STFT

Short Time Fourier Transform

تبدیل فوریه زمان کوتاه

مرجع: **فصل ۶ کتاب Advanced Topics in Signal Processing**

خروجی STFT : زمان و فرکانس

نمودار زمان فرکانس

آنالیز زمان فرکانس

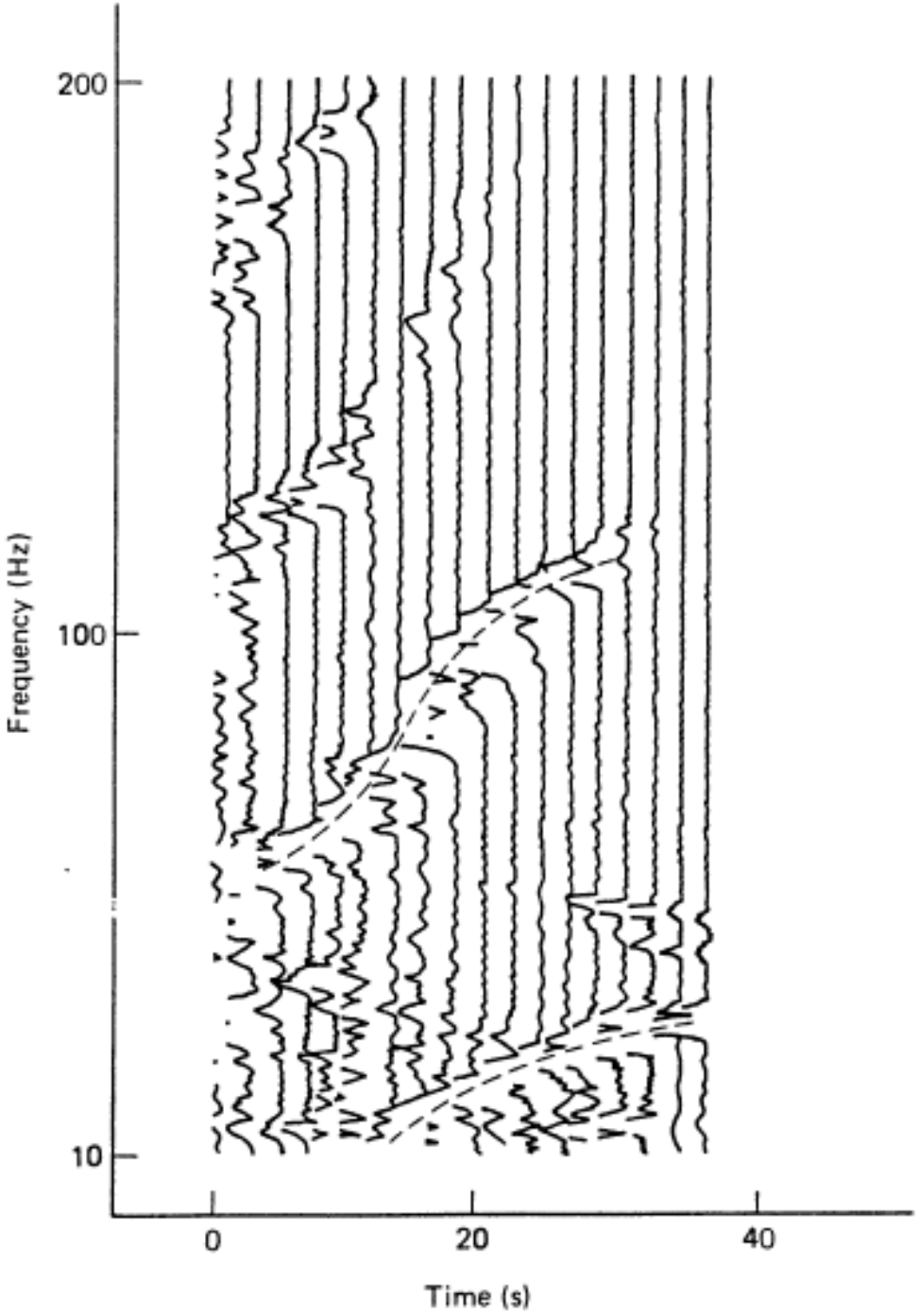


Figure 6.1 Consecutive Fourier transform magnitudes of an acoustic recording of a moving helicopter. Each plot contains a gap whenever the Fourier transform magnitude exceeds a threshold. By following such gaps from plot to plot, we can track the frequencies of highest energy. The frequency tracks of the two highest-energy harmonics are indicated by dashed lines.

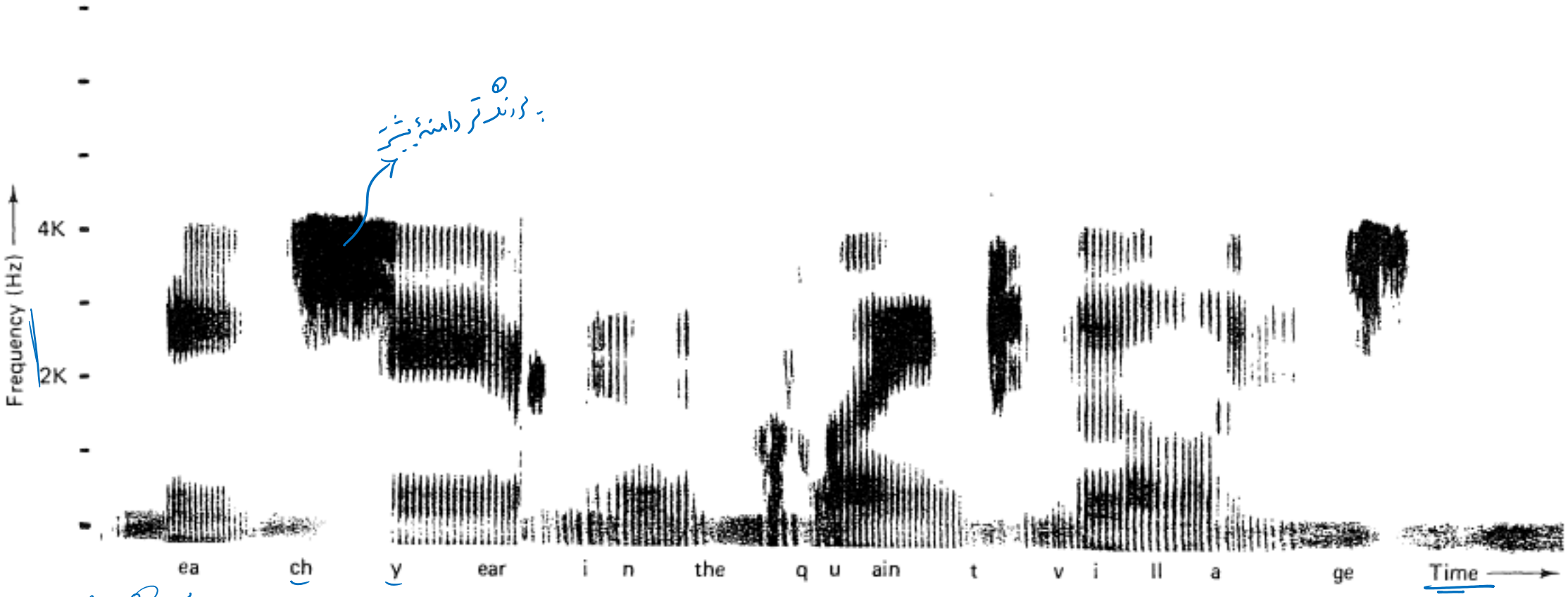


Figure 6.2 The spectrogram for the speech utterance "Each year in the quaint village."

Digital spectrogram: Discrete STFT

6.1 تبدیل فوریهٔ زمان کوتاه یک دنباله :

6.1.1 : از دیدگاه تبدیل فوریه :

تبدیل فوریهٔ زمان گسسته DTFT : $X(\omega) = \sum_{n=-\infty}^{+\infty} x(n) e^{-j\omega n}$

ω : متغیر پیوستهٔ فرکانس

اگر تبدیل فوریه را برای بازه های زمانی مختلف از $x(n)$ به دست آوریم به STFT می رسیم.

این بازه ها با ضرب $x(n)$ در دنبالهٔ شیفت یافتهٔ $w(n_0 - n)$ به دست می آیند.

رابطهٔ تبدیل فوریهٔ زمان کوتاه برای زمان n_0 : $X(n_0, \omega) = \sum_{n=-\infty}^{+\infty} \underbrace{x(n) w(n_0 - n)} e^{-j\omega n}$

$w(n)$: پنجره یا فیلتر آنالیز

بخش زمان کوتاه $x(n)$: $f_{n_0}(n) = x(n) w(n_0 - n)$

short time section

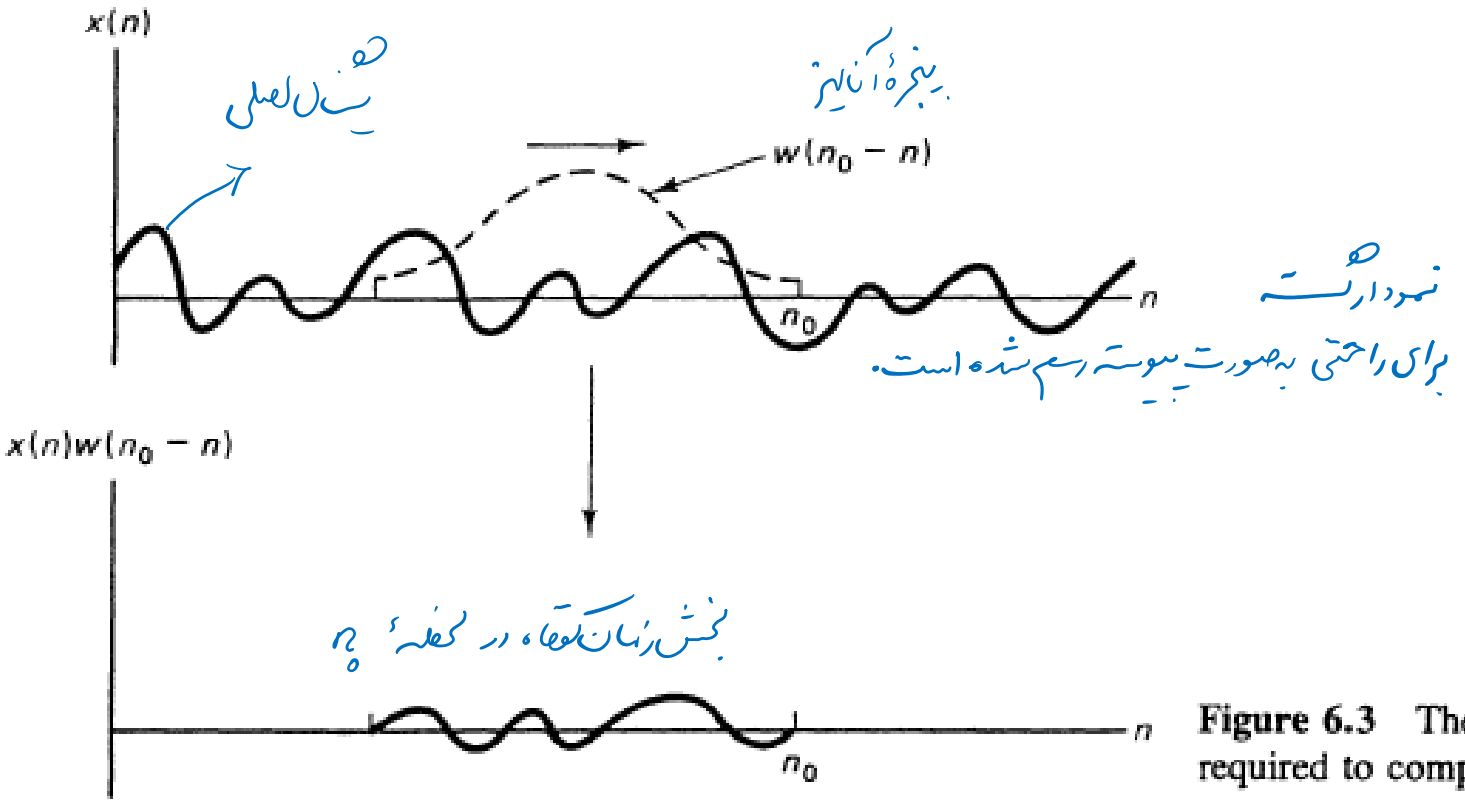


Figure 6.3 The series of operations required to compute a short-time section.

با انتقال پنجره برای تمام $n = n_0$ ها، مقدار $X(n_0, \omega)$ را به دست می‌آوریم. به این ترتیب STFT می‌شود:

STFT زمان گسسته :

Discrete Time STFT

$$X(n,\omega) = \sum_{m=-\infty}^{+\infty} \underbrace{x(m)\,w(n-m)}_{f_n(m)}\, e^{-j\omega m}$$

یکی از مشخصه‌های تبدیل STFT، پنجرهٔ آنالیز $w(n)$ است:

هرچه طول پنجره نسبت به طول سیگنال کوتاه‌تر باشد، دقت زمانی STFT بیشتر خواهد بود

هرچه طول پنجره بزرگتر انتخاب شود به تبدیل فوریهٔ معمول نزدیکتر می‌شویم.

بنابراین: طول پنجره نسبت به طول سیگنال باید کوتاه انتخاب شود مانند 256 در مقابل 30,000 مثال اسلاید.

مثال از پنجره آنالیز:

پنجرهٔ همینگ $w(n)$

این پنجره به طور معمول برای پردازش صحبت مورد استفاده قرار می‌گیرد.

طول پنجره: 256 نقطه

طول سیگنال: صحبت 3sec

نرخ نمونه برداری 10 kHz

تعداد نمونه ها: 30,000 نقطه

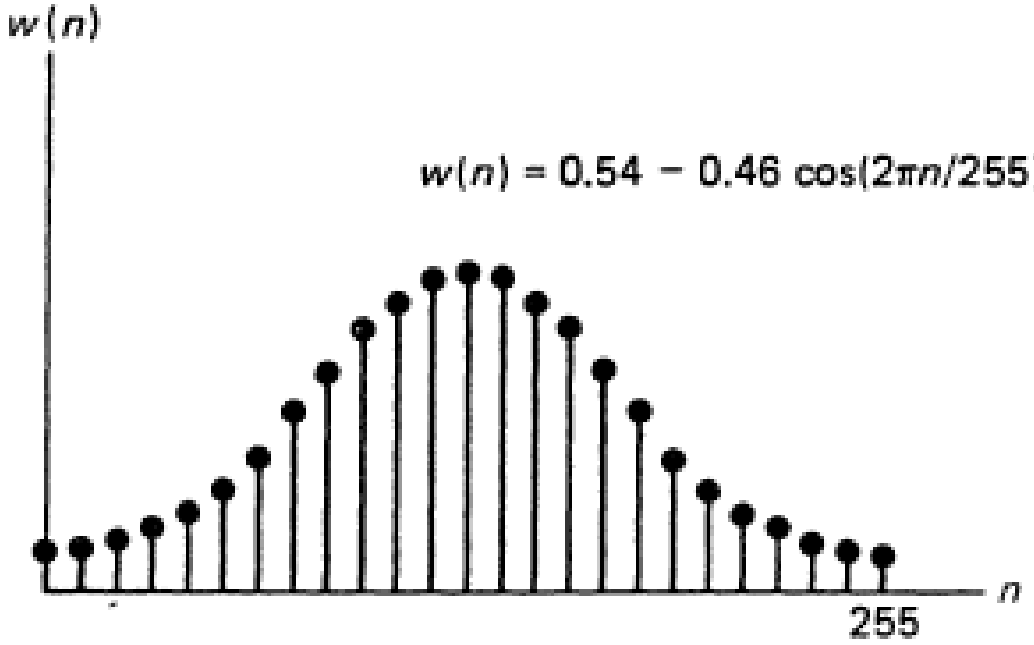


Figure 6.4 A 256-point Hamming window.

Discrete STFT

همانطور که DFT از DTFT به دست می‌آید. Discrete STFT نیز از تعریف Discrete Time STFT به دست می‌آید و در پردازش دیجیتال از Discrete STFT استفاده می‌شود.

