



# 多媒体技术

## 2.1.1 常见的媒体元素

---

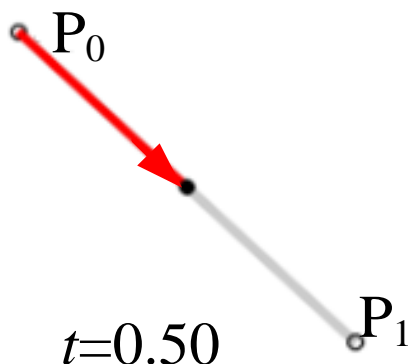
- 贝塞尔曲线
  - 贝塞尔：法国工程师Pierre étienne **Bézier** (1910—1999)的中译名
  - 贝塞尔曲线(Bézier curve): 用数学方法计算的曲线, 用来把多个点连成自由形态的光滑曲线或曲面
  - 只需少量的点就可定义大量的形状, 因此用于逼近一个给定形状时比其他数学方法更有效

## 2.1.1 常见的媒体元素

- 线性贝塞尔曲线

- 假设有两个点 $P_0$ 和 $P_1$ ，线性贝塞尔曲线(linear Bézier curve)是用函数 $B(t)$  描绘的一条直线，在数学上用下式表示：

$$B(t) = P_0 + (P_1 - P_0)t = (1-t)P_0 + P_1t, 0 \leq t \leq 1$$



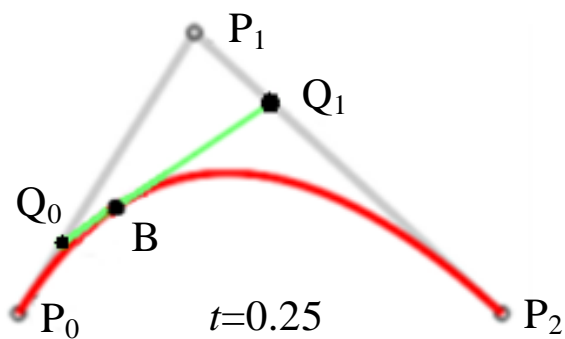
$t$ 用于描述从 $P_0$ 到 $P_1$ 方向走过的路径  
 $t=0.5$ 时 $B(t)$ 从 $P_0$ 向 $P_1$ 走了一半的路程  
当 $t$ 从0到1时，描绘的曲线是一条 $P_0$ 到 $P_1$ 的直线

## 2.1.1 常见的媒体元素

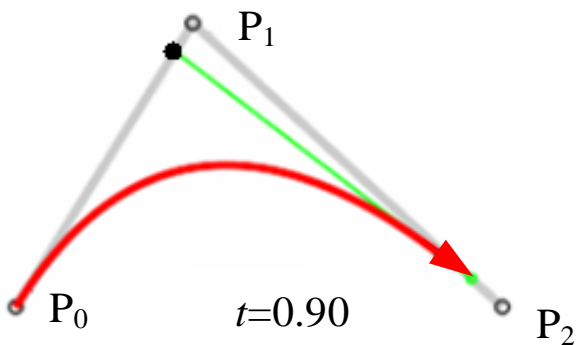
- 二次贝塞尔曲线

- 假设有三个点， $P_0$ 、 $P_1$ 和 $P_2$ ，二次贝塞尔曲线 (quadratic Bézier curve) 是用函数  $B(t)$  描绘的路径，在数学上用下式表示：

$$B(t) = (1-t)^2 P_0 + 2(1-t)tP_1 + t^2 P_2, \quad 0 \leq t \leq 1$$



(a) 二次贝塞尔曲线的构造



(b) 二次贝塞尔曲线的生成

## 2.1.1 常见的媒体元素

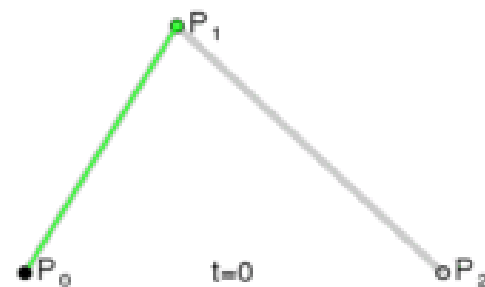
- 二次贝塞尔曲线

- 假设有三个点， $P_0$ 、 $P_1$ 和 $P_2$ ，二次贝塞尔曲线 (quadratic Bézier curve) 是用函数  $B(t)$  描绘的路径，在数学上用下式表示：

