

فصل ششم

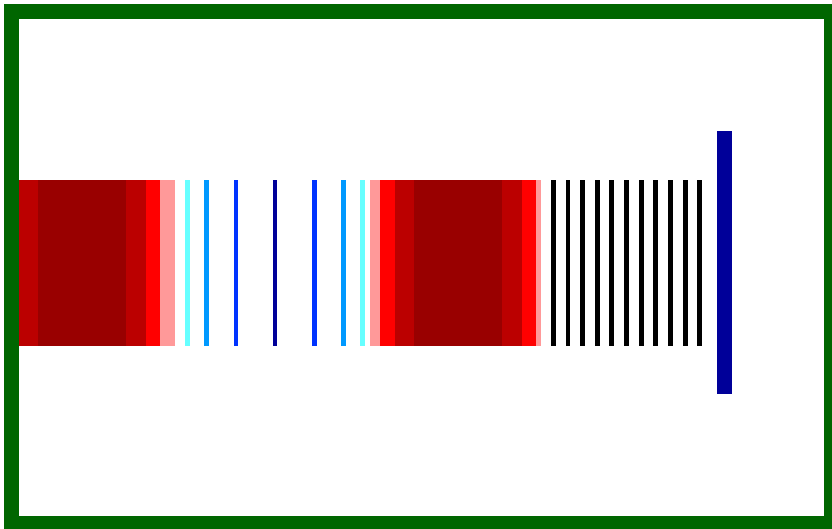
اصول پایه‌ای صدای دیجیتال

دیجیتال سازی صدا
رابط رقمی ابزار موسیقی (MIDI)
انتقال صدا



مفاهیم اولیه

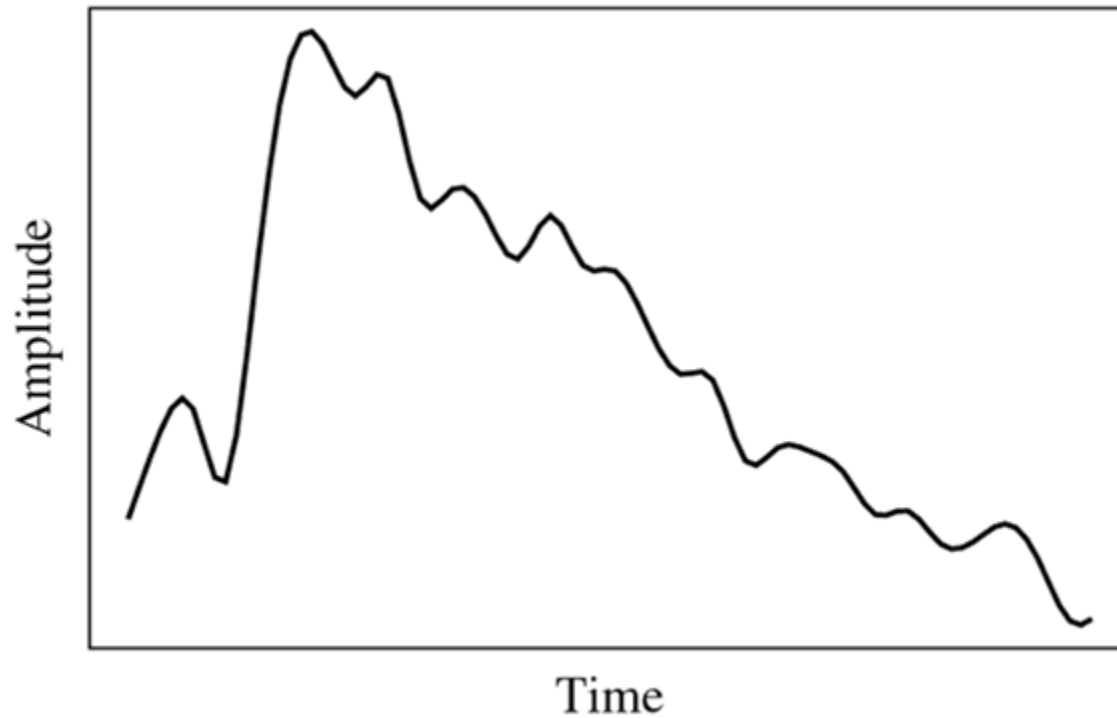
- صدا نوعی موج است، در هوا حرکت می کند و از تغییر فشار هوا تولید می شود.
- صدا کمیتی پیوسته محسوب می شود. برای دیجیتال سازی آن باید گسسته شود.





دیجیتال سازی صدا

◀ دیجیتال سازی (Digitization): دیجیتال سازی فرآیند تبدیل مقادیر پیوسته به یک رشته عدد صحیح است.