۱- یک پریود
۲- نمونه‌برداری
} DTFT از روی DFT :

$$X(k) = X(\omega)\Big|_{\omega=\frac{2\pi k}{N}} R_N(k)$$

$R_N(k) = u(k) - u(k-N)$ دنباله مستطیلی N نقطه‌ای

به همین ترتیب Discrete STFT نمونه‌برداری شده یک پریود از Discrete Time STFT روی فرکانس است:

Discrete STFT رابطه‌ی

$$X(n,k) = X(n,\omega)\Big|_{\omega=\frac{2\pi k}{N}} R_N(k)$$

از روی OT STFT

Discrete STFT:

$$X(n,k) = \sum_{m=-\infty}^{+\infty} \underbrace{x(m)w(n-m)}_{x_n(m)} e^{-j\frac{2\pi km}{N}} R_N(k)$$

کاهش نرخ نمونه برداری در خروجی Discrete STFT :

$X(nL, k)$ Decimation

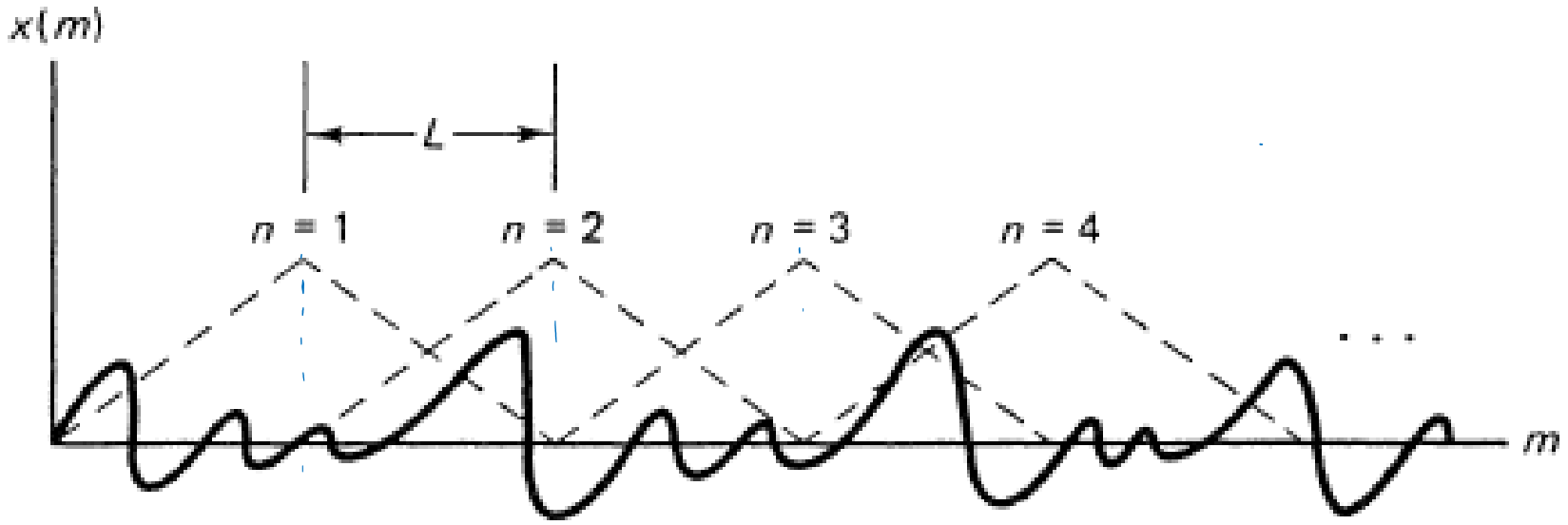


Figure 6.5 The analysis window positions used in computing $X(nL, k)$.

Discrete STFT . تبدیل DFT از دنباله‌های $f_n(m)$ است .

و با کاهش نرخ نمونه برداری ، به تبدیل DFT از دنباله‌های $f_{nL}(m)$ تبدیل می‌شود .

6.1.2 : تعریف STFT از دیدگاه فیلترینگ :

$$X(n,\omega_0)=\sum_{m=-\infty}^{+\infty} x(m)w(n-m)e^{-j\omega_0 m}$$

کانولوشن

$$\boxed{X(n,\omega_0)=\sum_{m=-\infty}^{+\infty}\left[x(m)e^{-j\omega_0 m}\right]w(n-m)} = \left[x(n)e^{-j\omega_0 n}\right] * w(n)$$

مدولاسیون با نمایی مختلط

پاسخ ضربه فیلتر پایین گذر

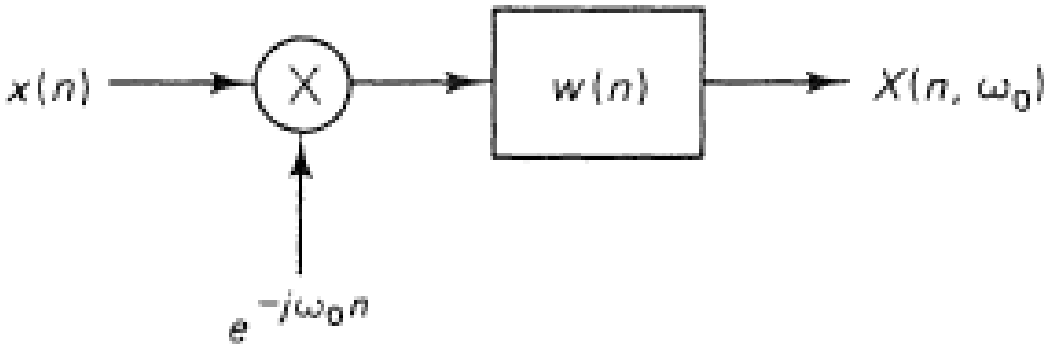


Figure 6.6 Filtering view of STFT analysis at frequency ω_0. Complex exponential modulation is followed by a lowpass filter.

نوع دوم :

$$X(n,\omega_0) = \underbrace{e^{-j\omega_0 n}}_{\text{مدولاسیون تمتع}} \underbrace{\left[x(n) * w(n) e^{j\omega_0 n}\right]}_{\text{فیلتر شدن}}$$

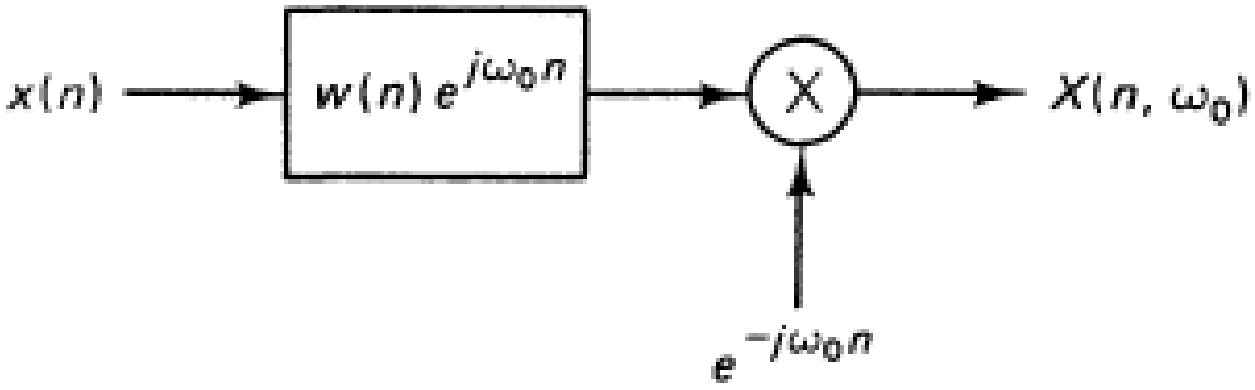


Figure 6.7 Alternative filtering view of STFT analysis at frequency ω_0. A bandpass filter is followed by complex exponential modulation.

همین دیدگاه فیلترینگ را می‌توان به Discrete STFT نیز اعمال کرد.

$X(n,k)$

بانک فیلتری Filter Bank

$$\omega_k = \frac{2\pi k}{N}$$

$k = 0, 1, \ldots, N-1$

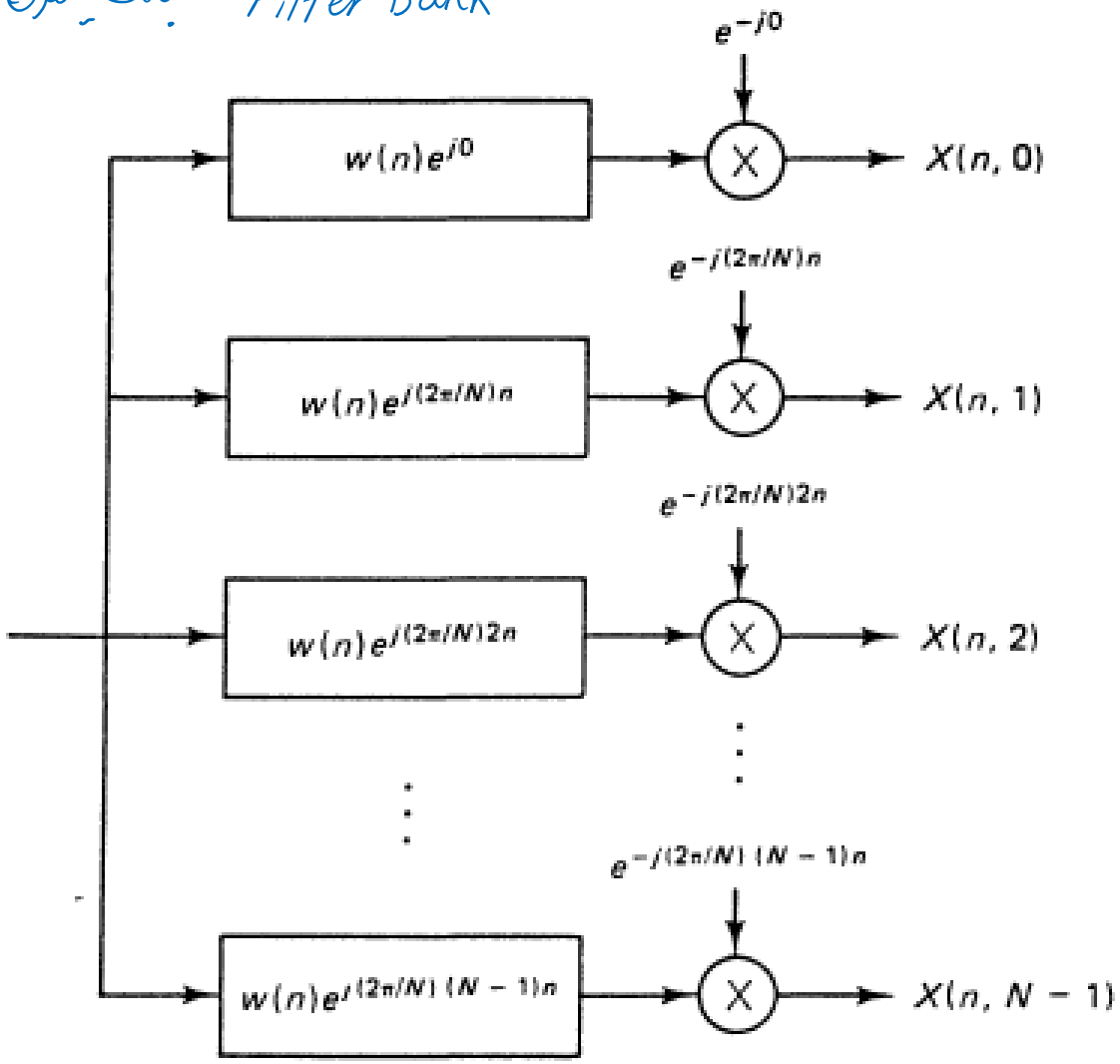


Figure 6.8 The discrete STFT as the output of a filter bank consisting of bandpass filters.

6.1.3 ویژگی‌های STFT DT :

TABLE 6.1 PROPERTIES OF THE DISCRETE-TIME STFT BASED ON THE FOURIER TRANSFORM INTERPRETATION

پریودیک	Property 1:	$X(n, \omega) = X(n, \omega + 2\pi)$
تقارن هرمیتی نسبت به ω	Property 2:	$x(n)$ real $\longleftrightarrow X(n, \omega) = X^*(n, -\omega)$
زوج نسبت به ω	Property 3:	$x(n)$ real $\longleftrightarrow \lvert X(n, \omega)\rvert = \lvert X(n, -\omega)\rvert$
فرد " " ω	Property 4:	$x(n)$ real $\longleftrightarrow \arg[X(n, \omega)] = -\arg[X(n, -\omega)]$
شیفت در حوزه n	Property 5:	$x(n - n_0) \longleftrightarrow e^{-j\omega n_0} X(n - n_0, \omega)$

ویژگی‌های Discrete STFT :

$$\omega = \frac{2\pi k}{N}$$

TABLE 6.2 PROPERTIES OF THE DISCRETE STFT

Property 1:	$X(n,k)$ is zero outside $0 \le k < N$
Property 2:	$x(n)$ real $\longleftrightarrow X(n,k) = X^*(n, N-k)$ for $0 < k < N$
Property 3:	$x(n)$ real $\longleftrightarrow \lvert X(n,k)\rvert = \lvert X(n, N-k)\rvert$ for $0 < k < N$
Property 4:	$x(n)$ real $\longleftrightarrow \arg[X(n,k)] = -\arg[X(n, N-k)]$ for $0 < k < N$
Property 5:	$x(n-n_0) \longleftrightarrow e^{-j(2\pi n_0 k/N)} X(n-n_0, k)$

دنباله حقیقی

ویژگی‌ها از دیدگاه فیلترینگ:

۱- در $\omega_0 = 0$ ، $X(n,\omega)$ به کانولوشن تبدیل می‌شود.

۲- اگر $x(n)$ طول N داشته باشد و طول $w(n)$، M باشد طول $X(n,\omega)$ در امتداد محور n، برابر با $n+M-1$ خواهد بود.

۳- پهنای باند دنباله $X(n,\omega_0)$ از پهنای باند $w(n)$ کوچکتر یا با آن برابر باشد. (از دیدگاه فیلترینگ خروجی فیلتر $w(n)$ است)

TABLE 6.3 PROPERTIES OF DISCRETE-TIME STFT BASED ON THE FILTERING INTERPRETATION

$\omega_0 = 0 \rightarrow$ Property 1:	$X(n, 0) = x(n) * w(n)$	کانولوشن
Property 2:	$x(n)$ length N, $w(n)$ length M, $X(n, \omega)$ length $N + M - 1$ along n	
Property 3:	Bandwidth of sequence $X(n, \omega_0) \leq$ Bandwidth of $w(n)$	
Property 4:	The sequence $X(n, \omega_0)$ has spectrum centered at the origin	
Property 5:	$x(n)$ causal, $w(n)$ causal, $X(n, \omega)$ causal in time	

۴- طیف دنباله $X(n,\omega_0)$ به مرکزیت مبدأ است. (اثبات از مدولاسیون)

۵- $x(n)$ علی، $w(n)$ علی، با توجه به کانولوشن $X(n,\omega)$ هم علی خواهد بود.

6.1.4 معکوس پذیری : Invertibility

هدف : از STFT به $x[n]$ برسیم.

همواره از روی discrete time STFT می‌توان به $x[n]$ رسید ولی discrete STFT نیاز به شرایط خاصی روی نرخ نمونه برداری دارد تا معکوس پذیر باشد.

معکوس پذیری DT STFT :

$X(n,\omega)$ در یک زمان مشخص : به تبدیل فوریه تبدیل می‌شود که $X(n,\omega)$ تبدیل فوریه دنباله $f_n(m) = x(m)\,w(n-m)$ است.

اگر $n=m$ برابر باشد تبدیل معکوس $x(n)w(0)$ خواهد شد.

اگر $w(0) \neq 0$ باشد می‌توان $x(n)$ را از رابطه زیر به دست آورد :

$$x(n) = \frac{1}{2\pi w(0)} \int_{2\pi} X(n,\omega)\, e^{j\omega n}\, d\omega$$

رابطه مشترک برای : discrete time STFT

در مقابل، discrete STFT همواره معکوس پذیر نیست.

برای مثال: پنجرهٔ آنالیز $w(n)$ با پهنای باند B را در نظر بگیرید. Fig 6.9.

← اگر $B < \frac{2\pi}{N}$ پهنای باند فیلتر آنالیز از نرخ نمونه برداری کمتر است و برخی مؤلفه‌های فرکانسی $x(n)$ از فیلتر عبور نمی‌کنند و بنابراین قابل بازیابی نخواهند بود.

← در حالتی که $x(n)$ در زمان با فاکتور L decimate شود و طول پنجرهٔ آنالیز (N_w) از L کوچکتر باشد، STFT معکوس پذیر نخواهد بود.

حالتِ اوّل :

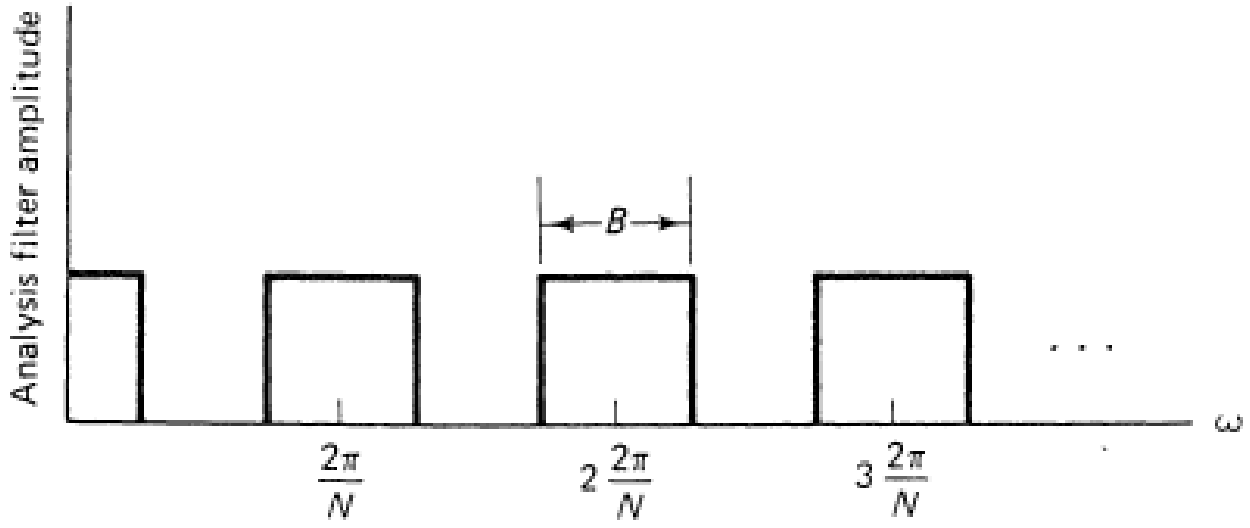


Figure 6.9 Undersampled STFT when the frequency sampling interval $2\pi/N$ is greater than the analysis filter bandwidth B.

