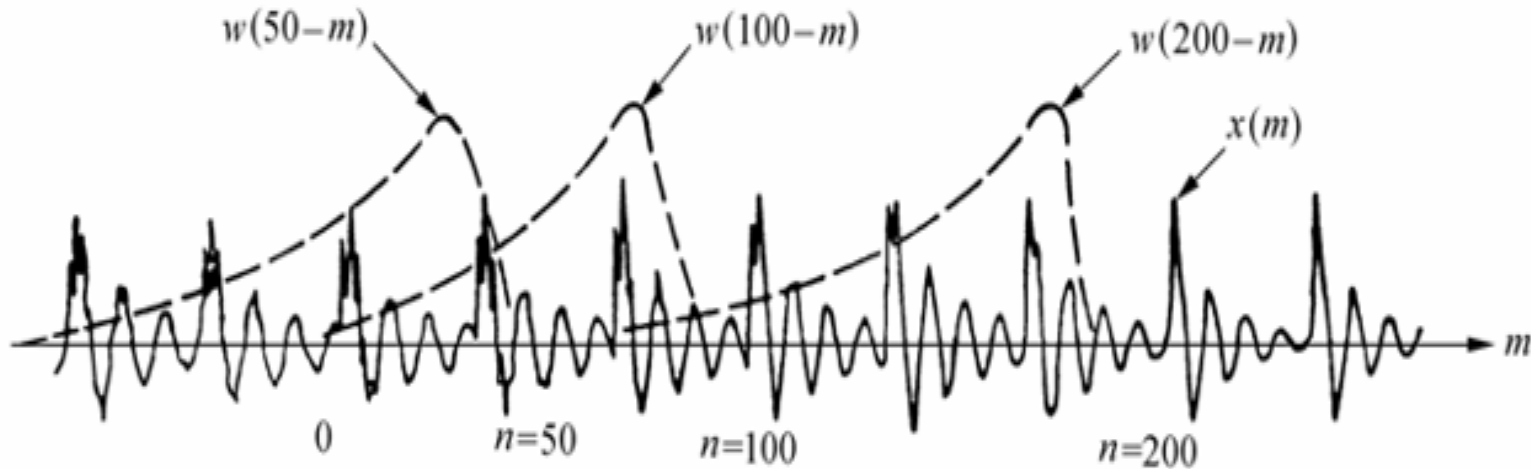


图 4-1 在几个 n 值上 $x(m)$ 与 $w(n-m)$ 的示意图

4.2.2 短时傅立叶变换--原则傅里叶变换的解释

- 根据功率谱定义，能够写出短时功率谱与短时傅里叶变换之间的关系 $\Psi S_n(e^{j\omega}) = X_n(e^{j\omega}) \bullet X_n^*(e^{j\omega}) = |X_n(e^{j\omega})|^2$



- 式中*表达复共轭运算。同步功率谱是短时自有关函数 Ψ



$$R_n(k) = \sum_{m=-\infty}^{\infty} w(n-m)x(m)w(n-k-m)x(m+k) \quad \text{的傅里叶变换。} \Psi$$

- 下面将短时傅里叶变换写为另一种形式。设信号序列和窗口序列的原则傅里叶变换为 Ψ

$$X(e^{j\omega}) = \sum_{m=-\infty}^{\infty} x(m)e^{-j\omega m}$$

$$W(e^{j\omega}) = \sum_{m=-\infty}^{\infty} w(m)e^{-j\omega m}$$

均存在。当n取固定值时， $w(n-m)$ 的傅里叶变换为 Ψ








$$\sum_{m=-\infty}^{\infty} w(n-m)e^{-j\omega m} = e^{-j\omega n} \bullet W(e^{-j\omega})$$