یک سیگنال آنالوگ

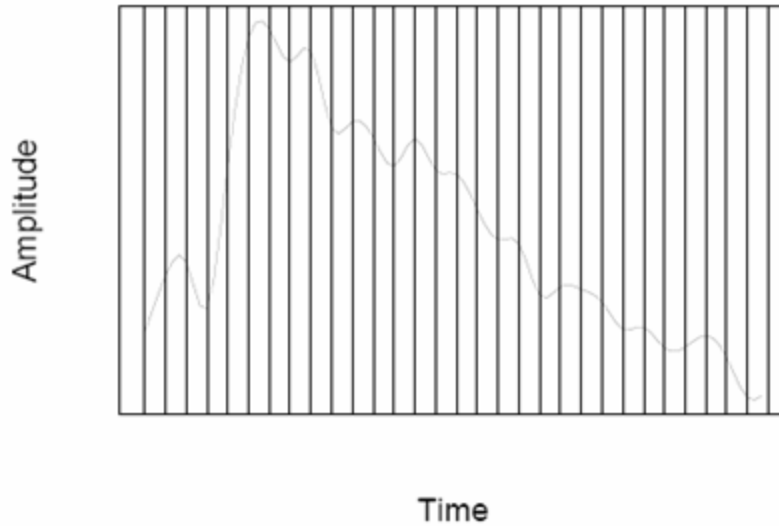
سیستم‌های چندرسانه‌ای



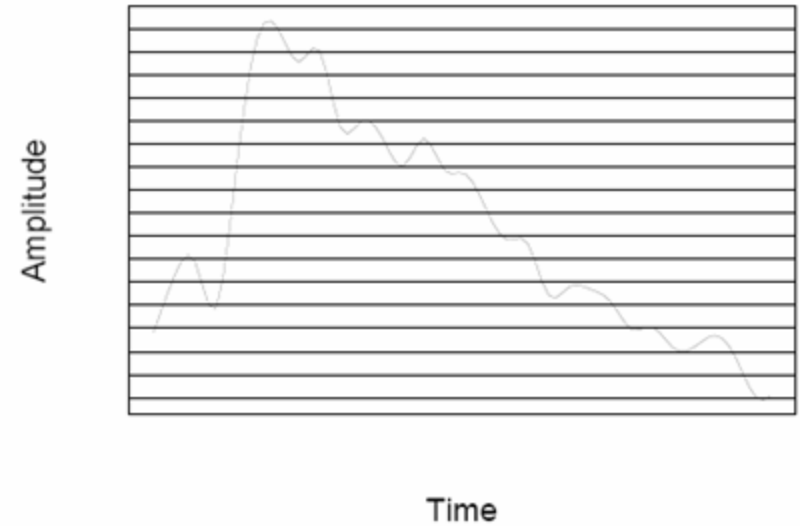
دیجیتال سازی صدا

- ◀ برای تبدیل کمیت پیوسته به گسسته نمونه برداری باید انجام شود.
- ◀ نرخ که بر اساس آن نمونه برداری انجام می شود «فرکانس نمونه برداری» نامیده می شود.
- ◀ برای صدا، نرخ معمولی نمونه برداری از ۸ کیلو هرتز (۸۰۰۰ نمونه در ثانیه) تا ۴۸ کیلو هرتز است.
- ◀ نمونه برداری در بعد دامنه یا اندازه ولتاژ، سطح بندی یا Quantization نامیده می شود.

دیجیتال سازی صدا (نمونه برداری و کوانتیزه کردن سیگنال)



نمونه برداری سیگنال آنالوگ در بعد زمان

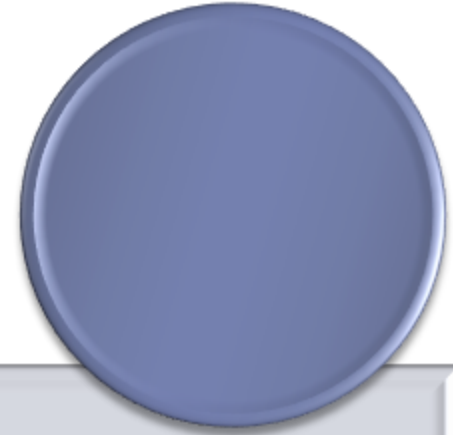


نمونه برداری سیگنال آنالوگ در بعد دامنه



دیجیتال سازی صدا

- برای این که تصمیم بگیریم چه طور داده های صوتی را دیجیتال سازی کنیم می بایست به سوالات زیر جواب دهیم:
۱. نرخ نمونه برداری (Sampling Rate) چه مقداری باشد؟
 ۲. داده ها چگونه سطح بندی شوند؟ (Quantization)
 ۳. قالب داده صوتی چیست؟