حالت دوم:

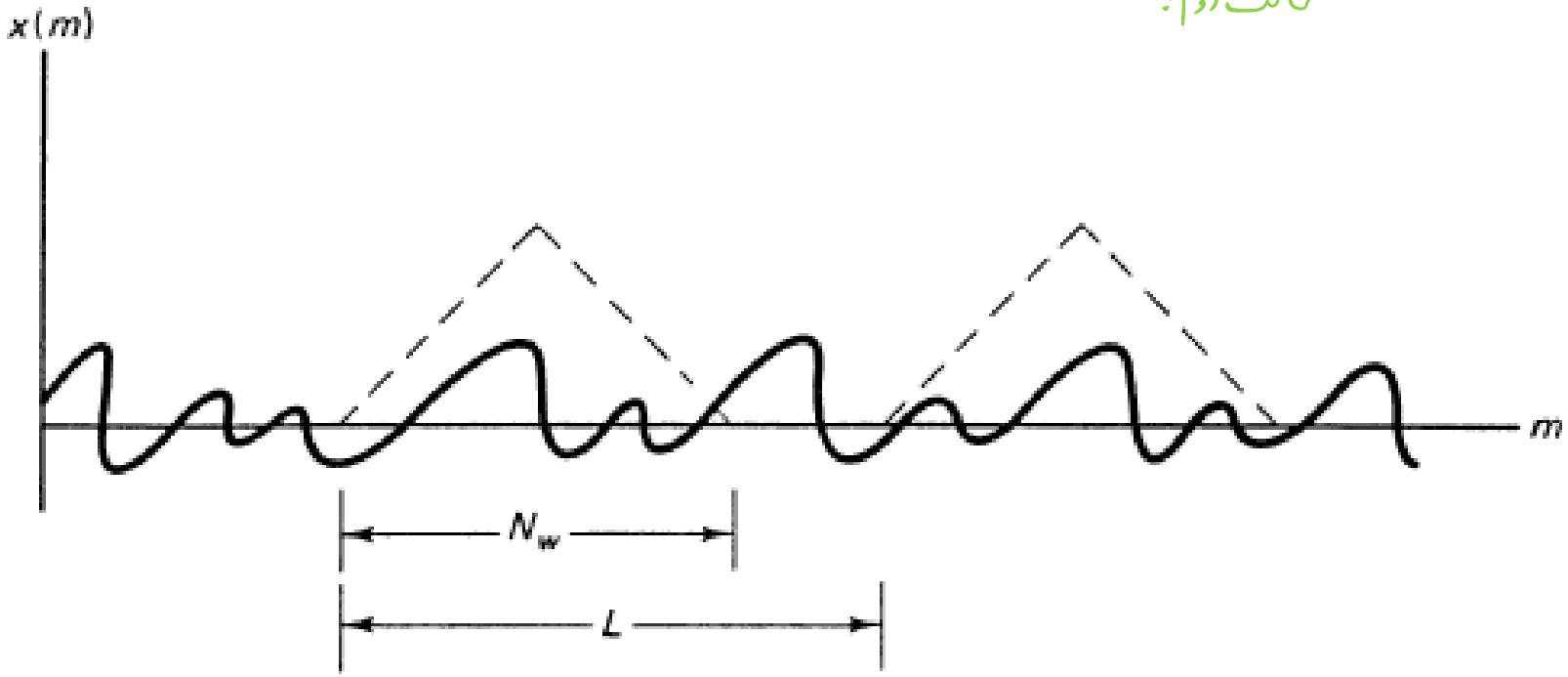


Figure 6.10 Underdecimated STFT when the decimation factor L is larger than the length N_w of the analysis window.

شرط های مکوس پذیری : Discrete STFT

$$\begin{cases} N_W > L & \text{decimation نرخ} \\ B > \frac{2\pi}{N} & \end{cases}$$

طول پنجره آنالیز

پهنای باند فیلتر آنالیز

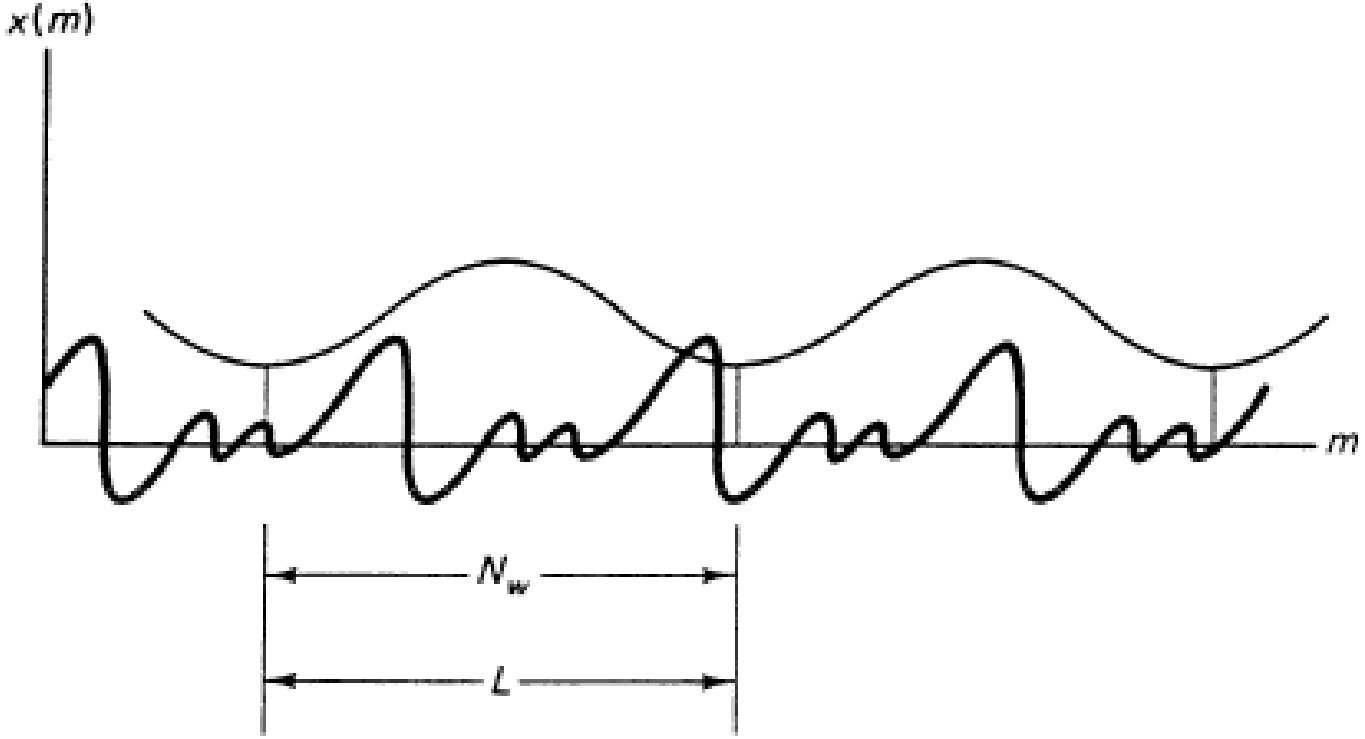
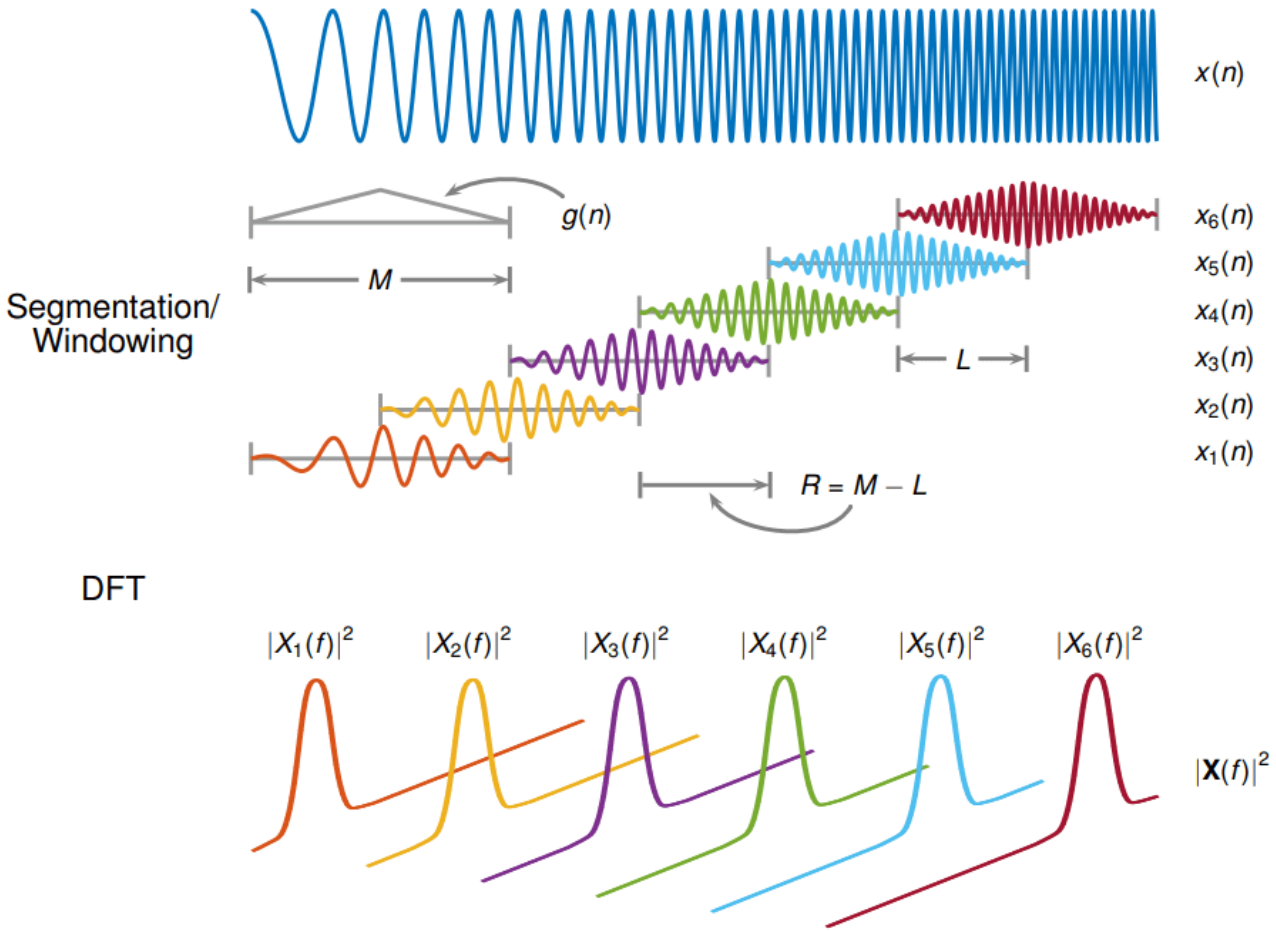


Figure 6.11 The decimation factor L equals the analysis window length N_w. The window positions are adjacent without any overlap or missed regions.

6.2 تاخیر

پیاده سازی در متلب

- `stft(x, fs, ‘window’, hamming(256))`



6.2 آنالیز فوریهٔ زمان کوتاه short Time Fourier Analysis

discrete STFT
۱- انتخاب پنجره مناسب
۲- الگوریتم‌های سریع

6.2.1 انتخاب پنجرهٔ آنالیز :

ویژگی‌های STFT به پنجرهٔ آنالیز انتخابی حساس است.

برقراری توازن مناسب بین طول پنجرهٔ زیاد برای وضوح فرکانسی بالا، و طول پنجرهٔ کوتاه برای زمان کوتاه بودن آنالیز

رابطهٔ STFT از دید تبدیل فوریه: تبدیل فوریهٔ $f_n(m) = x(m)\,w(n-m)$ ، STFT است.

$$\begin{cases} x(m) \xleftrightarrow{F} X(\omega) \\ w(n-m) \xleftrightarrow{F} W(-\omega)\,e^{-j\omega n} \end{cases}$$

m متغیر مستقل

ضرب در زمان
کانولوشن در فرکانس

$$X(n,\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\theta)\,e^{j\theta n}\,X(\omega+\theta)\,d\theta$$

$-\theta \to \theta$

تغییرات فرکانسی STFT برای هر زمان ثابت n را می‌توان به صورت منحنی هموار شده (smooth شده) تبدیل فوریهٔ سیگنال در نظر گرفت.

- وضوح فرکانسی بهتر به طول پنجرهٔ بزرگتری نیاز دارد.

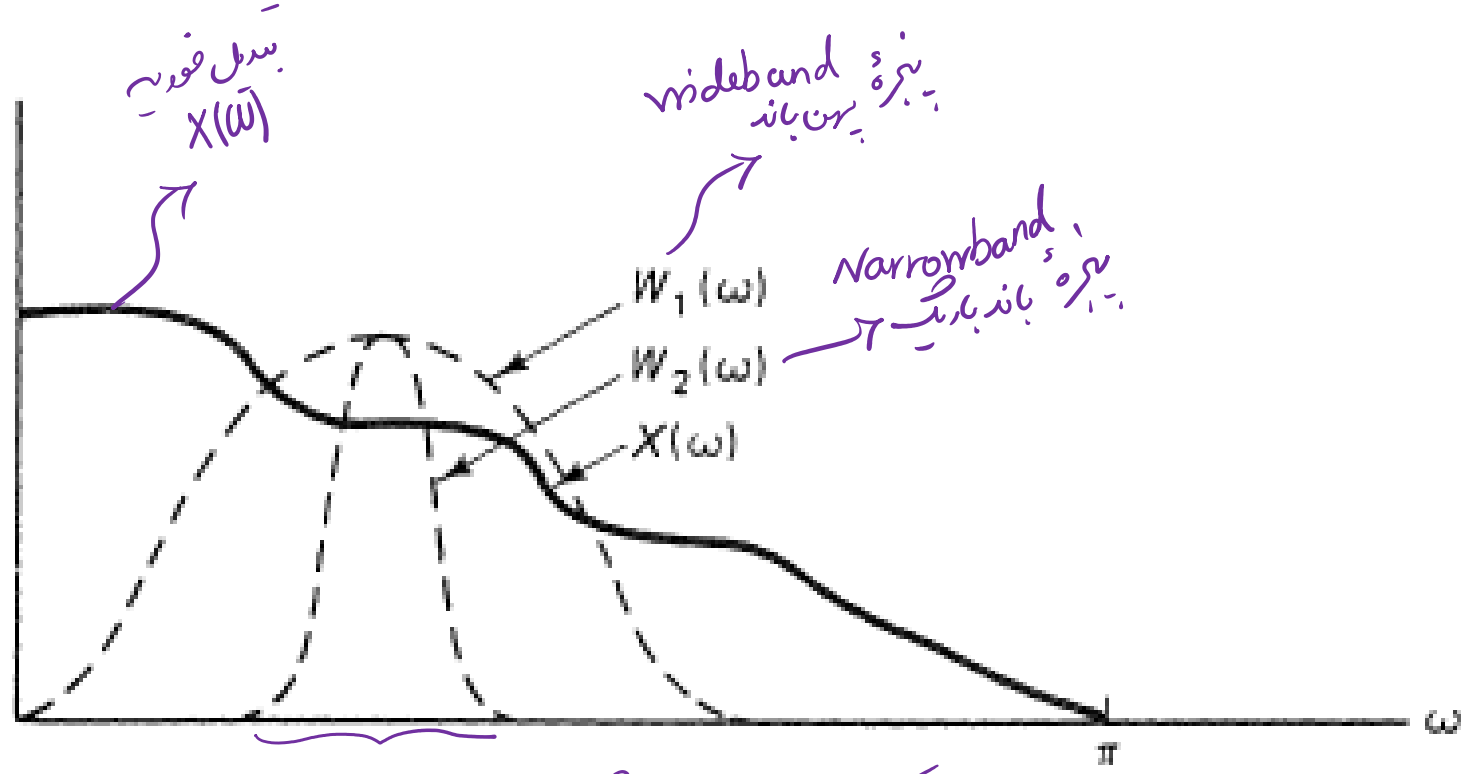


Figure 6.12 The window Fourier transform as a smoothing function for the Fourier transform of the underlying signal. A narrowband and a wideband window are illustrated.

هرچه $W(\omega)$ به ضربه نزدیکتر باشد یعنی باند باریک تر باشد. $X(n,\omega)$ وضوح فرکانسی بهتری خواهد داشت. بنابراین برای یک پنجرهٔ مشخص، وضوح فرکانسی به صورت معکوس با طول پنجره تغییر می‌کند.