4.2.2 短时傅立叶变换--原则傅里叶变换的解释

- 根据傅里叶变换的频域卷积定理，有 \Downarrow

$$X_n(e^{j\omega}) = X(e^{j\omega}) * [e^{-j\omega n} \bullet W(e^{-j\omega})]$$

4.2.2 短时傅立叶变换--原则傅里叶变换的解释

- 用波形乘以窗函数，不但为了在窗口边沿两端不引起急剧变化，使波形缓慢降为零，而且还相当于对信号谱与窗函数的傅里叶变换进行卷积。
- 为此窗函数应具有如下特征：
 - ① 频率辨别率高，即主瓣狭窄、锋利；（矩形窗）
 - ② 经过卷积，在其他频率成份产生的频谱泄漏少，即旁瓣衰减大。（海明窗）
 - 这两个要求实际上相互矛盾，不能同步满足。
- 窗口宽度 N 、取样周期 T 和频率辨别率 Δf 之间存在下列关系 $\Delta f = 1/NT$ 
- 可见：
 - 窗口宽度 \uparrow \rightarrow 频率辨别率 \uparrow 时间辨别率 \downarrow
 - 窗口宽度 \downarrow \rightarrow 频率辨别率 \downarrow 时间辨别率 \uparrow ，因而两者是矛盾的。

4.2.2 短时傅立叶变换--原则傅里叶变换的解释

- 第一种零点位置为 $2\pi/N$ ，显然它与窗口宽度成反比。
 - 矩形窗，虽然频率辨别率很高，但因为第一旁瓣的衰减只有13.2dB，所以不适用于频谱成份动态范围很宽的语音分析中。
 - 海明窗在频率范围中的辨别率较高，而且因为旁瓣的衰减不小于42dB，具有频谱泄漏少的优点，频谱中高频分量弱、波动小，因而得到较平滑的谱。
 - 汉宁窗是高次旁瓣低，第一旁瓣衰减只有30dB。
- 对语音波形乘以海明窗，压缩了接近窗两端的部分波形，等效于用作分析的区间缩短40%左右，所以，频率辨别率下降40%左右。所以，虽然在基音周期性明显的浊音频谱分析中，乘以合适的窗函数，也能克制基音周期与分析区间的相对相位关系的变动影响，从而得到稳定的频谱。因为乘以窗函数将造成分帧区间缩短，所觉得跟踪随时间变化的频谱，要求一部分区间反复移动。

4.2.2 短时傅立叶变换--原则傅里叶变换的解释

为使 $X_n(e^{j\omega})$ 准确再现 $X(e^{j\omega})$ 的特性, $W(e^{j\omega})$ 相对于 $X(e^{j\omega})$ 来说应是一个冲激函数。

N 越大, 窗傅里叶变换 $W(e^{j\omega})$ 的主瓣越窄, $X_n(e^{j\omega})$ 越接近于 $X(e^{j\omega})$ 。

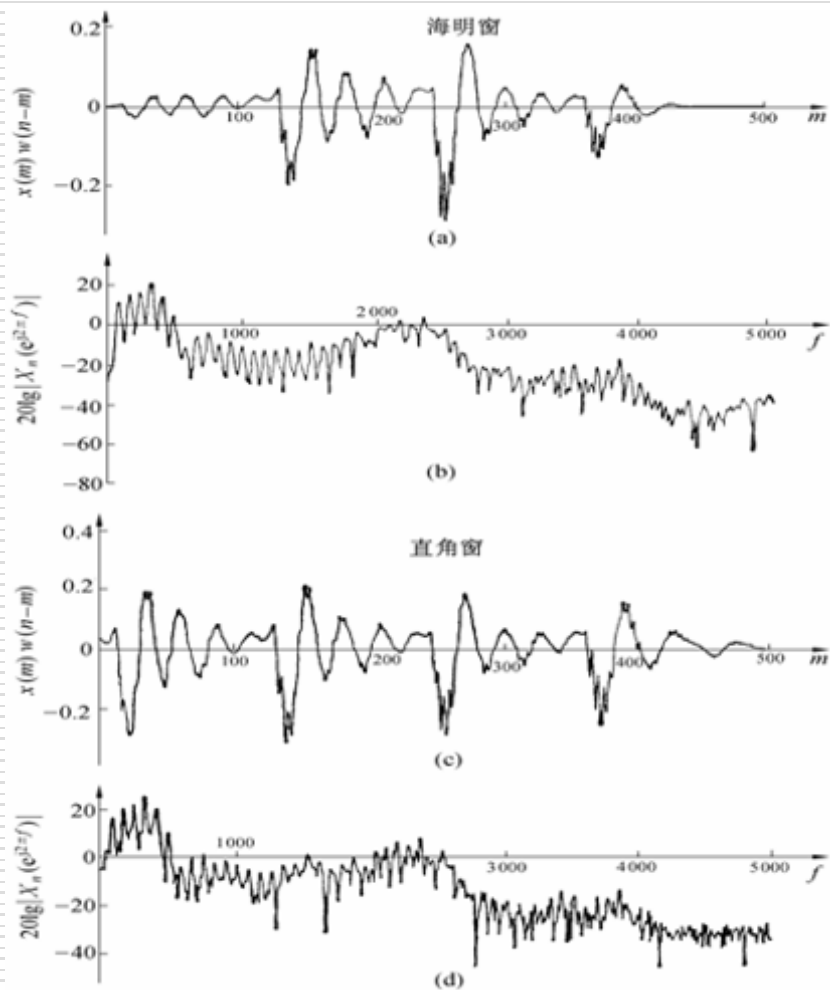
当 $N \rightarrow \infty$ 时, $X_n(e^{j\omega}) \rightarrow X(e^{j\omega})$ 。