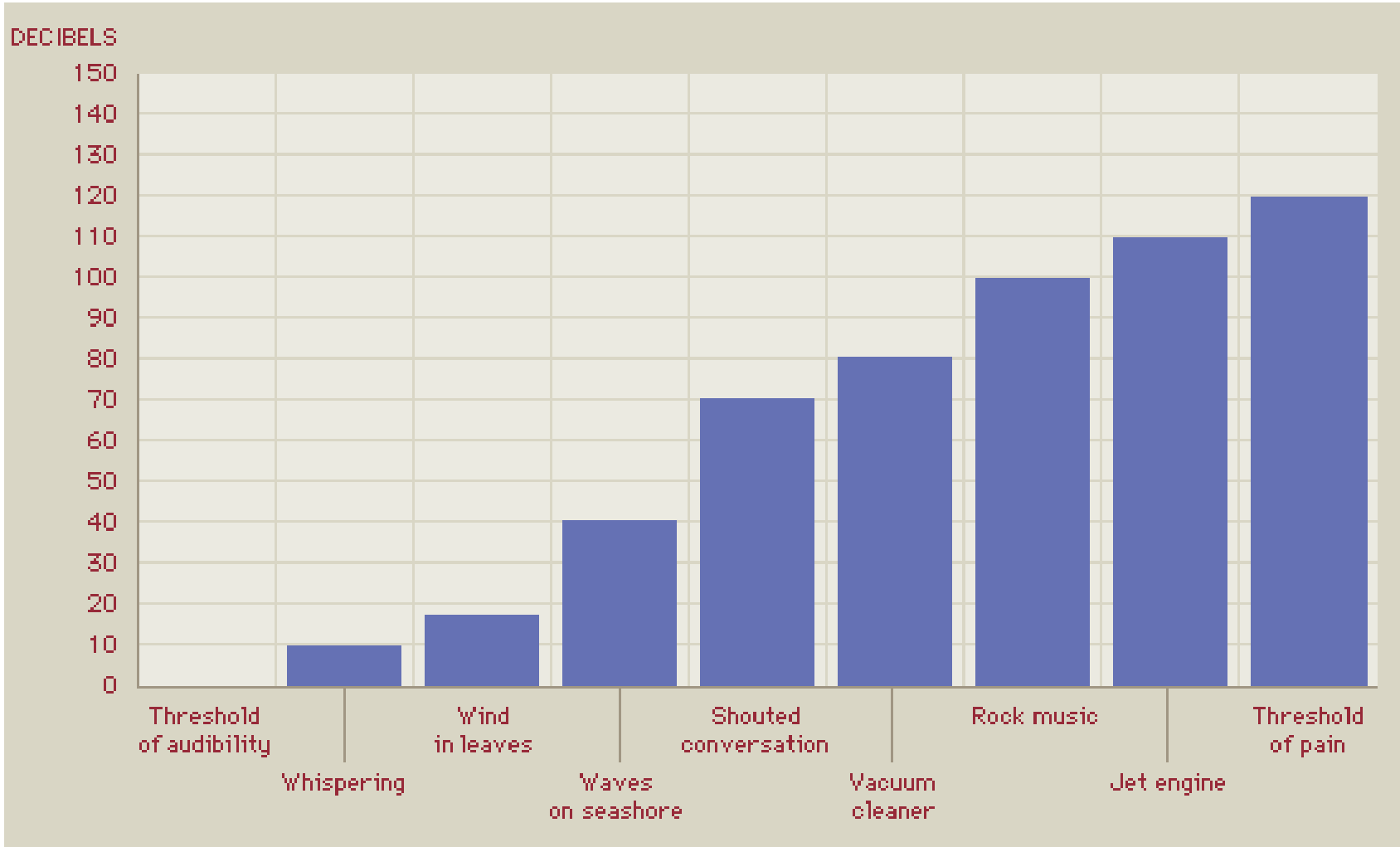


قضیه نایکوپیست

برای رقمی سازی بدون اتلاف باید فرکانس نمونه برداری
دو برابر بیشترین فرکانس موجود در سیگنال باشد



بلندی صوت



بلندی صدا با دسی بل اندازه گیری می شود (دسی بل بیانگر یک نسبت و بدون واحد است).

دسی بل اندازه گیری به صورت لگاریتمی است.

اگر دامنه صدایی صد برابر شود، صد برابر بلندتر شنیده نمی شود.

اگر شدت صدا 6db افزایش یابد صدا دو برابر به نظر می رسد.