سیگنال chirp

از $\frac{\pi}{4}$ به $\frac{\pi}{8}$

مثالی از مسأله (trade-off)

بین وضوح فرکانسی و طول پنجره

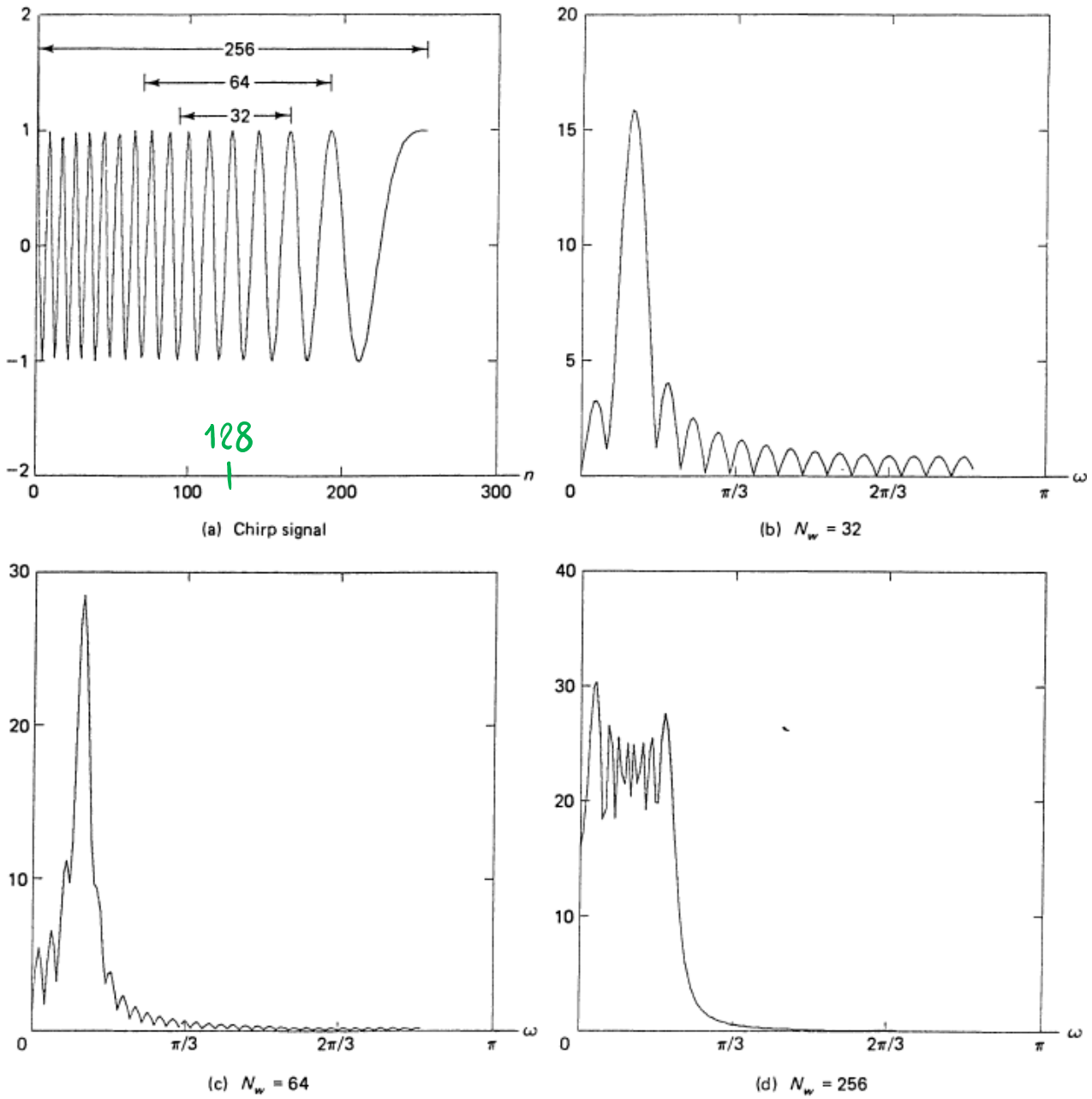


Figure 6.13 Effect of the length of the analysis window on the Fourier transform of a linearly frequency-modulated sinusoid of length 256 samples whose frequency increases from $\pi/4$ to $\pi/8$. The Fourier transform is taken for sections centered around the 128th sample (frequency $3\pi/16$). Transforms are shown for window lengths 32, 64, and 256.

6.2.2 پیاده‌سازی تبدیل فوریهٔ زمان کوتاه با تبدیل فوریهٔ سریع FFT :

همانطور که پیش از این گفته شد Discrete STFT را می‌توان به صورت DFT بخش‌های زمان کوتاه که متناظر با شیفت‌های پنجرهٔ آنالیز می‌باشند، درنظر گرفت.

بنابراین راه حل اول برای آنالیز STFT، پیاده‌سازی با DFT است که $X(n,\omega)$ را برای هر مقدار n مشخص، می‌سازد.

هزینهٔ محاسباتی DFT : N^2 بنابراین برای محاسبهٔ موثرتر از FFT با $N=2^\nu$ استفاده می‌کنیم.

فرض کنید پنجرهٔ آنالیز طول N_W دارد :

$$X(n,k)=\sum_{m=-\infty}^{+\infty}\underbrace{w(n-m)\,x(m)}_{f_n(m)}\,e^{-j\frac{2\pi mk}{N}}=\sum_{m=-\infty}^{+\infty}f_n(m)\,e^{-j\frac{2\pi mk}{N}}$$

بخش‌های زمان کوتاه.

حالت اول: اگر N (تعداد نمونه‌های فرکانسی) از N_W (طول پنجره) بزرگتر باشد:

در این حالت هر یک از بخش‌های زمان کوتاه $f_n(m)$ برای یک n مشخص زنباله با طولی که دوری است که طول آن از N کوچکتر و یا مساوی است $N \geq N_W$. در این حالت $X(n,k)$ را می‌توان با FFT، N نقطه‌ای از $f_n(m)$ که تعداد لازم صفر به انتهای آن اضافه شده است به دست آورد.

Zero-Padding

حالت دوم: اگر N کوچکتر از N_W باشد:

از آنجا که $X(n,\omega)$ تبدیل فوریه $f_n(m)$ است، نمونه برداری فرکانسی یکنواخت با گام نمونه برداری $\frac{2\pi}{N}$ در حوزه زمان متناظر با کانولوشن $f_n(m)$ با قطار ضربه پریودیک با پریود N است که با رابطه زیر قابل بیان است:

$$\tilde{f}_n(m) = \sum_{k=-\infty}^{+\infty} f_n(m+kN)$$

$\tilde{f}_n(m)$ دنباله پریودیک شده $f_n(m)$ با پریود N

به این دنباله، اصطلاحاً نسخه time-aliased $f_n(m)$ گفته می‌شود. که با تبدیل FFT، N نقطه‌ای از یک پریود آن می‌توان N نمونه discrete STFT را به دست آورد.

اگر نخواهیم $X(nL,k)$ را بسازیم، $\tilde{f}_n(m)$ را به ازای هر L امین بخش زمان گفته، به دست می‌آوریم:

$$\tilde{f}_{nL}(m) = \sum_{k=-\infty}^{+\infty} f_{nL}(m+kN)$$

6.2.3 : پیاده سازی با روش بانک فیلتری — Filter Bank Approach

- از دیدگاه پیاده سازی فیلتری STFT استفاده می کند.
- چندین پیاده سازی برای آن ارائه می کنیم.
- در برخی حالت ها این نوع پیاده سازی از FFT بهره می دارد
  - طول پنجره نامحدود.
  - برخی از نمونه های خروجی لازم باشد.

$$X(n,k) = e^{-j\omega_k n}\left[\underbrace{x(n) * w(n)e^{j\omega_k n}}_{\text{فیلترینگ میان گذر}}\right] \; ; \quad \omega_k = \frac{2\pi k}{N} \; ; \; k = 0, 1, \dots, N-1$$

Fig 6.8

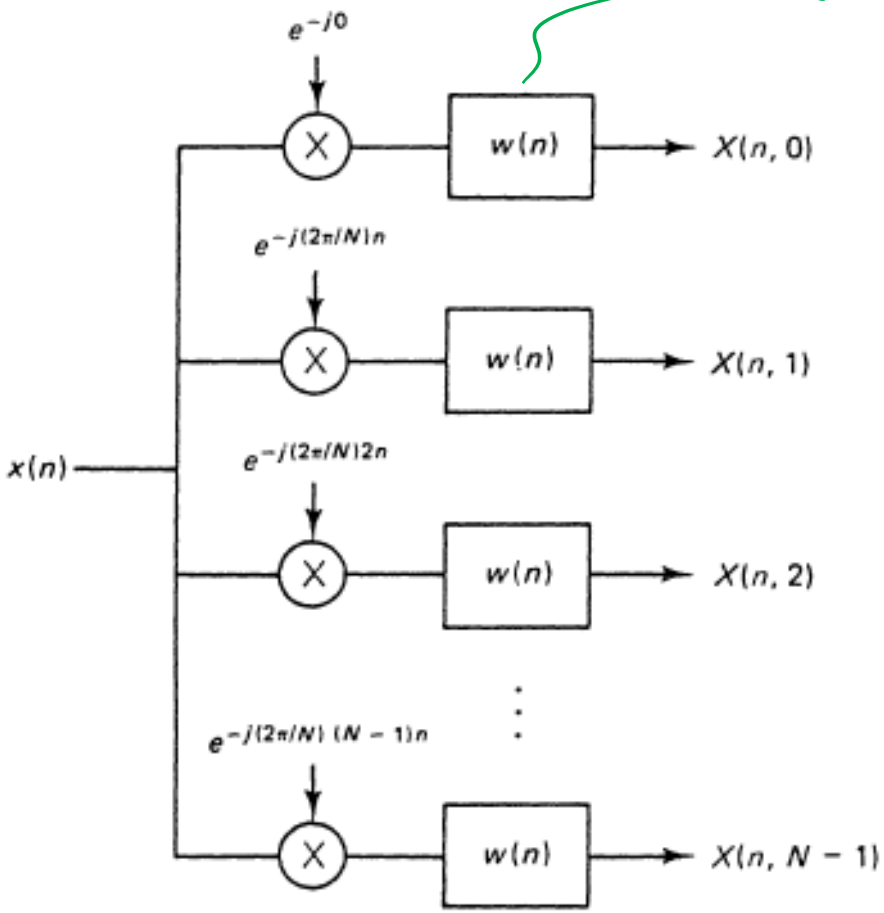
اگر خروجی هر فیلتر در حوزه زمان به ضریب L، decimate شود، $X(nL, k)$ به دست می آید.

- پیاده سازی دیگری نیز با جایگزینی فیلتر میان گذر با فیلتر پایین گذر حاصل می شود:

$$X(n,k) = \underbrace{x(n)e^{-j\omega_k n} * w(n)}_{\text{فیلترینگ پایین گذر}} \; ; \quad \omega_k = \frac{2\pi k}{N} \; , \; k = 0, 1, \dots, N-1$$

Fig 6.14

فیلتر پایین گذر



خروجی هر شاخه یکی از نمونه های discrete STFT را نتیجه می دهد.

Figure 6.14 The discrete STFT as the output of a filter bank consisting of lowpass filters.

- در پیاده‌سازی با فیلترهای میان‌گذر و پایین‌گذر به مدولاسیون‌های مختلف (ضرب در $e^{-j\omega_k n}$) نیاز است.

می‌توان این مدولاسیون را با عملیات حقیقی پیاده‌سازی کرد:

پایین‌گذر

$$X(n,k) = x(n)e^{-j\omega_k n} * w(n) = [x(n)\cos\omega_k n * w(n)] + j[x(n)\sin(\omega_k n) * w(n)]$$

$$= X_r(n,k) + jX_i(n,k)$$

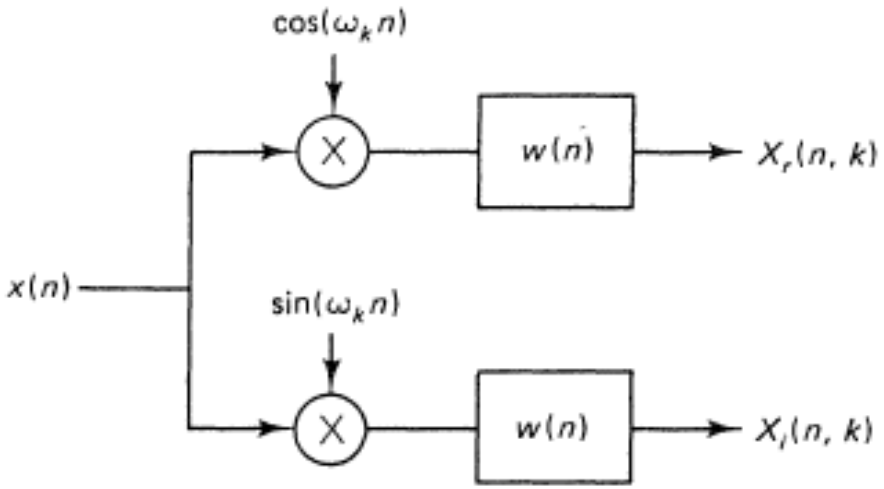


Figure 6.15 Implementation with real operations of each branch of the lowpass filter bank for STFT analysis.

$$X(n,k) = e^{-j\omega_k n}\left[x(n) * w(n)e^{j\omega_k n}\right] = (\cos\omega_k n - j\sin(\omega_k n))\left[x(n) * w(n)\{\cos\omega_k n + j\sin\omega_k n\}\right]$$

$$= (\cos\omega_k n - j\sin\omega_k n)\left[\underbrace{x(n) * w(n)\cos\omega_k n}_{\hat{X}_r(n,k)} + j\underbrace{x(n) * w(n)\sin\omega_k n}_{\hat{X}_i(n,k)}\right]$$

میان‌بر :

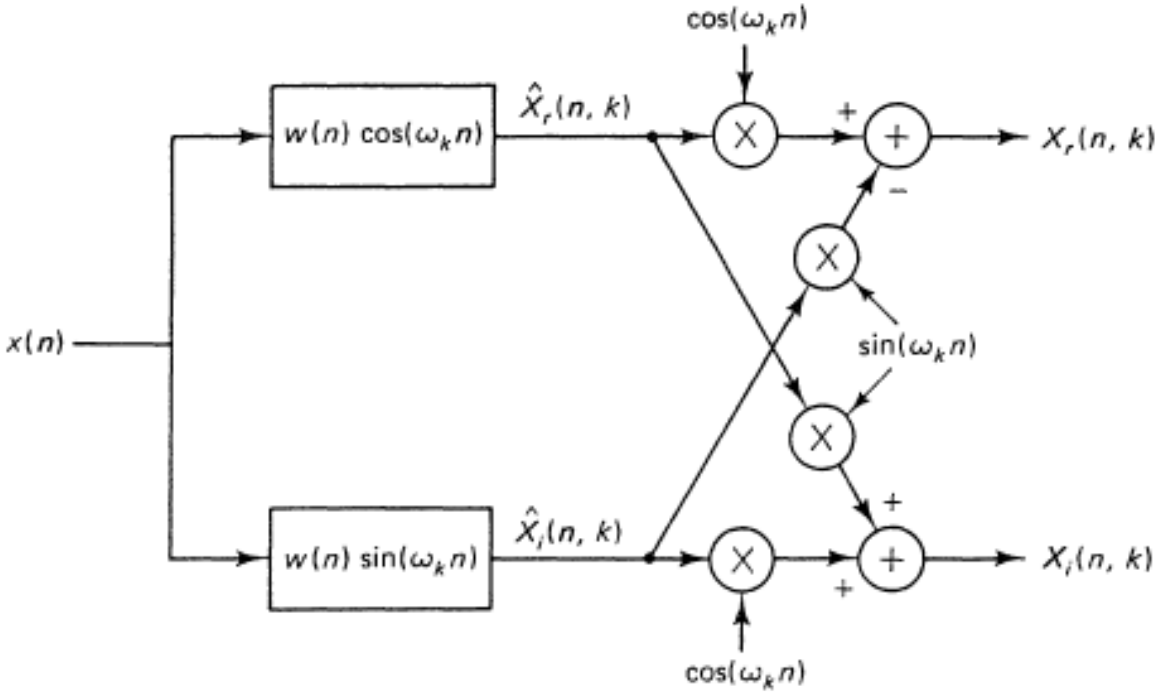


Figure 6.16 Implementation with real operations of each branch of the bandpass filter bank for STFT analysis.

$$= \left[\hat{X}_r(n,k)\cos\omega_k n + \hat{X}_i(n,k)\sin\omega_k n\right] + j\left[\hat{X}_i(n,k)\cos\omega_k n - \hat{X}_r(n,k)\sin\omega_k n\right]$$

$$= X_r(n,k) + jX_i(n,k)$$

نکته: درصورت حذف عبارت مدولاسیون در رابطه فیلتر میان‌گذر، خروجی نمونه‌های $X(n,w)e^{jwn}$ خواهند بود که در برخی کاربردها این نمونه‌ها کافی است. برای مثال عبارت فاز خطی e^{jwn} در سیستم‌های طیفی مانند پهنای باند و پارامترهای LPC، در آنالیز زمان کوتاه هیچ اثر صحبت تأثیری ندارد. بنابراین در این قبیل کاربردها، می‌توان با حذف عبارت‌های مدولاسیون، ساختارهای پیاده‌سازی را ساده‌تر کرد.