$$B(t) = (1-t)^2 P_0 + 2(1-t)tP_1 + t^2 P_2, \quad 0 \leq t \leq 1$$

当  $t$  从 0 变到 1 时：

- ✓ 通过  $P_0$  和  $P_1$  计算出点  $Q_0$ ， $Q_0$  从  $P_0$  变化到  $P_1$  的过程中，描绘一条线性贝塞尔曲线
- ✓ 通过  $P_1$  和  $P_2$  计算出点  $Q_1$ ， $Q_1$  从  $P_1$  变化到  $P_2$  的过程中，描绘一条线性贝塞尔曲线
- ✓ 通过  $Q_0$  和  $Q_1$  计算出点  $B$



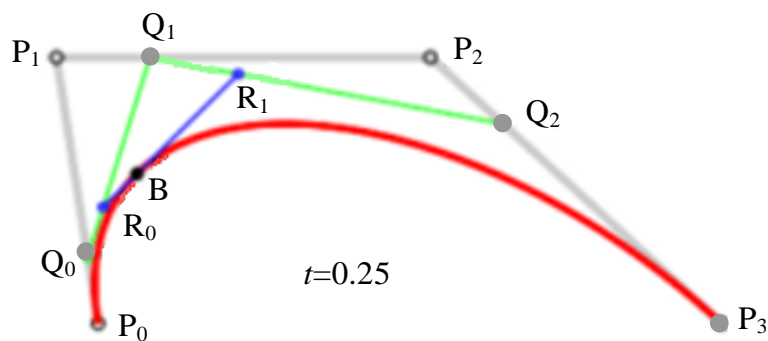
公式推导

## 2.1.1 常见的媒体元素

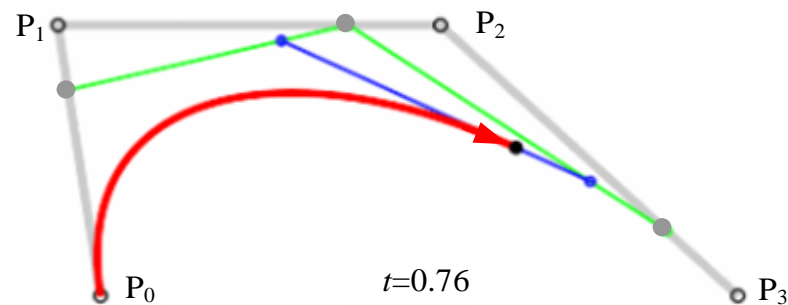
- 三次贝塞尔曲线

- 假设在三维空间的平面上有四个点， $P_0$ 、 $P_1$ 、 $P_2$ 和 $P_3$ ， $P_0$ 到 $P_3$ 的三次贝塞尔曲线(cubic Bézier curve)用函数 $B(t)$ 描绘，在数学上用下式表示：

$$B(t) = (1-t)^3 P_0 + 3(1-t)^2 t P_1 + 3(1-t) t^2 P_2 + t^3 P_3, \quad 0 \leq t \leq 1$$



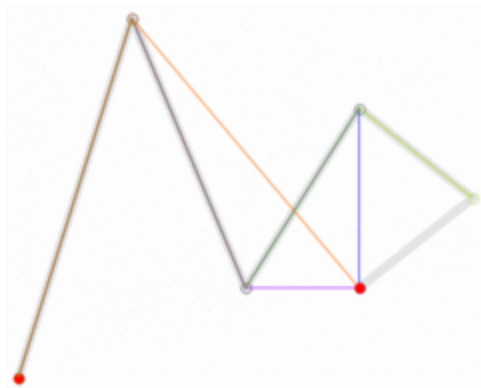
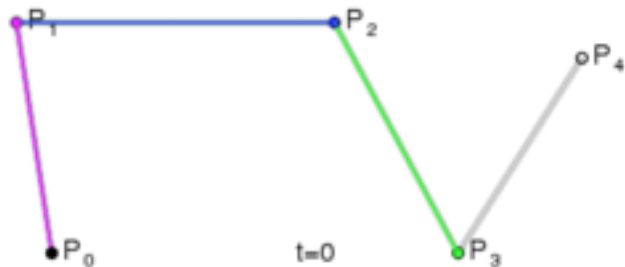
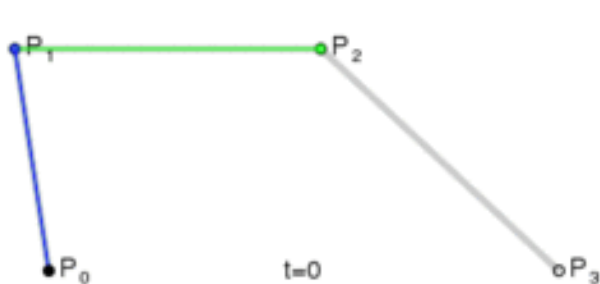
(a) 三次贝塞尔曲线构造