اگر I شدت صدا (Sound Intensity) باشد، بلندی صدا از رابطه زیر به دست می آید:

$$L_I = 10 \log_{10} \left(\frac{I_1}{I_0} \right), I_0 = 10^{-12} \text{ W/m}^2$$

آستانه شنوایی

وات بر متر مربع

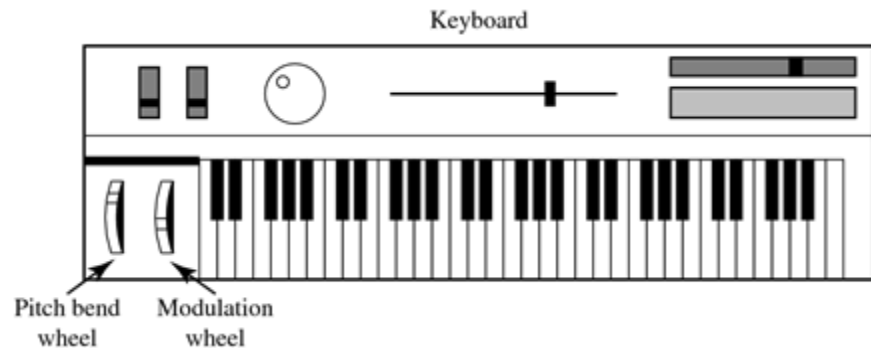
رابطه رقمی ابزار موسیقی (MIDI)



رابطه رقمی ابزار موسیقی

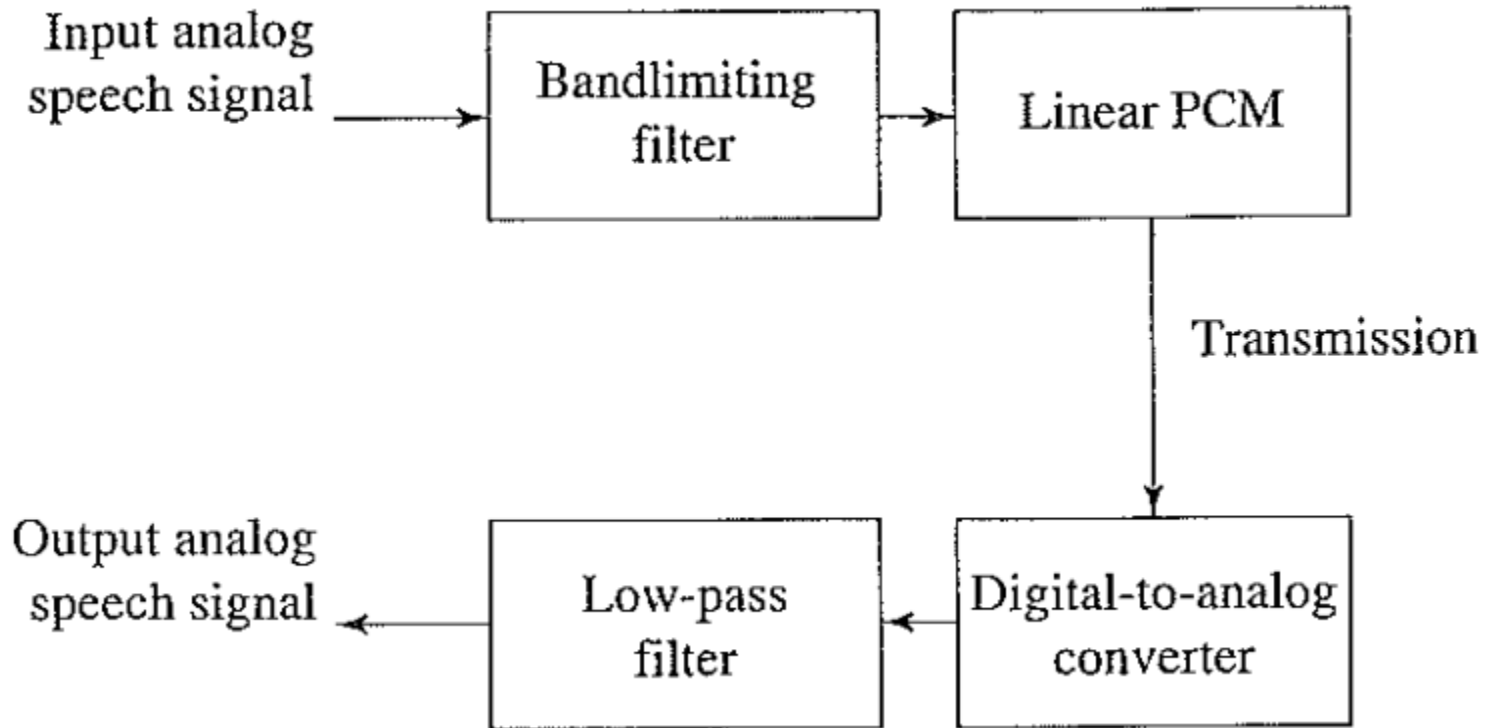
Musical Instrument Digital Interface (MIDI)