نکته: در دو حالت فیلتر بانک برتری دارد:

① حالتی که طول پنجره آنالیز (N_w) نامحدود باشد. در این حالت می‌توان فیلترها را به صورت بازگشتی پیاده‌سازی کرد.

مثال: فرض کنید پاسخ ضربه فیلتر: $w(n) = (0.9)^n u(n)$

$y(n) = 0.9\, y(n-1) + x(n)$ پیاده‌سازی بازگشتی

که به دو عملیات برای هر نمونه نیاز دارد. درصورتی که اگر از پیاده‌سازی FFT استفاده کنیم هر نمونه به $\log N$ عملیات نیاز دارد. مثال ($\log_2 128 = 7$)

② وقتی که تنها تعداد کمی از نمونه‌های STFT را نیاز داشته باشیم.

Filter Bank: $4 \times 256 = 1024$

تعداد نمونه‌های لازم — فیلتر FIR با طول 256

FFT: $N \log N = 256 \log 256 = 2048$

6.3 : سنتز (تولید) صوت زمان کوتاه :

همواره نمی‌توان سیگنال اصلی را از discrete STFT بازسازی کرد زیرا $X(n,k)$ همواره معکوس‌پذیر نیست ولی $X(n,\omega)$ همواره معکوس‌پذیر است.

discrete time STFT

راه متداول این است که روابط سنتز برای discrete time STFT را استفاده و تعمیم گسسته آن را برای بازسازی سیگنال $x(n)$ از $X(n,k)$ به کار بریم.

دو روش متداول و کلاسیک برای سنتز زمان کوتاه:

۱- روش FBS : Filter Bank Summation

۲- روش OLA : overlap-Add

روش FBS : Filter Bank Summation :

از 6.2 معکوس پذیری → به شرط $w(0) \neq 0$

$$x(n) = \left[\frac{1}{2\pi w(0)}\right] \int_{-\pi}^{\pi} X(n,\omega) e^{j\omega n} d\omega$$

FBS : نسخه‌ی گسسته شده‌ی رابطه‌ی بالا :

$$y(n) = \left[\frac{1}{N w(0)}\right] \sum_{k=0}^{N-1} X(n,k) e^{j\frac{2\pi k}{N} n}$$

$$y(n) = \left[\frac{1}{N w(0)}\right] \sum_{k=0}^{N-1} \sum_{m=-\infty}^{+\infty} x(m) w(n-m) e^{-j\frac{2\pi k m}{N}} e^{j\frac{2\pi k n}{N}}$$

از دیدگاه فیلترینگ →

$$y(n) = \left[\frac{1}{N w(0)}\right] x(n) * \sum_{k=0}^{N-1} w(n) e^{j\frac{2\pi k}{N} n}$$

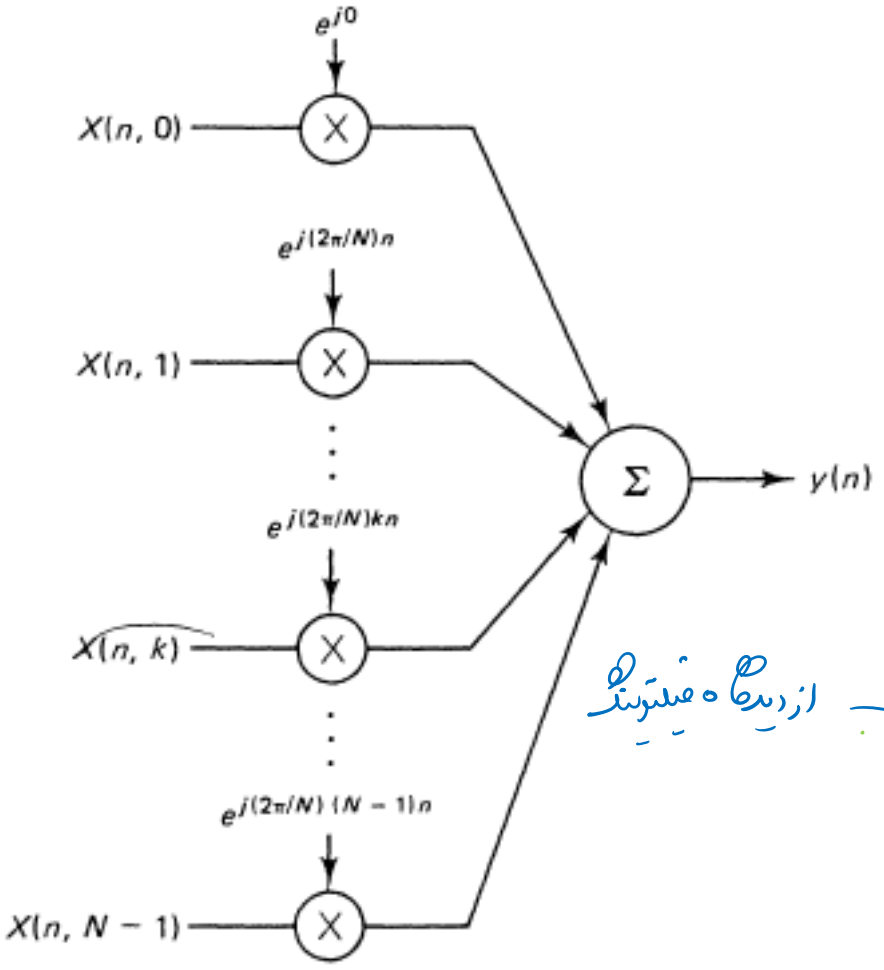


Figure 6.17 Filter bank summation procedure for signal synthesis from the discrete STFT.

$N w(0) \delta(n)$

$$y(n) = \left[\frac{1}{N w(0)}\right] x(n) * w(n) N \sum_{r=-\infty}^{+\infty} \delta(n - rN)$$

$$y(n) = x(n) \Rightarrow N w(0) \delta(n) = w(n) N \sum_{r=-\infty}^{+\infty} \delta(n - rN)$$

شرط FBS در گستره‌ی زمان

شرط FBS در حوزهٔ زمان :

$$w(n)\,N\sum_{r=-\infty}^{+\infty}\delta(n-rN) = N\,w(0)\,\delta(n)$$

← اگر $N_W < N$ باشد : این رابطه همواره برقرار است .

پس برای پنجره های با طول کوچک به شرطی که طول پنجره از تعداد نمونه های فرکانسی کوچکتر باشد FBS همواره قابل استفاده است .

اگر $N_W > N$ باشد : رابطه زمانی برقرار است که :

$$w(rN) = 0 \qquad \text{for} \qquad r = -1, 1, -2, 2, -3, 3, \ldots$$

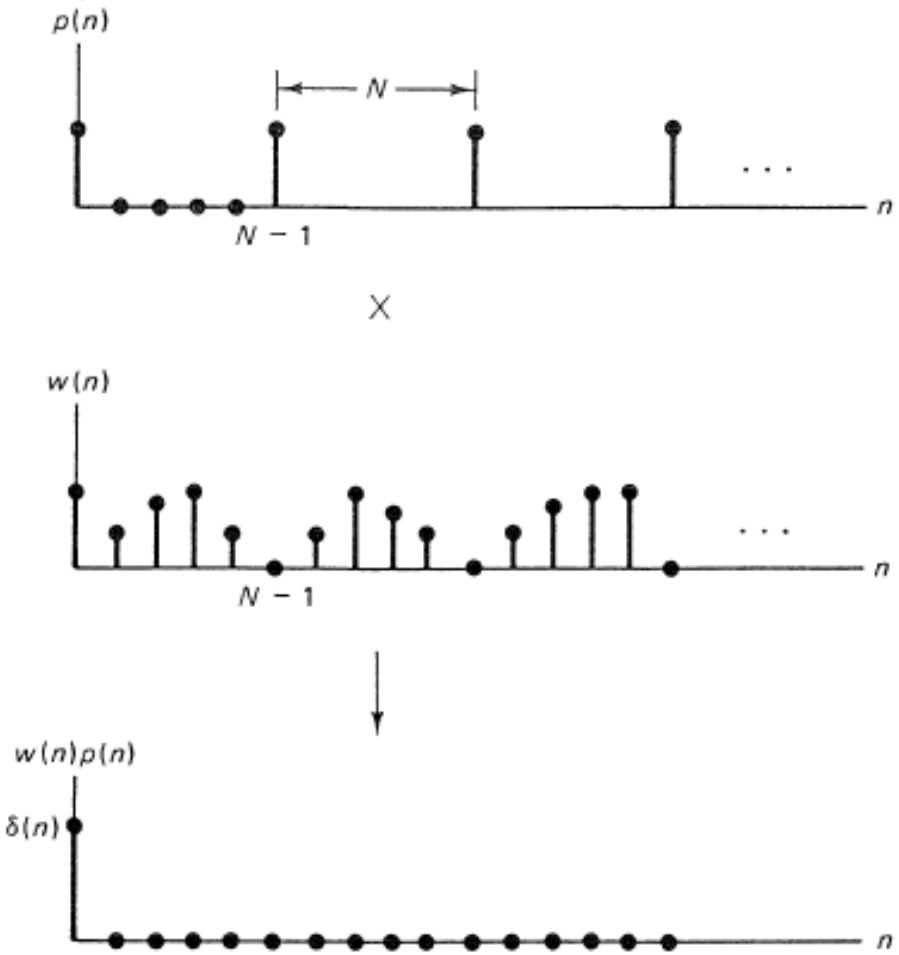
شرط FBS در حوزهٔ فرکانس :

$$\sum_{k=0}^{N-1} W\left(\omega - \frac{2\pi k}{N}\right) = N\,W(0)$$

روش FBS با $L=1$ ، به N_W عملیات برای هر نمونهٔ $x(n)$ نیاز دارد.

اگر $L \gg 1$ باشد، از آنجا که به عملیات درونیابی برای خروجی فیلتر نیاز است به N_W^2 عملیات برای هر نمونهٔ $x(n)$ نیاز است .

$$\sum_{r=-\infty}^{+\infty} \delta(n-rN)$$



بررسی شرط FBS در حوزهٔ زمان :

$$w(n) \underbrace{\sum_{r=-\infty}^{+\infty} \delta(n-rN)}_{p(n)} = w(0)\delta(n)$$

Figure 6.18 Example of an analysis window and how it satisfies the FBS constraint. The analysis window length is longer than the frequency sampling factor.

بررسی شرط FBS در حوزهٔ فرکانس :

$$\sum_{k=0}^{N-1} W\left(\omega - \frac{2k\pi}{N}\right) = N W(0)$$

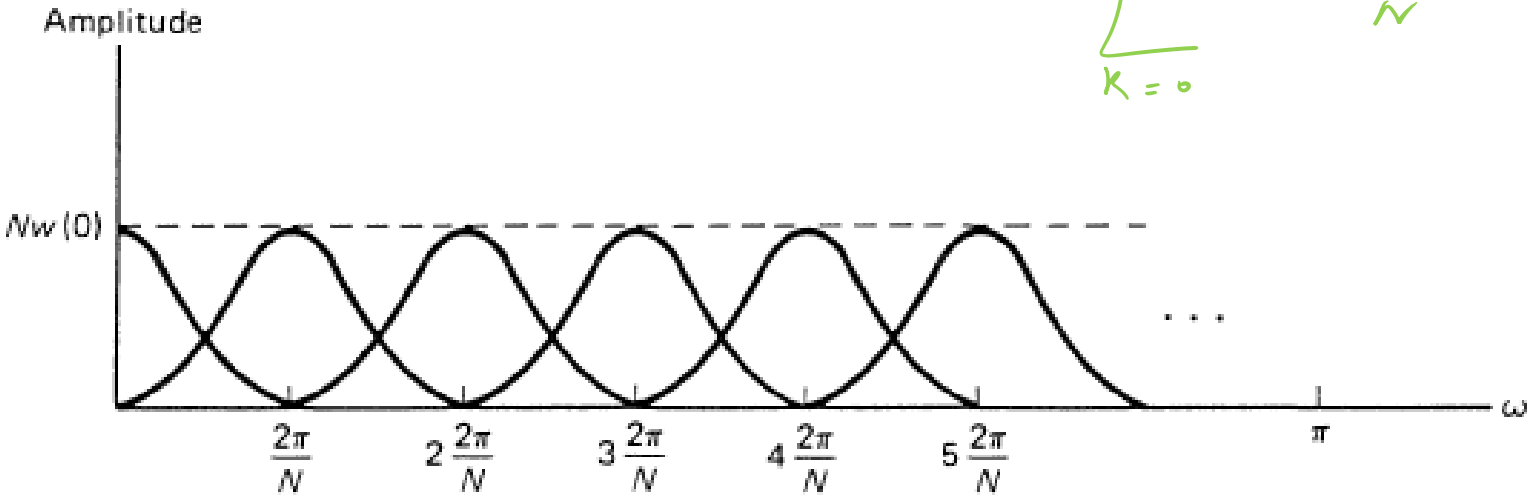


Figure 6.19 The FBS constraint visualized in the frequency domain.

روش OLA : Overlap-Add :

سنتز STFT:
- روش FBS : از دیدگاه فیلترینگ
- روش OLA : " " تبدیل فوریه

ساده‌ترین راه برای سنتز $x(n)$ از STFT : استفاده از تبدیل معکوس (IDFT)

در این روش برای هر زمان ثابت، DFT معکوس از باند فرکانسی متناظر آن زمان گرفته و بر پنجرهٔ آنالیز تقسیم می‌کنیم. ولی این روش در عمل کارایی چندانی ندارد زیرا کوچکترین تغییر در STFT باعث می‌شود که سیگنال تولید شده بسیار متفاوت از سیگنال اصلی باشد. برای مثال اگر STFT، با ضریب فاز $e^{j\omega_0 n}$ که ω_0 نامعلوم است شیفت پیدا کند، تقسیم آن بر پنجرهٔ آنالیز بدون در نظر گرفتن این مقدار شیفت باعث ایجاد اعوجاج در سیگنال خروجی تولید شده می‌شود.

روش OLA که در ادامه توضیح داده می‌شود شرکتی shift یافتهٔ سیگنال اصلی را بدون اعوجاج نشان می‌دهد.

از دیدگاه تبدیل فوریه می‌توان روش OLA را بررسی کرد. در روش OLA، DFT معکوس از زمان کوتاه‌ها گرفته می‌شود ولی به جای تقسیم بر پنجرهٔ آنالیز برای هر بازهٔ زمان کوتاه، از تکنیک هم‌پوشانی و جمع کردن بین بازه‌های زمان کوتاه استفاده می‌شود. این روش به شرطی کار می‌کند که پنجرهٔ آنالیز طوری طراحی شود که روش هم‌پوشانی و جمع کردن به طور موثری بتواند پنجرهٔ آنالیز را از بسامد تولید شده حذف کند.

ایده اصلی روش OLA از رابطه بین دنباله و STFT زمان گسسته آن قابل درک است:

$$x(n) = \left[\frac{1}{2\pi W(0)}\right] \int_{-\pi}^{\pi} \sum_{r=-\infty}^{+\infty} X(r,\omega)\, e^{j\omega n}\, d\omega \qquad (6.28)$$

نسخه STFT زمان گسسته

$$W(0) = \sum_{n=-\infty}^{+\infty} w(n)$$

تبدیل فوریه پنجره ← بزرگ

پنجره

$$X(n,\omega_0) = \left[x(n)\, e^{-j\omega_0 n}\right] * w(n) \qquad (6.9)$$

از دیدگاه فیلترینگ STFT ←

تبدیل فوریه از دو طرف و نتیجه در فرکانس صفر نسبت به n →

$$\tilde{X}(\phi,\omega)\Big|_{\varphi=0} = X(\omega)\, W(0)$$

$\tilde{X}(\phi,\omega)$ تبدیل فوریه $X(n,\omega)$ نسبت به n است.

$$\tilde{X}(\phi,\omega)\Big|_{\varphi=0} = \sum_{r=-\infty}^{+\infty} X(r,\omega) = X(\omega)\, W(0)$$

مجموع نمونه‌های $X(n,\omega)$

$$\longrightarrow X(\omega) = \frac{1}{W(0)} \sum_{r=-\infty}^{+\infty} X(r,\omega)$$

و با تبدیل معکوس از $X(\omega)$ به $x[n]$ می‌رسیم.