但是 **N 值太大时, 信号的分帧已失去了意义**, 尤其是 N 值大于语音的音素长度时, $X_n(e^{j\omega})$

已不能反映该语音音素的频谱了。

应折衷选择窗宽 N。

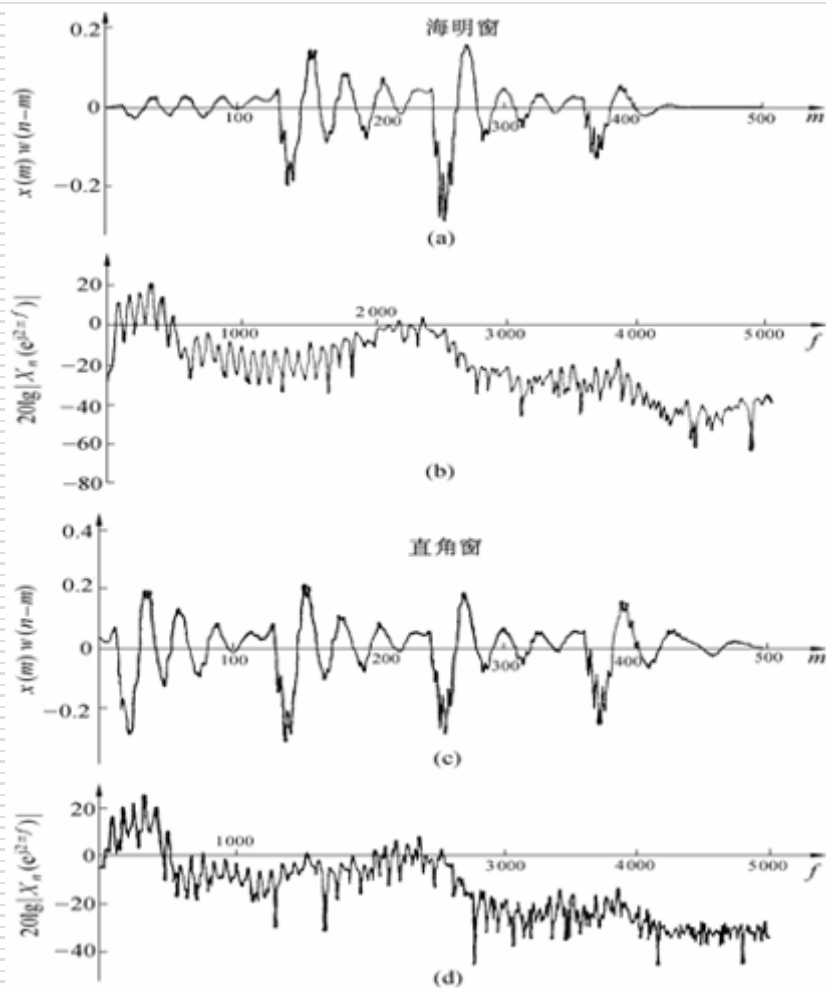
4.2.2 短时傅立叶变换--原则傅里叶变换的解释



- 其中图(a)是海明窗的窗选信号,图(b)是其对数功率谱;图(c)是矩形窗下的窗选信号,图(d)是其对数功率谱。
- 从图(a)能够明显看出时间波形的周期性,此周期性一样在图(b)中体现出来。图中基频及其谐波在频谱中体现为等频率间隔的窄峰。图(b)中的频谱大约在300~400Hz附近有较强的第一共振峰,而约在2023Hz附近有一种相应于第二、三共振峰的宽峰。另外,还能在3800Hz附近看到第四个共振峰。最终,因为声门脉冲谱的高频衰减特征,频谱在高频部分体现出下降的趋势。