(b) 三次贝塞尔曲线的生成

## 2.1.1 常见的媒体元素

- 三次贝塞尔曲线



## 2.1.1 常见的媒体元素

- 数据可视化

- 将文本数据以**视觉**的形式来**呈现**，如图表或地图等，以帮助人们了解这些数据的意义



全球24小时航班流量图

多 媒体 技术

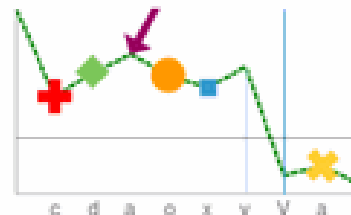
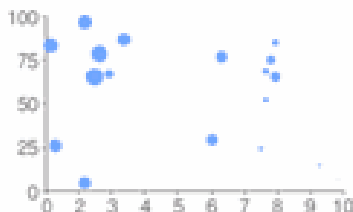
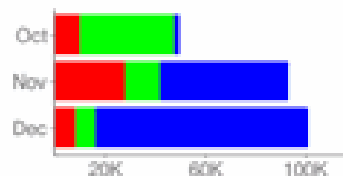
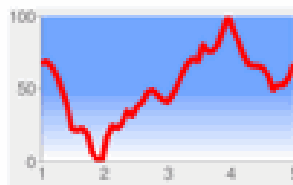


## 2.1.1 常见的媒体元素

- 数据可视化

- Google charts

- Google chart API, 制图服务接口, 用来为统计数据自动生成图片
    - 用于在web上可视化数据, 可在web页面上嵌入 javascript



# 2.1.1 常见的媒体元素

---

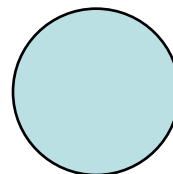
- 图形

- 图形的概念

- 一般指用计算机绘制的画面，如直线、圆、圆弧、矩形、任意曲线和图表等。
    - 图形的格式是一组描述点、线、面等几何图形的大小、形状及其位置、维数的指令集合。
    - 在图形文件中只记录生成图的算法和图上的某些特征点，也称矢量图。

- line ( x1, y1, x2, y2, color )

- circle ( x, y, r, color )



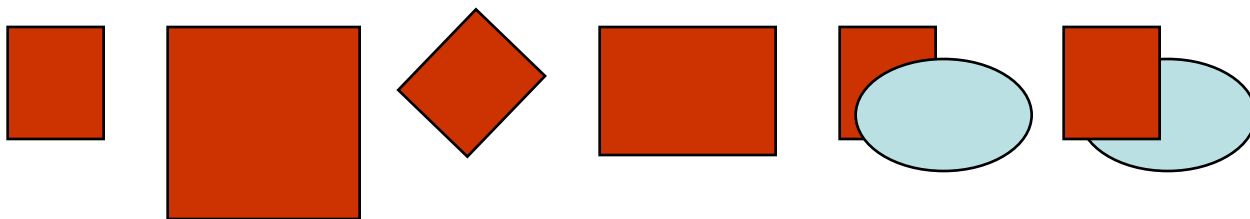
## 2.1.1 常见的媒体元素

---

- 图形

- 图形的特征

- 图形是对图像进行抽象的结果（人工或自动）；
    - 图形的矢量化使得有可能对图中的各个部分分别进行控制（放大、缩小、旋转、变形、扭曲、移位等）；
    - 图形的产生需要计算时间。