روش OLA نسخه‌ی گسسته شده رابطه‌ی (6.28) a است:

$$\omega = \frac{2\pi k}{N}$$

$$y[n] = \frac{1}{W(0)} \sum_{p=-\infty}^{\infty} \underbrace{\left[\frac{1}{N} \sum_{k=0}^{N-1} X(p,k)\, e^{j\frac{2\pi k n}{N}}\right]}_{y_p(n) = x(n)W(p-n)}$$

$$\Rightarrow y[n] = \frac{1}{W(0)} \sum_{p=-\infty}^{+\infty} x(n)W(p-n) = x(n) \underbrace{\frac{1}{W(0)} \sum_{p=-\infty}^{+\infty} W(p-n)}_{=1}$$

بنابراین در صورتی $y(n)=x(n)$ برابر خواهد شد اگر:

شرط OLA روی پنجره‌ی آنالیز

$$\sum_{p=-\infty}^{+\infty} W(p-n) = W(0)$$

اگر discrete STFT در زمان با ضریب L، decimate شود، می‌توان به طور مشابه نشان داد که شرط پنجره‌ی آنالیز باید به صورت زیر باشد:

شرط OLA روی پنجره‌ی آنالیز

$$\sum_{p=-\infty}^{+\infty} W(pL-n) = \frac{W(0)}{L}$$

و رابطه‌ی سنتز

$$x(n) = \left[\frac{L}{W(0)}\right] \sum_{p=-\infty}^{+\infty} \left[\frac{1}{N} \sum_{k=0}^{N-1} X(pL,k)\, e^{j\frac{2\pi k}{N} n}\right]$$

شرط OLA :

$$\boxed{\sum_{p=-\infty}^{+\infty} w(pL-n) = \frac{W(0)}{L}} \qquad (6.37)$$

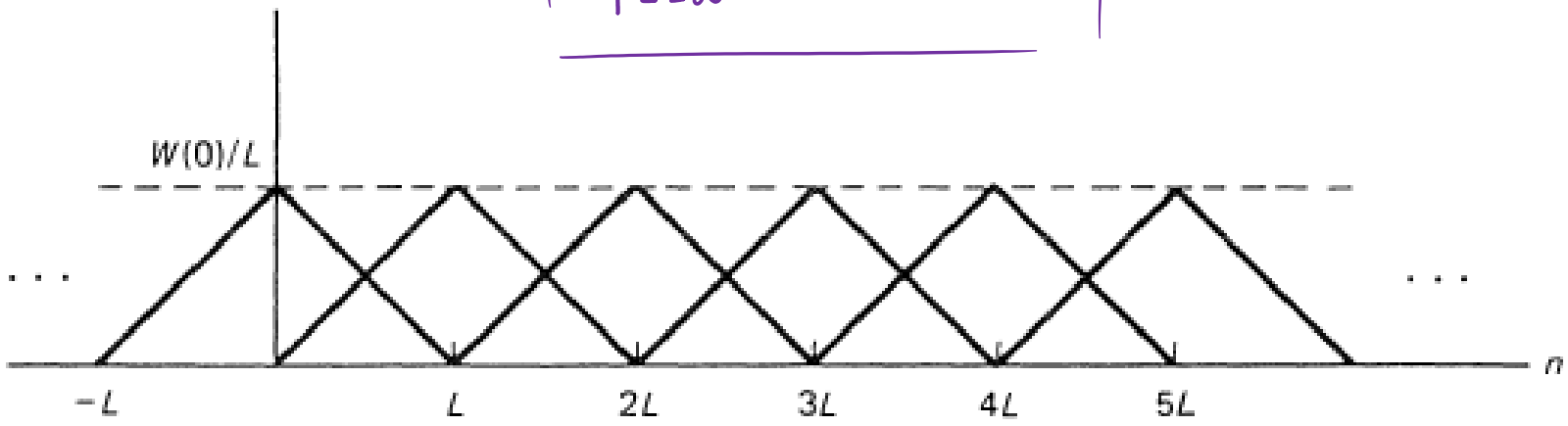


Figure 6.20 The OLA constraint visualized in the time domain.

نکته: شرط FBS در حوزهٔ فرکانس دوگان شرط OLA در حوزهٔ زمان است.

شرط FBS :

$$\sum_{k=0}^{N-1} W\left(\omega - \frac{2\pi k}{N}\right) = N W(0)$$

حاصل جمع شیفت یافته‌های تبدیل فوریه پنجره آنالیز باید مقدار ثابتی باشد

در روش FBS دیدیم که برای پنجره های طول که و اگر طول پنجره از ضریب نمونه برداری فرکانسی کوچکتر باشد، شرط FBS همواره برقرار است.

در مقایسه می توان نشان داد در روش OLA اگر پنجره آنالیز پهنای باند کوچک با فرکانس پیشینه کمتر از $\frac{2\pi}{L}$ داشته باشد شرط OLA برآورده می شود (L : ضریب decimation)

برای مثال پنجره ای با پهنای باند $2\omega_c$ را در نظر بگیرید.

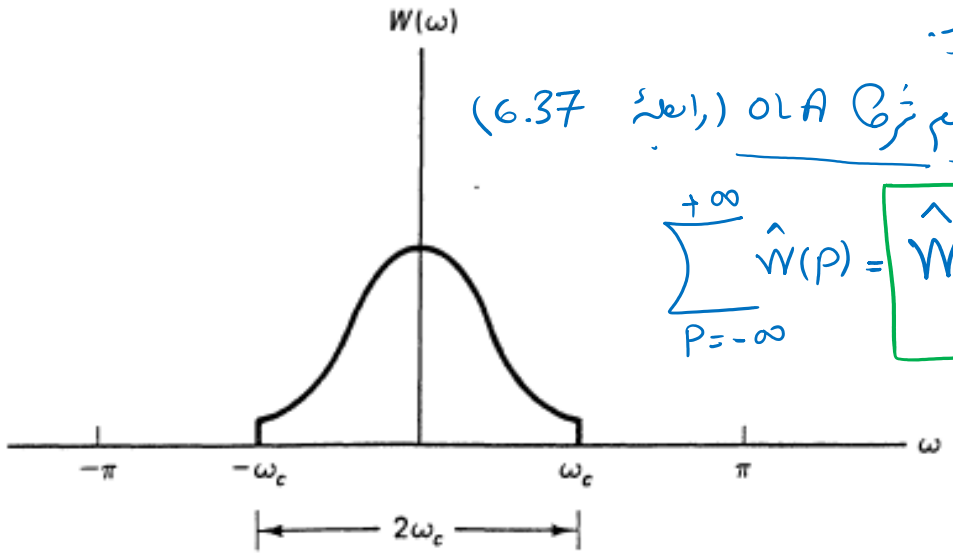


Figure 6.21 Example of a Fourier transform of an analysis window with bandwidth $2\omega_c$.

اگر دنباله $w(pL-n)$ را با $\hat{w}(p)$ نمایش دهیم شرط OLA (رابطه 6.37) را می توان به فرم زیر نوشت:

$$\sum_{p=-\infty}^{+\infty} \hat{w}(p) = \boxed{\hat{W}(0) = \frac{W(0)}{L}} \quad (6.39)$$

$\hat{W}(\omega)$ تبدیل فوریه $\hat{w}(n)$ است.

از آنجایی که $\hat{w}(p)$ نمونه برداری شده $w(p-n)$ است، می توان نوشت:

$$\hat{W}(\omega) = \frac{1}{L}\sum_{k=-\infty}^{+\infty} e^{-j(\omega-\frac{2\pi}{L})n}\, W(\omega-\frac{2\pi}{L}k)$$

درصورتی که هم پوشانی بین $W(\omega)$ و $W(\omega-\frac{2\pi}{L}k)$ در $\omega=0$ وجود نداشته باشد به رابطه 6.39 می رسیم.

نتیجه: هر پنجره آنالیز با پهنای باند کوچک که فرکانس پیشینه آن از $\frac{2\pi}{L}$ کمتر است شرط OLA را برآورده می کند.

شرط عدم هم پوشانی: $\boxed{\omega_c < \frac{2\pi}{L}}$

مقایسه از نظر پیچیدگی محاسباتی:

به عملیات از $O(N_W)$ برای هر نمونه $x(n)$ نیاز دارد: FBS با $L=1$ و طول پنجره N_W

" " " " OLA : " " $O(N_W \log_2 N_W)$ "

$$L \gg 1 \rightarrow O(N_W^2) : FBS$$

$$L \gg 1 \rightarrow O(\log_2 N_W) : OLA$$

نتیجه: برای نرخ decimation های زیاد روش OLA بسیار کارآمدتر از FBS است.

روش های دیگر:

- 6.3.3: تعمیم یافته FBS
- 6.3.4: Helical Interpolation Method
- 6.3.5: Weighted OLA

6.4 اندازهٔ تبدیل فوریهٔ زمان کوتاه

STFTM

Short Time Fourier Transform Magnitude

در برخی کاربردها مانند پردازش گفتار spectrogram که می‌تواند به اندازهٔ STFT مرتبط شود نقش عمده‌ای دارد. نمایش‌های مشابه spectrogram مانند اندازهٔ STFT نیز در این زمینه مورد استفاده قرار می‌گیرند. اندازهٔ STFT (STFTM) یک تابع نامنفی زمان-فرکانس time-freq می‌باشد. در حالی که STFT یک تابع با مقدار مختلط در حالت کلی است و برای کاربردهایی مانند اصلاح مقیاس زمانی صحبت، تخمین فاز این تابع از نظری بسیار پیچیده است. ولی روش‌های وجود دارد که بسیار پردازش شده را تنها از روی اندازهٔ STFT (STFTM) استخراج می‌کنند و بنابراین مشکل تخمین فاز را حل می‌کنند. در این بخش اندازهٔ STFT را به عنوان نمایش دیگری در حوزهٔ زمان-فرکانس معرفی می‌کنیم و روابط آنالیز و سنتز آن را بررسی می‌کنیم.

همانند STFT زمان گسسته، STFTM زمان گسسته نیز متقارن و پریودیک با پریود 2π است. تفاوت آنها در این است که STFTM حقیقی است.

STFTM می‌تواند با تابع دیگری با عنوان خودهمبستگی زمان کوتاه. (short time autocorrelation) مرتبط شود:

$$r(n,m)$$

- تابع خودهمبستگی زمان کوتاه:

$$r(n,m) = \frac{1}{2\pi}\int_{-\pi}^{\pi} |X(n,\omega)|^2 e^{j\omega n}\, d\omega$$

$$|X(n,\omega)|^2 = \sum_{m=-\infty}^{+\infty} r(n,\overset{m}{\cancel{\omega}})\, e^{-j\omega n}$$

$|X(n,\omega)|^2$ تبدیل فوریه $r(n,\overset{m}{\cancel{\omega}})$ است.

$$r(n,m) = [x(m)w(m-n)] * [x(-m)w(-m-n)]$$

$$= \sum_{p=-\infty}^{+\infty} x(p)w(p-n)x(p-m)w(p-m-n)$$

در حالت کلی بخش زمان کوتاه. $f_n(m) = x(m)w(n-m)$ را نمی‌توان از $r(n,m)$ به دست آورد.

دنباله زمان کوتاه.

اندازه طیف فوریه زمان کوتاه.

خود همبستگی زمان کوتاه.

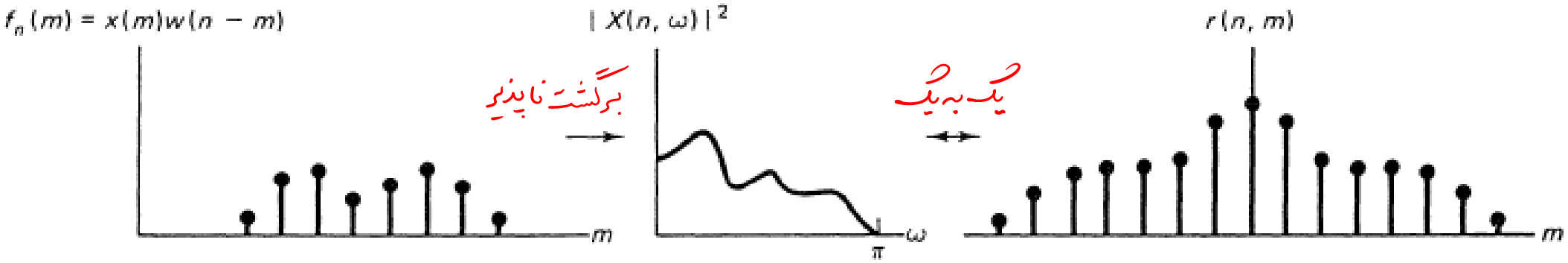


Figure 6.23 Illustration of the noninvertibility of the short-time Fourier transform magnitude for a fixed n. The Fourier transform magnitude is one-to-one with the autocorrelation.

در حالت کلی STFTM نمایش یکتایی از سیگنال ها را ارائه نمی دهد.

نشان خواهیم داد که خود همبستگی های بخش های زمان کوتاه در صورتی که هم پوشانی جزئی داشته باشند می توانند برای به دست آوردن بخش های زمان کوتاه مورد استفاده قرار گیرند. بنابراین تحت شرایط مشخصی می توان STFTM را به عنوان نمایش یکتایی از سیگنال در نظر گرفت.

6.7.1 نمایش پیشانی:

هدف: بررسی مسأله مشخص کردن STFTM زمان گسسته که بتوان از آن برای نمایش پیشانی دنباله استفاده کرد.

STFTM پیشانی برای سیگنال $x(n)$ و $-x(n)$ STFTM یکسانی دارند.

شرایط خاص پتیل پیشانی STFTM

- اگر پنجرهٔ آنالیز روی بازهٔ طول محدود (N_w) غیرصفر باشد، دنباله‌های یکطرفه می‌توانند به طور یکتا (با ابهام در علامت $\pm$) نمایش داده شوند به شرطی که لااقل صفرهای پیاپی بین دو نمونهٔ غیرصفر در سیگنال، حداکثر N_w-2 باشد. به عبارت دیگر علامت تنها یک نمونه از سیگنال را داشته باشیم، پتیل کامل یکتا خواهد بود.

$x(n) = x_1(n) + x_2(n)$

$x(n) = x_1(n) - x_2(n)$

$x(n) = -x_1(n) + x_2(n)$

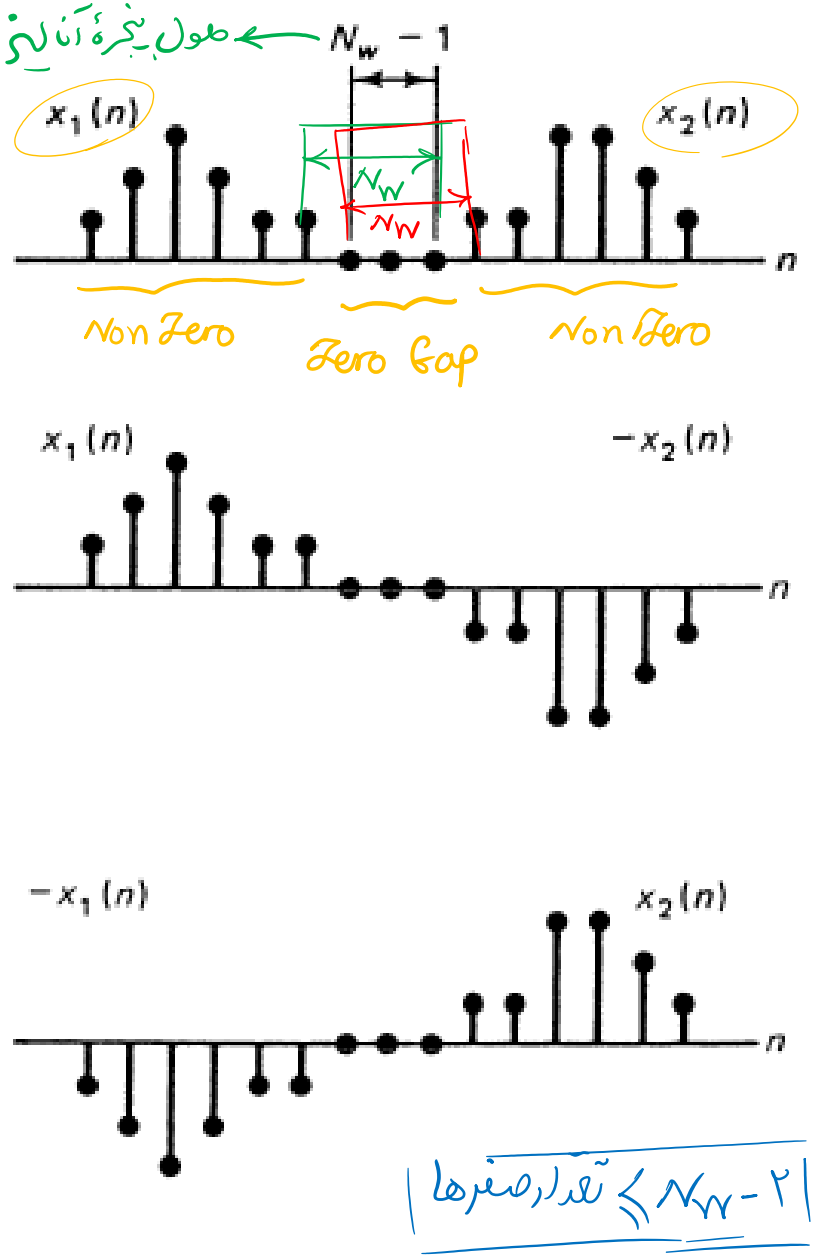


Figure 6.24 Three sequences with the same STFTM.

$x(n) = x_1(n) + x_2(n)$

اگر صفرهای بین x_1 و x_2 به اندازه‌ای زیاد باشد که پنجرهٔ آنالیز در هیچ موقعیتی نتواند هم نمونه‌های دنباله $x_1(n)$ و هم نمونه‌های دنباله $x_2(n)$ را در بر بگیرد، اندازهٔ STFTM یکتا نخواهد بود.