- ◀ MIDI یک زبان اسکرپتی است که رویدادها (Events) را برای تولید صدا کد و ذخیره می‌کند. به عنوان نمونه یک فایل MIDI ممکن است شامل مقادیر گام‌های یک نت، مدت و حجم صدای آن باشد.
- ◀ MIDI توسط صنعت موسیقی الکترونیکی برای کنترل ابزارهای تولید موسیقی از قبیل ترکیب‌کننده‌ها و کارت‌های صدا استاندارد شده است.
- ◀ کامپیوترها باید واسط مخصوص MIDI داشته باشند، اما واسط MIDI در اغلب کارت‌های صدا گنجانده شده است. همچنین کارت صدا باید تبدیل‌کننده D/A و A/D داشته باشند.





انتقال صوت





فیلتر کردن صوت (Audio Filtering)

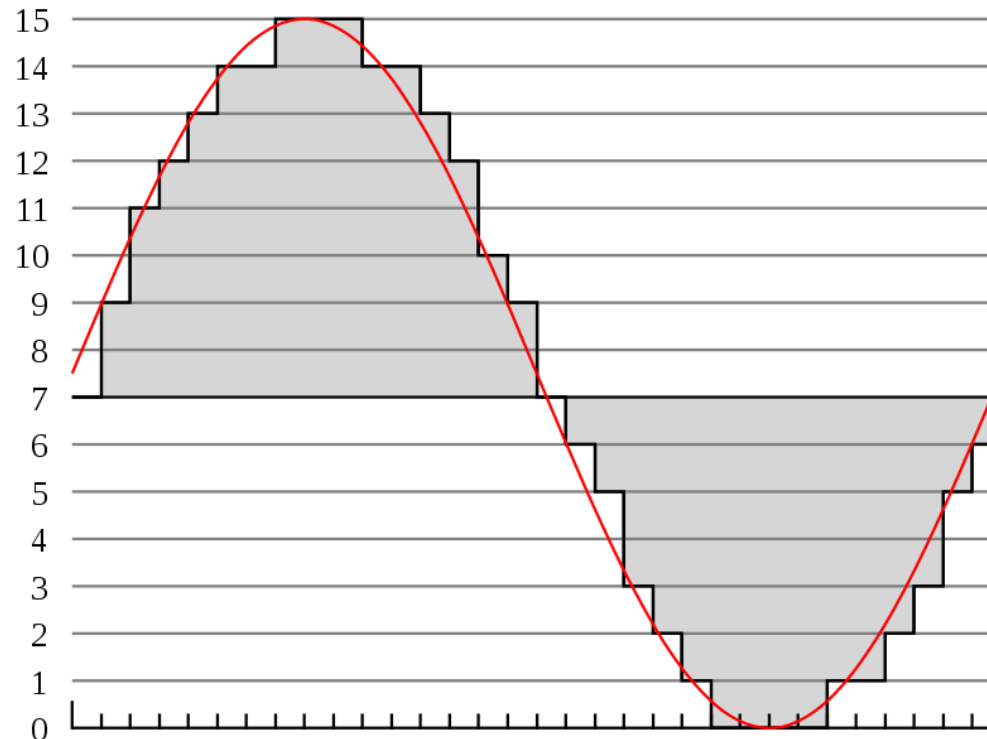
- ◀ قبل از نمونه‌گیری و تبدیل آنالوگ به دیجیتال، سیگنال صوت را معمولاً به منظور رفع فرکانس‌های ناخواسته فیلتر می‌کنند.
- ◀ محدوده شنوایی انسان 20Hz تا 20KHz است.
- ◀ حداکثر فرکانس موجود در صدای انسان در حدود 4KHz است.
- ◀ برای سیگنال تلفنی فرض می‌شود ماکزیمم فرکانس 4KHz است. بنابراین نرخ نمونه‌برداری 8kHz در نظر گرفته می‌شود.