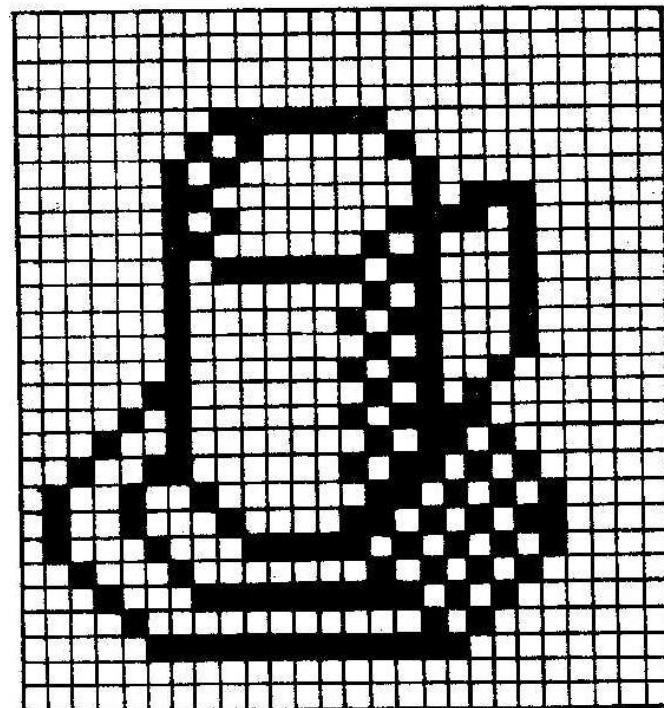


# 2.1.1 常见的媒体元素

- 图像

- 图像的概念

- 图像是指由输入设备捕捉的**实际场景画面**，或以数字化形式存储的任意画面。
    - 静止的图像是一个矩阵，由一些排成行列的点组成，这些点称之为**像素点**（pixel），这种图像称为**位图**（bitmap）。



# 2.1.1 常见的媒体元素

- 图像

- 图像的主要技术参数

- 1. 分辨率

- 每英寸包含的像素总数
      - 水平方向的像素数 \* 垂直方向的像素数
      - 每毫米的线数或行数

- 屏幕分辨率

- 计算机显示器屏幕显示图像的最大显示区
      - 用“水平像素数 × 垂直像素数”表示，如**640 × 480**表示显示屏分成**480**行，每行显示**640**个像素，整个显示屏含有**307200**个显像点。



# 2.1.1 常见的媒体元素

---

- 图像

- 图像的主要技术参数

- 图像分辨率

- 数字化图像的大小，图像精细程度的度量方法
      - 对同样尺寸的一幅图，如果像素数目越多，则说明图像的分辨率越高，看起来就越逼真。相反，图像显得越粗糙

- 像素分辨率

- 像素的宽高比，一般为1：1

- 打印分辨率

- 每英寸能印刷的点数

# 2.1.1 常见的媒体元素

---

- 图像

- 图像的主要技术参数

- 2. 图像深度

- 每个图像的最大颜色数
      - 1位：黑白2色；
      - 4位：16色；
      - 8位：256色；
      - 24位：真彩色
      - 32位：alpha通道，用256级灰度记录图像中的透明度信息，定义透明、不透明和半透明区域