给出了 $N=500$ 时(取样率10 kHz,窗连续时间50 ms)时直角窗及海明窗下浊音语音的频谱。

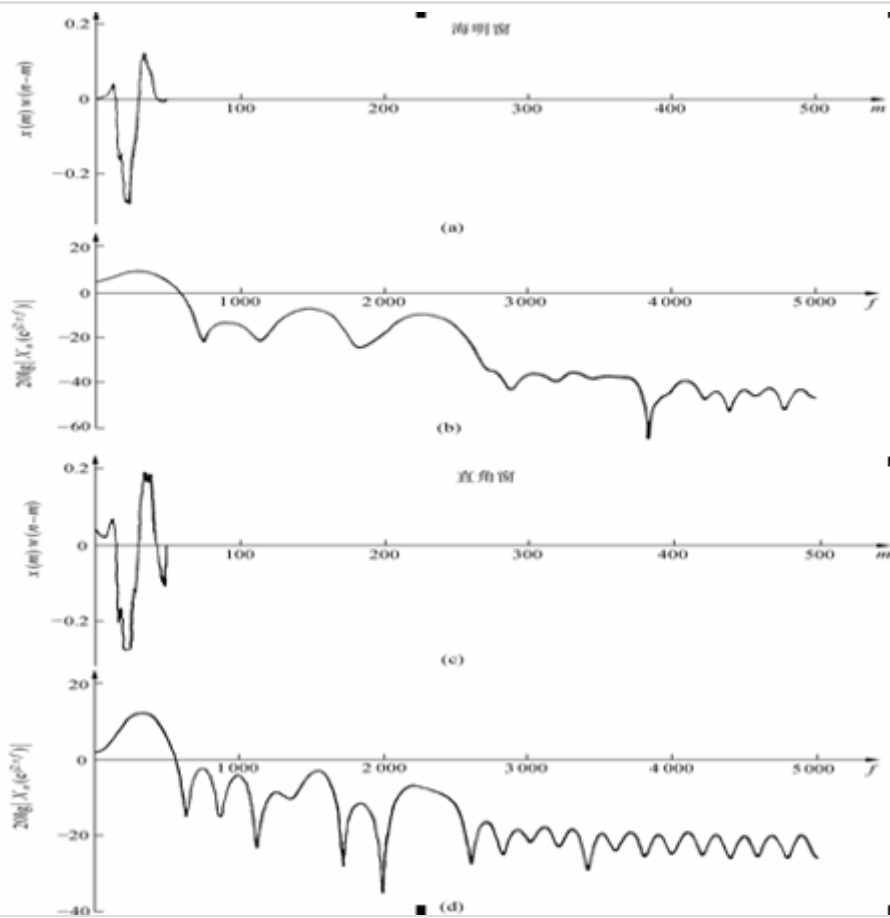
4.2.2 短时傅立叶变换--原则傅里叶变换的解释



- 将图(b)和图(d)比较可看出它们在基音谐波、共振峰构造以及频谱粗略形状上的相同性,一样也能看到其频谱之间的差别。
- 最明显的是图(d)中基音谐波峰利度增长,这主要是因为矩形窗频率辨别率较高。
- 另一差别是矩形窗较高的旁瓣产生了一种类似于噪声的频谱。这是因为相邻谐波的旁瓣在谐波间隔内的相互作用(有时加强有时抵消),因而在谐波间产生了随机变化。这种相邻谐波间不希望有的“泄漏”抵消了其主瓣较窄的优点,
- 所以在语音频谱分析中极少采用矩形窗。🔥