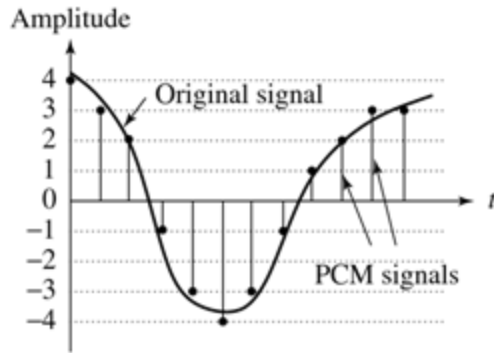
$$8000 \times 8\text{bit} = 64000\text{bit} = 64\text{kbps}$$

تعیین محدوده کار
برای مودم

◀ برای انتقال سیگنال صدا ابتدا آن را کوانتیزه کرده و با استفاده از PCM دیجیتال می‌کنیم.

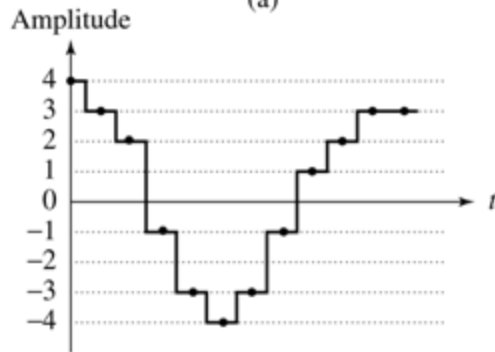
PCM=Pulse Code Modulation





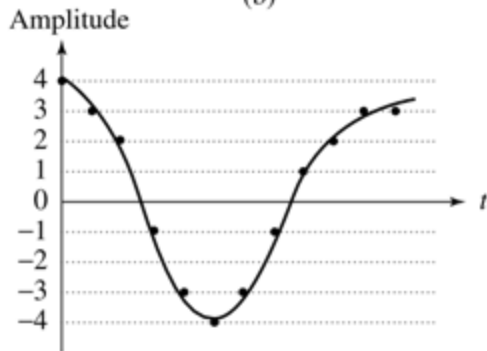
(a)

(a) سیگنال اصلی و نمونه‌های PCM



(b)

(b) سیگنال رمزگشایی شده پله‌ای



(c)

(c) سیگنال بازسازی شده
(با استفاده از فیلتر پایین گذر)

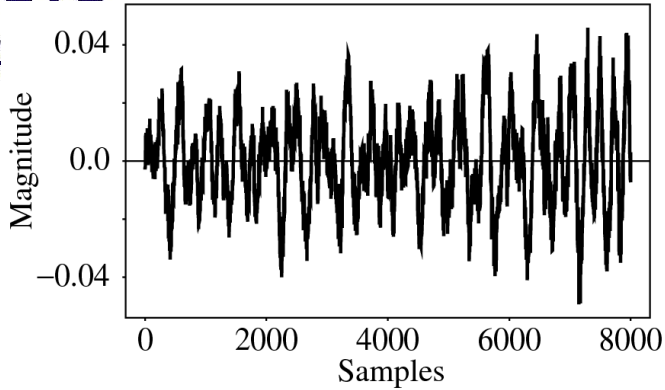


نرخ داده صوتی در کاربردهای متداول (بدون فشرده سازی)

Quality	Sample rate (KHz)	Bits per sample	Mono / Stereo	Data rate (Uncompressed) kb/sec
Telephone	8	8	Mono	64
AM radio	11.025	8	Mono	88.2
FM radio	22.05	16	Stereo	705.6
CD	44.1	16	Stereo	1536

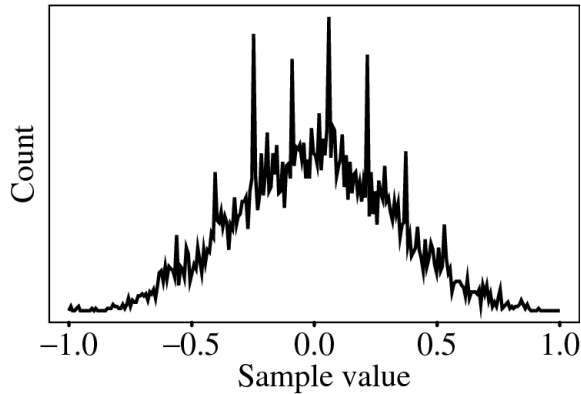


انتقال صوت (کدگذاری تفاضلی صدا)



◀ در طول زمان اطلاعات مشابه فراوانی وجود دارد، که به آن افزونگی زمانی (Temporal Redundancy) گفته می‌شود.

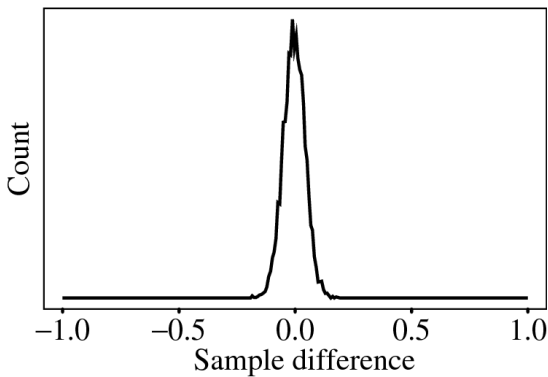
◀ با تفاضل گیری نمونه‌های متوالی سیگنال می‌توان این افزونگی را کاهش داد.