图像文件大小：（高×宽×灰度位数）/ 8 字节

640\*480的256色的图像大小为 $640*480*8/8 = 307000$  Byte

## 2.1.1 常见的媒体元素

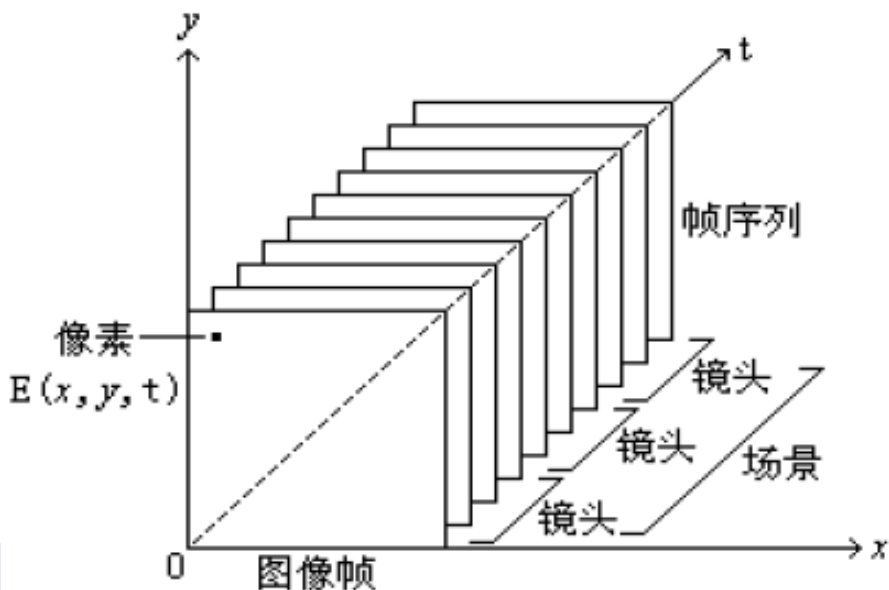
- 视频

- 视频的概念

- 视频源于**电视技术**，它由连续的画面组成。这些画面以一定的速率连续地投射在屏幕上，使观察者具有图像连续运动的感觉。

视频是连续的图像  
图像是离散的视频

视觉滞留效应：人眼在观察景物时，光信号传入大脑神经，需经过一段短暂时间，光的作用结束后，视觉形象并不立即消失，一般1/24秒





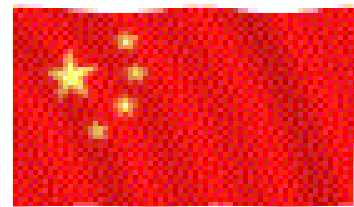
# 2.1.1 常见的媒体元素

---

- 动画

- 动画的概念

- 基本原理与电影、电视相同，利用人眼“视觉滞留效应”的视觉原理，即人眼看到一幅画或一个物体后，在1/24s内不会消失
    - 运动的图画，实质是一幅幅静态图像连续播放。动画的连续播放既指时间上的连续，也指图像内容上的连续，即播放的相邻两幅图像之间内容相差不大。