给出了 $N=500$ 时(取样率10 kHz,窗连续时间50 ms)时直角窗及海明窗下浊音语音的频谱。

4.2.2 短时傅立叶变换--原则傅里叶变换的解释



- 图4-3给出了 $N=50$ 的比较成果(取样率与图4-2中相同,因而窗口连续时间为5ms)。
- 因为窗口很短,因而时间序列(图(a)和(c))及信号频谱(图(b)和(d))均不能反应信号的周期性。
- 与图4-2相反,图4-3只大约在400、1400及2200Hz频率上有少许较宽的峰值。它们与窗内语音段的前三个共振峰相相应。比较图4-3(b)及(d)的频谱后,再次表白矩形窗能够得到较高的频率辨别率。

4.2.2 短时傅立叶变换--原则傅里叶变换的解释

□ 结论:

- 窗口宽度与短时傅里叶变换特征之间的关系
 - 用窄窗可得到好的时间辨别率
 - 用宽窗能够得到好的频率辨别率。
 - 但因为采用窗的目的是要限制分析的时间以使其波形特征没有明显变化，因而要折衷考虑。

4.2.3 短时傅立叶变换--滤波器的解释一

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} [x(m)e^{-j\omega m}]w(n-m)$$

- $w(n)$ -----一种滤波器的单位函数响应
- $X_n(e^{j\omega})$ -----该滤波器的输出
- $x(n)$ -----滤波器的输入
- 过程:调制+滤波

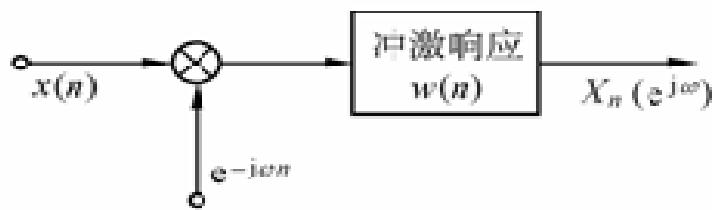


图4-4 短时傅里叶变换滤波器解释的第一种形式(a)复数运算

4.2.3 短时傅立叶变换--滤波器的解释二

$$X_n(e^{j\omega}) = |X_n(e^{j\omega})| e^{j\theta_n(\omega)} = a_n(\omega) - jb_n(\omega)$$

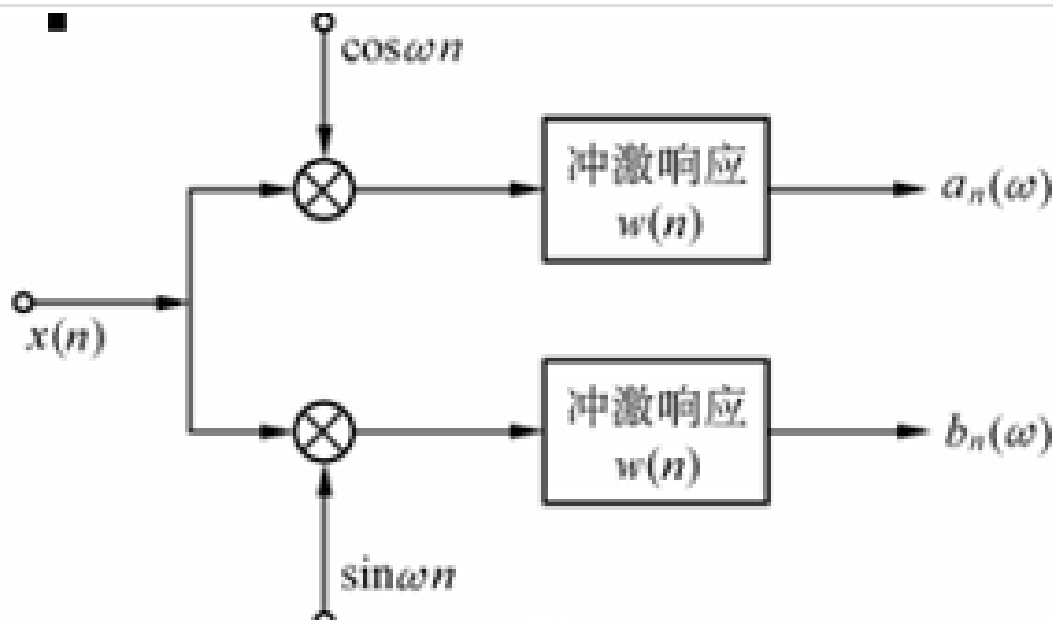


图4-4 短时傅里叶变换滤波器解释的第一种形式 (b)只有实数运算

4.2.3 短时傅立叶变换--滤波器的解释三

□ 令 $m = n - m'$

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} w(m')x(n-m')e^{-j\omega(n-m')} = e^{-j\omega n} \sum_{m=-\infty}^{\infty} w(m')x(n-m')e^{j\omega m'}$$