زیرا در این حالت STFTM برای $x(n)$ مجموع STFTM، x_1 و x_2 خواهد شد.

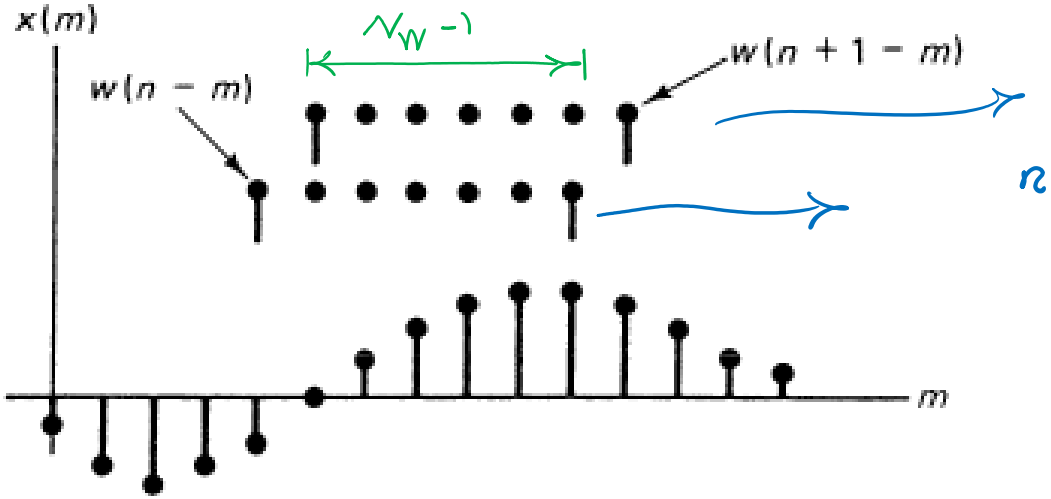
* در نتیجه برای شرایط یکتا بودن STFTM یکی از کمبودیت‌ها این است که تعداد صفرهای بین بخش‌های غیرصفر نباید از حد مشخصی بیشتر شود.

تعیین دنباله یکطرفه $x(n)$ از روی STFTM آن:

چون STFTM به تنهایی این قابلیت را ندارد و به اطلاعات فضایی نیاز داریم که این اطلاعات فضایی از هم پوشانی بین نمونه‌های پشت سر هم موقعیت های مختلف پنجره آنالیز به دست می آید

اگر پنجره آنالیز با طول N_w و غیر صفر باشد N_w-1 نمونه اول بخش زمان کوتاه، در زمان $n+1$ و N_w-1 نمونه آخر بخش زمان کوتاه در زمان n برابر است.

بنابراین با برون یابی (extrapolation) می توان آخرین نمونه بخش $n+1$ را از روی N_w-1 نمونه اول به دست آورد و با تکرار این روند، $x(n)$ را بازیابی کرد.



پنجره آنالیز در زمان $n+1$

پنجره آنالیز در زمان n

Figure 6.25 Illustration of the consistency required among adjacent short-time sections. Note the samples that are common to the adjacent sections.

هدف : محاسبهٔ $x(n)$ از بردن پایی N_W-1 نمونه مشخص ولهٔ اندازهٔ $|X(n,\omega)|$ یا $r(n,m)$ مدول آن :

$$r(n, N_W-1) = [w(0)\,x(n)][w(N_W-1)\,x(n-(N_W-1))]$$

$$\boxed{x(n) = \frac{r(n, N_W-1)}{w(0)\,w(N_W-1)\,x(n-(N_W-1))}}$$

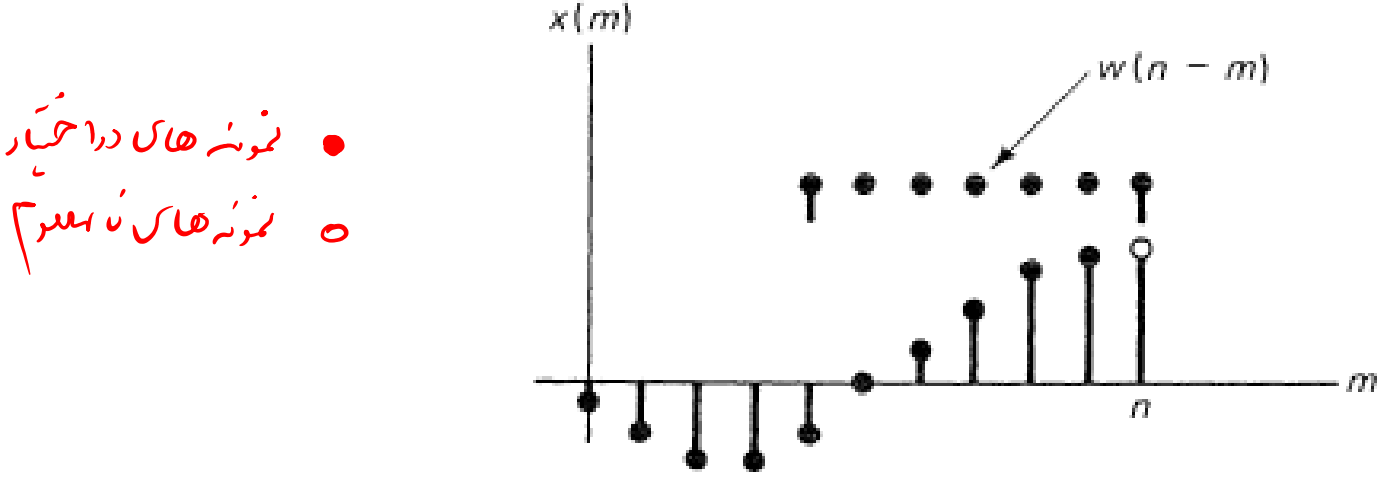


Figure 6.26 Illustration of the samples under $w(n-m)$.

پنجرهٔ آنالیز : مستطیلی شکل

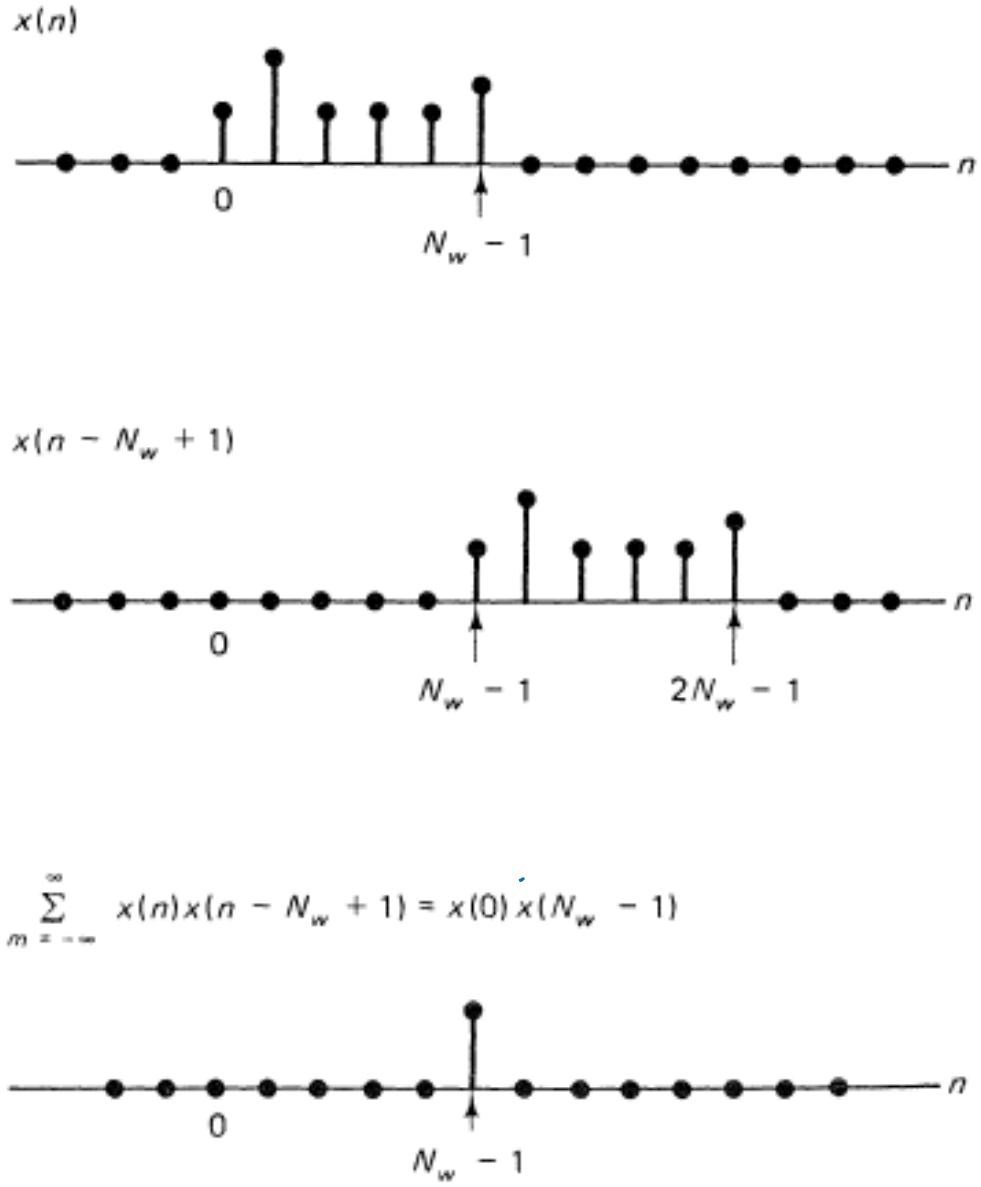


Figure 6.27 Computation of the last nonzero autocorrelation sample (assuming a rectangular analysis window).

6.4.2 آنالیز STFTM:

discrete time STFTM $|X(n,\omega)|$

discrete STFTM $|X(n,k)|$

پیاده سازی STFTM با عملیات حقیقی :

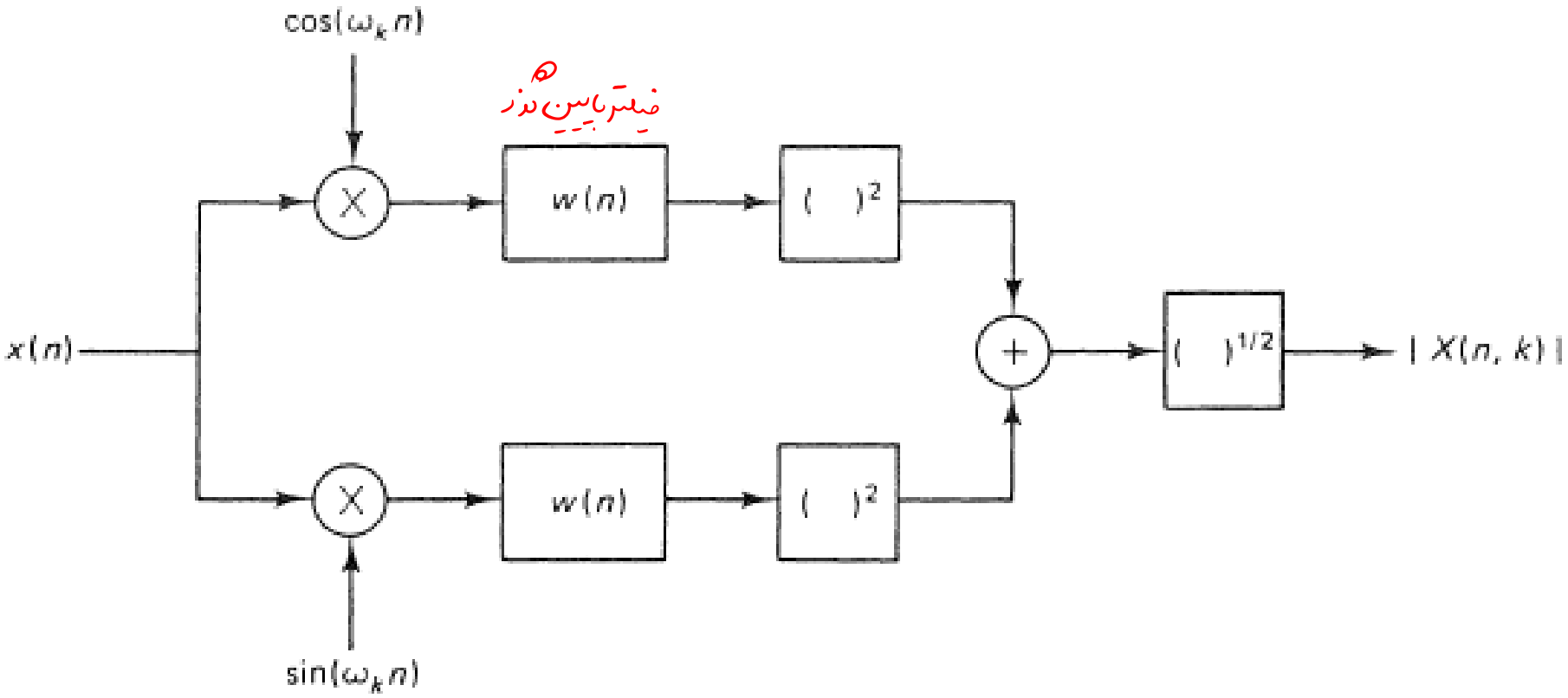


Figure 6.28 Implementation with real operations of each branch of the lowpass filter bank for STFTM analysis.

با فیلتر بانک میان‌گذر :

فیلتر میان‌گذر

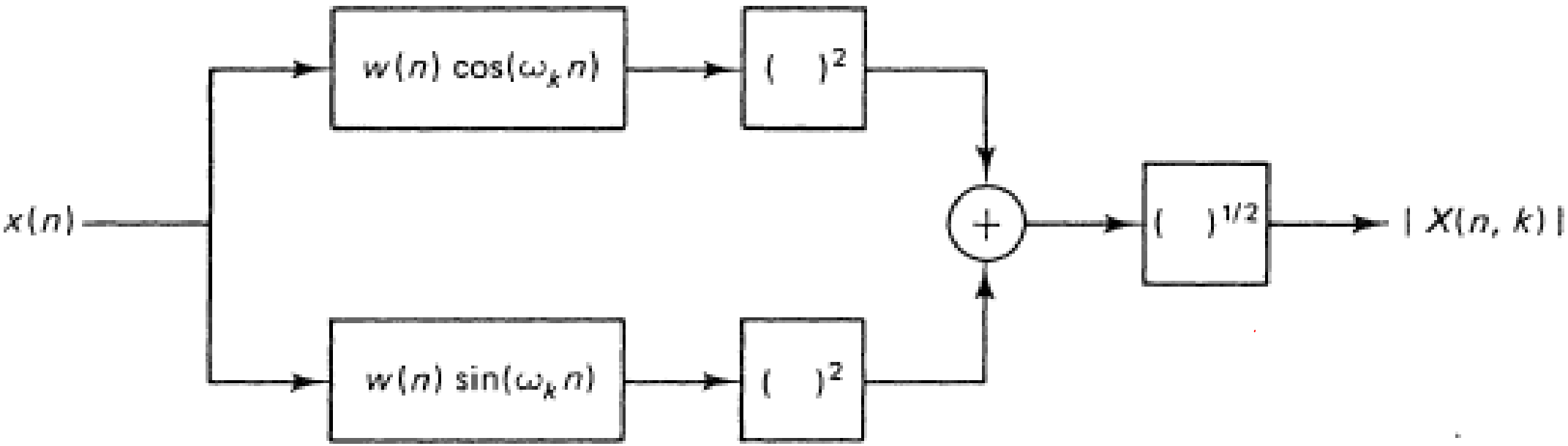


Figure 6.29 Implementation with real operations of each branch of the bandpass filter bank for STFTM analysis.