(a) سیگنال اصلی

(b) هیستوگرام سیگنال

(c) هیستوگرام سیگنال تفاضلی





◀ کدگذاری پیش‌بینی‌کننده بدون اتلاف

◀ (Lossless Predictive Coding)

◀ در این روش نمونه بعدی را از نمونه‌های قبل پیش‌بینی می‌کنیم و فقط تفاضل مقدار اصلی و پیش‌بینی‌شده را ارسال می‌کنیم.

$$\hat{f}_n = \lfloor \frac{1}{2}(f_{n-1} + f_{n-2}) \rfloor$$

$$e_n = f_n - \hat{f}_n$$

◀ می‌توان از پیش‌بینی‌کننده‌های دیگری هم با توجه به فضای مساله استفاده کرد.

◀ مثال: برای انتقال رشته $f_1, f_2, f_3, f_4, f_5 = 21, 22, 27, 25, 22$ با کدگذاری پیش‌بینی‌کننده با این رابطه مقادیر خطا به این شکل خواهند بود:

$$f_1 = 21, \quad \hat{f}_2 = 21, \quad e_2 = 22 - 21 = 1$$

$$\hat{f}_3 = \lfloor \frac{1}{2}(f_2 + f_1) \rfloor = \lfloor \frac{1}{2}(22 + 21) \rfloor = 21$$

$$e_3 = 27 - 21 = 6$$

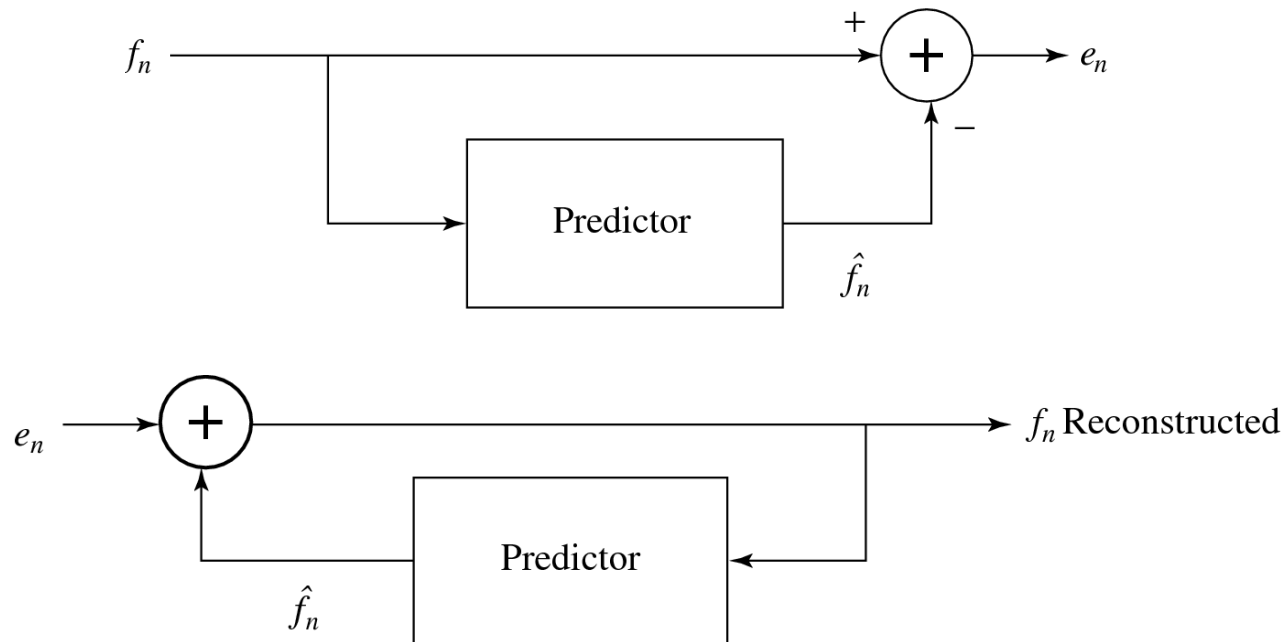
$$\hat{f}_4 = \lfloor \frac{1}{2}(f_3 + f_2) \rfloor = \lfloor \frac{1}{2}(27 + 22) \rfloor = 24$$

$$e_4 = 25 - 24 = 1$$

$$\hat{f}_5 = \lfloor \frac{1}{2}(f_4 + f_3) \rfloor = \lfloor \frac{1}{2}(25 + 27) \rfloor = 26$$

$$e_5 = 22 - 26 = -4$$

◀ شماتیک دیاگرام کدکننده و کدگشای پیش‌بینی‌کننده



دستور wavread

```
clear all  
close all  
clc
```

```
[wavemusic FS] = wavread('music.wav');  
wavplay(wavemusic, FS);  
wavplay(wavemusic, FS/4);  
wavplay(wavemusic, FS*4);  
% FS = sample frequency in Hertz
```