$$\text{令: } \tilde{X}_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} w(m')x(n-m')e^{j\omega m'}$$

$$\text{所以: } X_n(e^{j\omega}) = e^{-j\omega n} \bullet \tilde{X}_n(e^{j\omega})$$

4.2.3 短时傅立叶变换--滤波器的解释

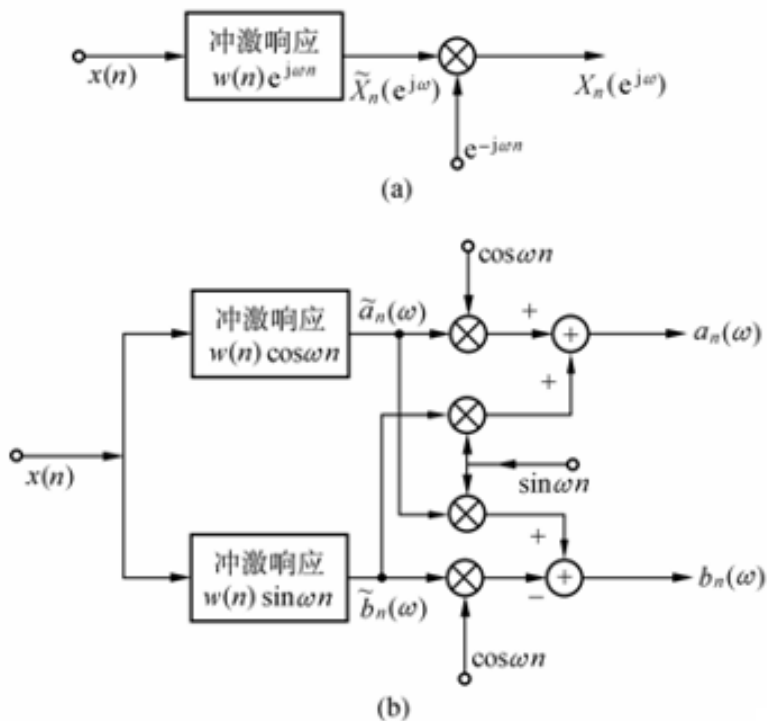


图 4-5 短时傅里叶变换滤波器解释的另一种形式
(a)复数运算(b)只有实数运算

- $W(e^{j\omega})$ 为窄带低通滤波器。第一种形式为低通滤波器；
- 因为第二种形式中的滤波器单位函数响应为 $w(n)(e^{j\omega n})$ ，所以它为带通滤波器。

4.2.3 短时傅立叶变换--滤波器的解释

$$|X_n(e^{j\omega})| = [a_n^2(\omega) + b_n^2(\omega)]^{1/2} = |\tilde{X}_n(e^{j\omega})| \cdot |e^{-j\omega n}|$$

$$|\tilde{X}_n(e^{j\omega})| = [\tilde{a}_n^2(\omega) + \tilde{b}_n^2(\omega)]^{1/2}$$

- 假如将 $w(n)$ 的滤波运算除外，短时傅里叶变换实际上是对信号的幅度调制。
 - 第一种形式是在输入端进行调制， $x(n)$ 乘以 $e^{-j\omega n}$ 相当于将 $x(n)$ 的频谱从 ω 移到零频处；而 $w(n)$ (直角窗或海明窗等)为窄带低通滤波器。
 - 后一种形式是在输出端进行调制，此时先对信号进行带通滤波，滤波器的单位函数响应为 $w(n)e^{-j\omega n}$ ，而调制后输出的是中心频率为 ω 的短时谱。🔥

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/626133151142010211>