Sequential Extrapolation Approach روش برونیابی ترتیبی

6.4.3 نسخه STFTM :

STFTM از $x(n)$

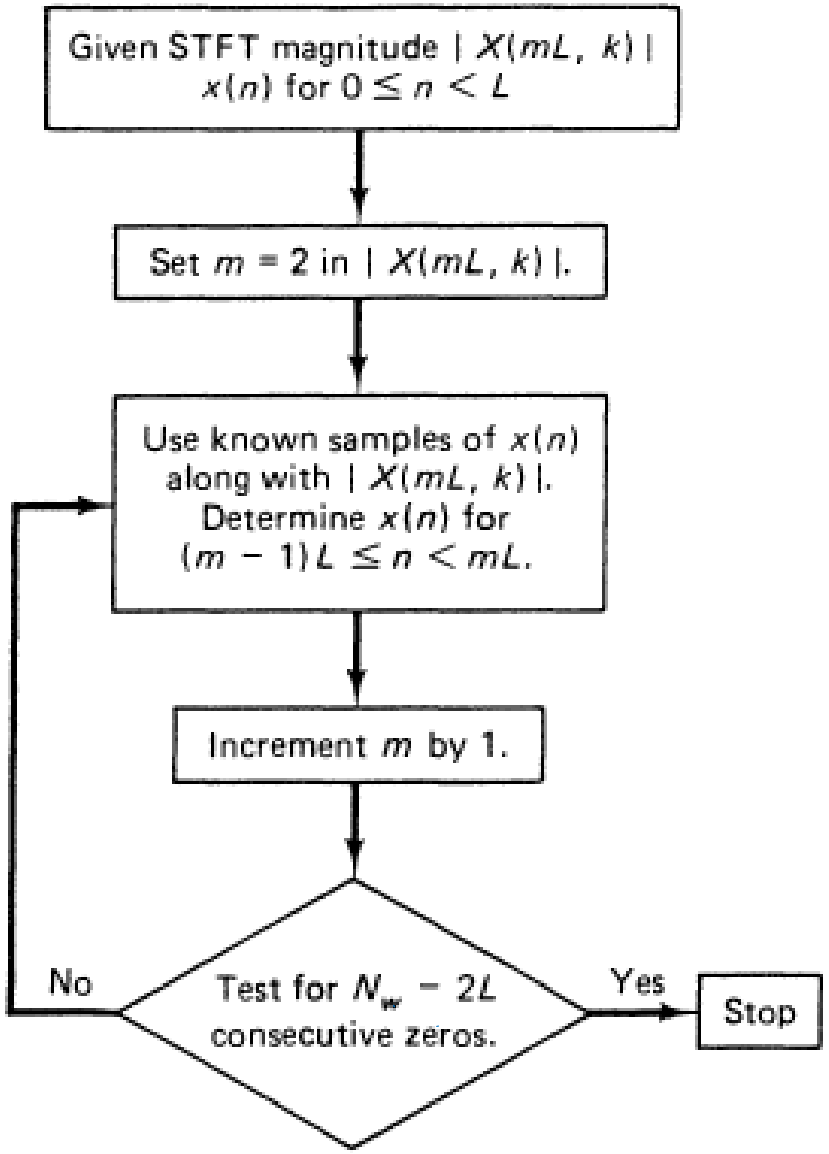


Figure 6.30 The sequential extrapolation approach for STFTM synthesis.

6.7. کاربردهای STFT

6.7.1 پردازش گفتار Speech Processing

دو کاربرد:

- تغییر مقیاس زمانی time-scale modification ← بخش سریع تر یک صحبت ضبط شده برای مرور کردن تر
  " کندتر " " " برای درک بهتر (آموزش زبان)
- کاهش پهنای باند Bandwidth Reduction ← بیان شدن با نمایش فشرده تر

مدل سیگنال صحبت

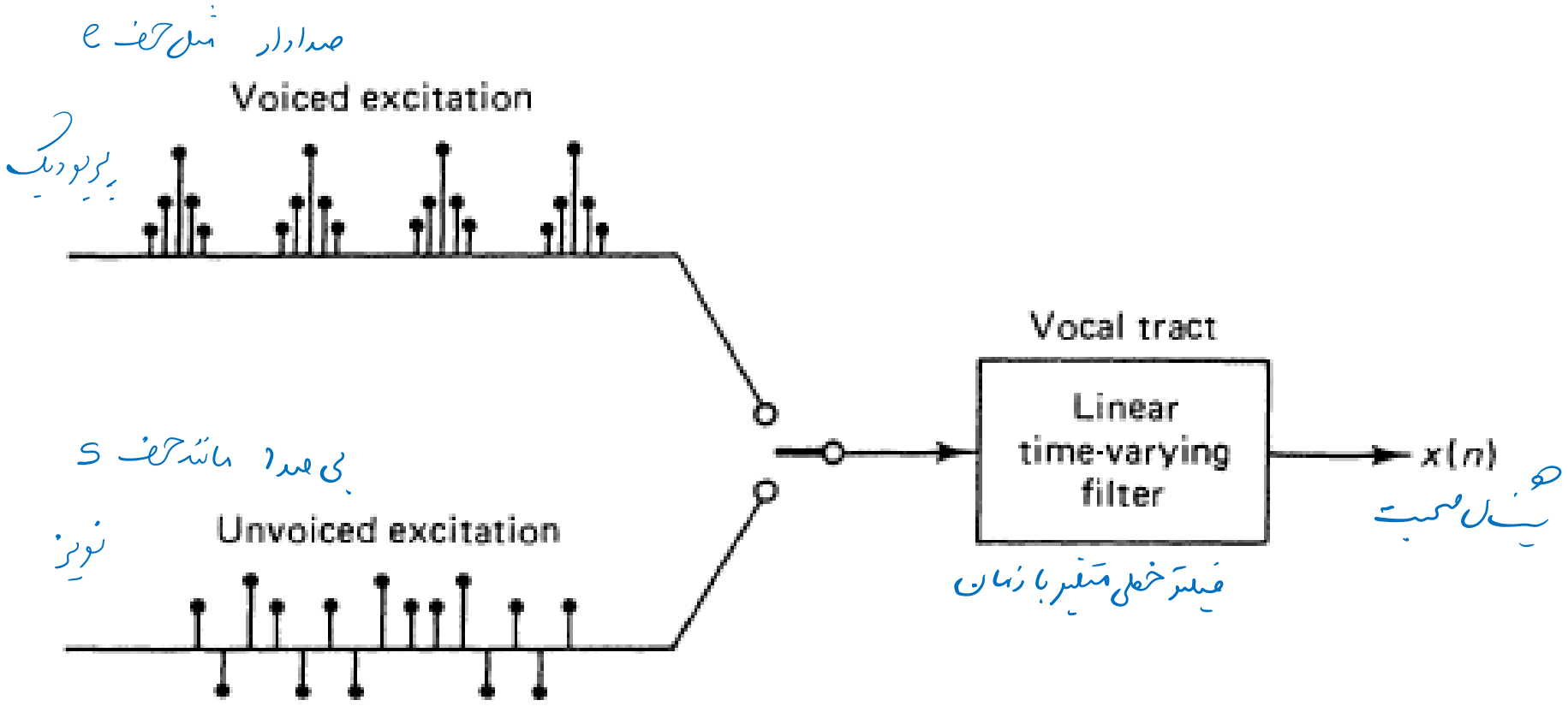


Figure 6.33 Speech production model.

- هر کدام از صداها حدود 20ms طول می‌کشند که می‌توان فرکانس سیگنال در این بازه ثابت می‌ماند.
- از خاصیت ایستا بودن (stationary) سیگنال در این بازه می‌توان برای کاهش پهنای باند استفاده کرد.
- تغییرات تدریجی سیگنال در حوزه تغییر مقیاس زمانی قابل استفاده است.

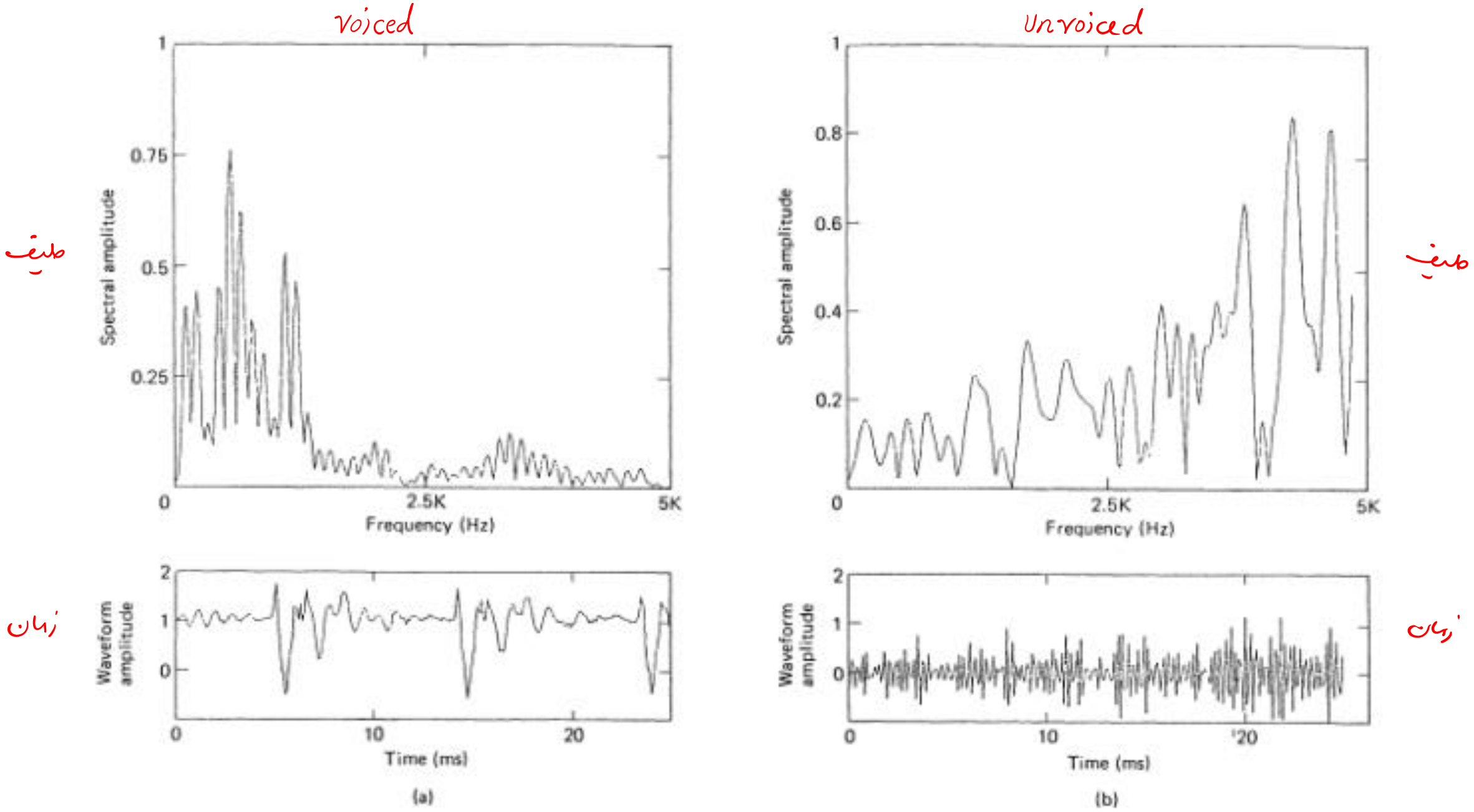


Figure 6.34 Fourier transform magnitudes of (a) typical voiced and (b) typical unvoiced sounds.

: vocoder

voice coder

هدف از طراحی vocoder ها، کاهش پهنای باند لازم برای انتقال سیگنال صحبت

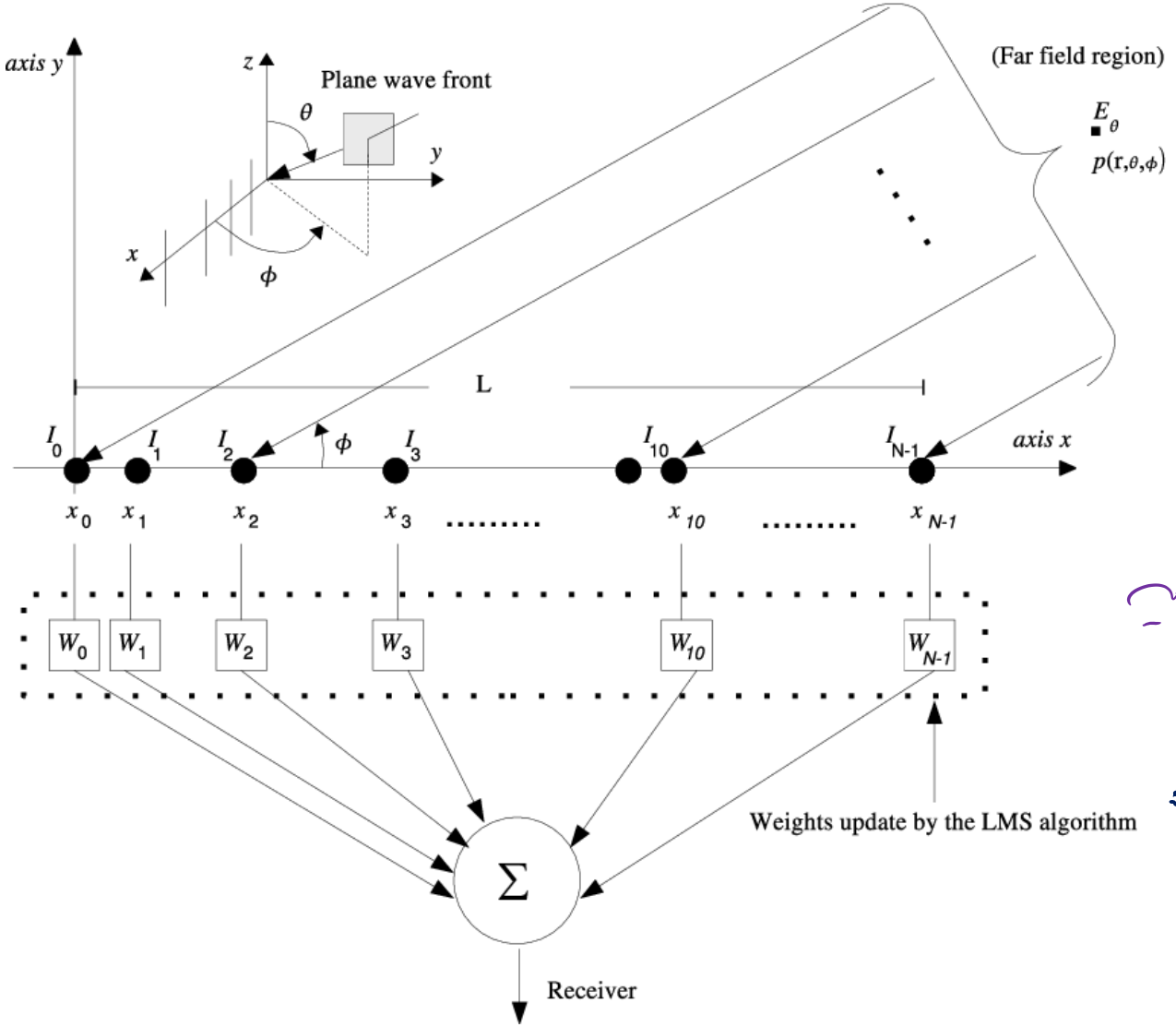
مثال: صحبت با 8000 sample/sec و 12 بیت کوانتیزه: 96000 bit/sec که به کانال با پهنای زیادی نیاز دارد.

نمونه برداری

STFT مدل پارامترهای آن یا discrete STFT در هر بازه 20ms و کوانتیزه کردن نمونه های آن

6.7.2 : پردازش سیگنال آرایه‌ای

Array Signal Processing



سونار و حادثه‌های جغرافیایی

سیگنال دریافتی سنسور iام

مدل ساده

$$x(n) = \frac{1}{N}\sum_{i=0}^{N-1} x_i(n-n_i)$$

تأخیر سنسور iام

N : تعداد حسگرها

فرکانس

براساس تبدیل فوریه

$$X(\omega) = \frac{1}{N}\sum_{i=0}^{N-1} X_i(\omega)\, e^{-j\omega n_i}$$

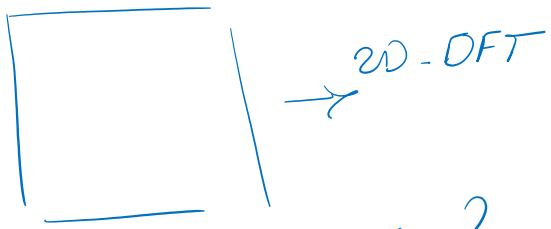
با توجه به تغییرات کند محتوای فرکانسی سیگنال‌های دریافتی از حسگرها

می‌توان از STFT برای پردازش سیگنال‌های آرایه‌ای استفاده کرد:

$$X_n(\omega) = \frac{1}{N}\sum_{i=0}^{N-1} X_i(n,\omega)\, e^{-j\omega n_i}$$

با گسسته‌سازی می‌توان به رابطه discrete STFT رسید.

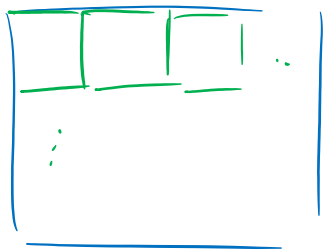
6.7.3 پردازش تصویر:



فرض غلط: در کاهش نویز با فیلتر ویینر فرض می‌شود که محتوای فرکانسی تصویر در جهت های مختلف ثابت است.

روش بهتر: در عمل فرض بالا برقرار نیست یعنی هر تصویر ممکن است از زیرتصویرهای مختلف با بافت های متفاوت تشکیل شده باشد.

بنابراین بهتر است که تصویر را به چنین subimage ها تقسیم کنیم و هر یک از آنها را جداگانه پردازش کنیم.



می‌توان زیربخش ها را به قدری کوچک انتخاب کرد که محتوای فرکانسی یکنواختی داشته باشند.

سازوکار لازم برای این راه حل: 2D-STFT یا مشهور.

$$X(n_1,n_2,\omega_1,\omega_2)=\sum_{m_1}\sum_{m_2}x(m_1,m_2)\,w(n_1-m_1,n_2-m_2)\,e^{-j(\omega_1 m_1+\omega_2 m_2)}$$

$w(n_1,n_2)$ پنجره دوبعدی