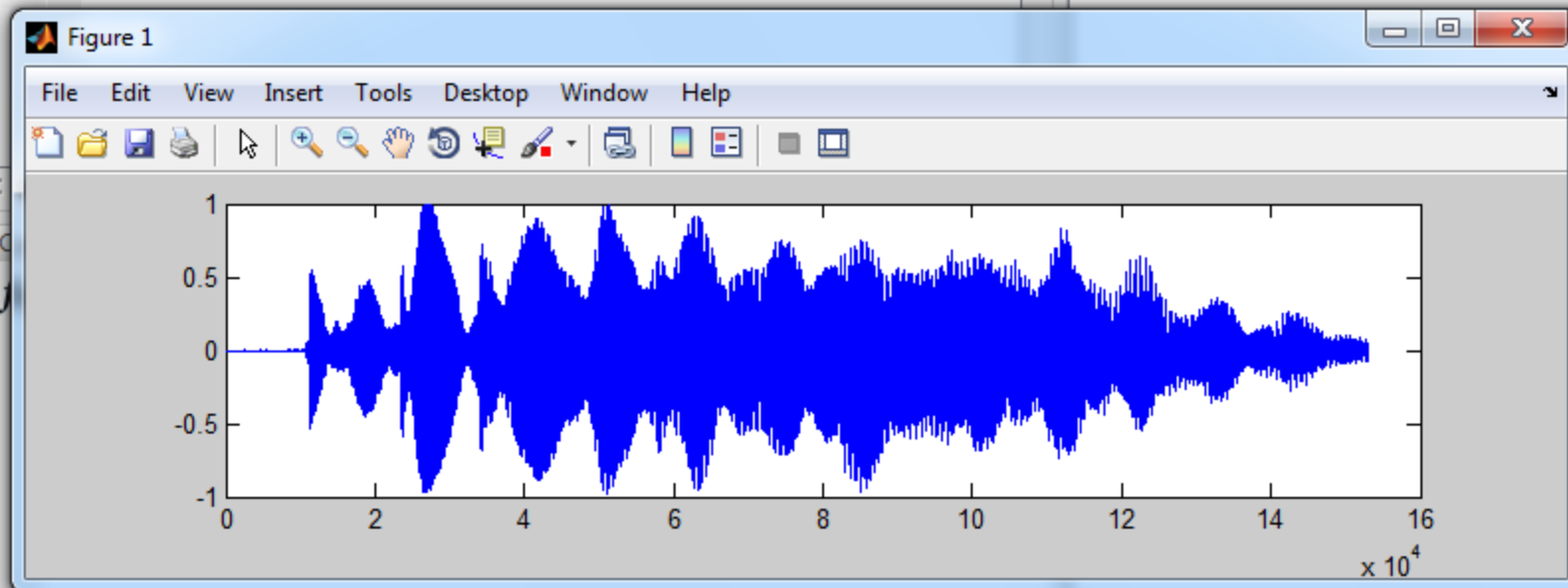
```
figure; plot(wavemusic(:,1));  
-- %
```



موسیقی با فرکانس مختلف و تصویر آن

```
1 clear all
2 close all
3 clc
4 [wavemusic FS] = wavread('music.wav');
5 wavplay(wavemusic, FS);
6 wavplay(wavemusic, FS/4);
7 wavplay(wavemusic, FS*4);
8 % FS = sample frequency in Hertz
9 figure; plot(wavemusic(:,1)');
10 % --
```

FS	48000	48000	48000
wavemusic	<152947x2 double>	-0.9784	1.00





صدای انسان با فرکانس مختلف و تصویر آن

```
1 - clear all
2 - close all
3 - clc
4 - [wavemusic FS] = wavread('sound.wav');
5 - wavplay(wavemusic, FS);
6 - wavplay(wavemusic, FS/1.25);
7 - wavplay(wavemusic, FS*1.75);
8 - % FS = sample frequency in Hertz
9 - figure; plot(wavemusic(:,1)');
```

FS	48000	48
wavemusic	<117542x2 double>	-0