# 2.1.1 常见的媒体元素

- 音频

- 分类

- 波形声音

- 包含了所有的声音形式，它可以把任何声音都进行采样量化，并恰当地恢复出来

- 语音

- 人的说话声虽是一种特殊的媒体，但也是一种波形，所以和波形声音的文件格式相同。
      - 具有语言内涵和人类约定的特殊媒体

- 音乐

- 符号化了的的声音，乐谱可转变为符号媒体形式

下雨天留客天留我不留



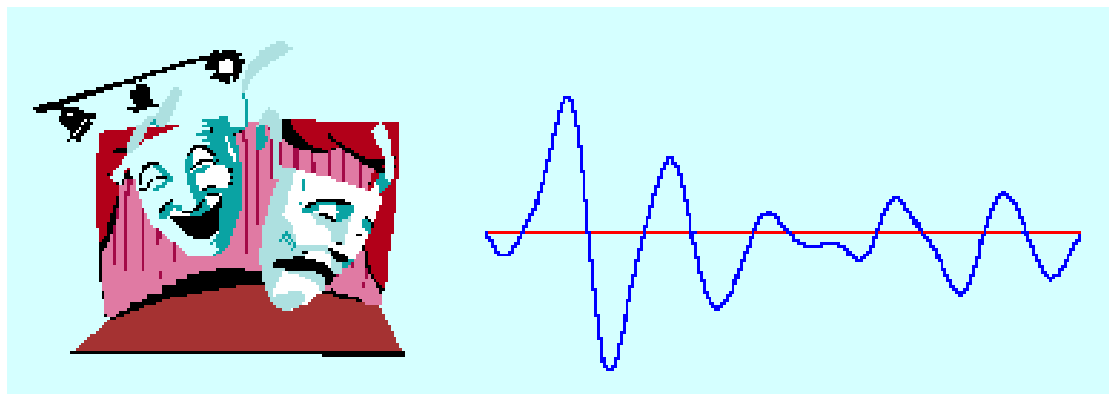
## 2.2 听觉媒体技术

# 2.2.1 声音心理学

## 1. 声音的量纲

声音是听觉器官对声波的感知。

最简单的情况，声音的振动是一种**正弦波**，声音的变化必须确定三件事：**频率**（变化的速度）、**幅度**（产生的压力）、**相位**（何时开始）。



## 2.2.1 声音心理学

---

另外一方面，人们可以感觉到声音的**强弱**，可以感觉到歌唱家音调的**高低**。

- **声音的强弱**主要体现在**声波振动的幅度**大小上。
- **音调的高低**主要体现在**声波振动的频率**上。

因此，声音的量纲分为声音的**物理量纲**和**心理量纲**。

# 声音的量纲

心理变量	首要的物理变量	次要的物理变量
响度	声强度	声波频率
音调	声波频率	声强
音色	声波复合	—
音量	频率和强度	—
密度	频率和强度	—
谐和（流畅或粗糙）	谐波结构	音乐技巧
噪声	强度	频率组合，各种时间参量
骚扰声	多强度技术	频率组合，无意义

## 2.2.1 声音心理学

---

- 物理量纲
  - 可以用精确的值来描述
- 心理量纲
  - 对某一具体声音得来的心理印象却不容易说明白，因为心理印象要由被测者的经验而定
- 两者关系
  - 不是线性的
  - 不是孤立的
  - 不是不变的

## 2.2.1 声音心理学

---

例如：

- 声音的**响度**取决于**强度**和**频率**两个因素，
  - 如果频率不变，强声显得比弱声要响些。
  - 如果强度不变，过高频率的声音和过低频率的声音似乎比中频的声音听起来都要弱一些。
- 更高或更低的频率上，人耳听不到，
  - 由此可见，响度依赖于频率，原因是人耳能反应的频率范围是有上限和下限的。



## 2.2.1 声音心理学

---

- 声音分类

- 按照频率的不同分为

- 次声：频率低于20Hz;
    - 超声：频率高于20kHz;
    - 可听声：位于20Hz-20kHz频率范围内。也称为音频，相应的波形可称为音频信号。
      - 电话语音：20Hz-3.4kHz
      - 调幅广播：50Hz-7kHz
      - 调频广播：20Hz-15kHz
      - 宽带音频：20Hz-20kHz

## 2.2.1 声音心理学

---

- 声音的**物理特性**
  - 频率、声压和声强、动态范围、频谱.....
- 声音的**心理特性**
  - 音调、响度、音色、掩蔽效应、方位感、空间感.....
- 声音的**音质**
  - 频带宽度、信噪比、数据量.....

## 2.2.1 声音心理学

---

- a. 音强

- 又叫响度，表示声音能量的强弱程度，主要取决于声波振幅的大小。一般用声压或声强来计量，单位是分贝（dB）。
- 常说的“音量”也是指响度。

## 2.2.1 声音心理学

---

- 声压

- 声波引起某处媒质 **压强的变化量** 称为该处的声压。（单位为 **Pa**（帕斯卡），即  $\text{牛顿}/\text{米}^2$  是压强的量纲）

## 2.2.1 声音心理学

- 声压级

— 将声压的有效值以对数的形式表示声音强弱的数值称为声压级。

$$SPL = 20 \lg ( P_{rms} / P_{ref} ) \quad (\text{单位用分贝dB})$$

$P_{rms}$  : 计量点的声压有效值

$P_{ref}$  : 零声级的参考声压值 ( $P_{ref}=2 \times 10^{-5}$ 帕)

## 2.2.1 声音心理学

- 声压级

— 将声压的有效值以对数的形式表示声音强弱的数值称为声压级。

$$SPL = 20 \lg ( P_{rms} / P_{ref} ) \quad (\text{单位用分贝dB})$$

1、 $P_{ref}$ 为具有正常听力的年轻人对1kHz的声音刚好能察觉的声压值。

2、声压级实际上是一种相对量，是某点的声压与零声压的比，是描述声音变化的动态范围的物理量。

## 2.2.1 声音心理学

- 声强

- 是在传播方向上，单位时间内通过单位面积的声能量，记为 $I$ ，单位为瓦/米<sup>2</sup>，W/m<sup>2</sup>。

- 声强级

- 表示声强与参考声强 $I_0$ （取 $I_0 = 10^{-12} \text{W/m}^2$ ）的相对关系，记为 $L_I$ ，即：

$$L_I = 10 \lg \frac{I}{I_0} \quad (\text{dB})$$

## 2.2.1 声音心理学

---

- b. 音调

- 音乐也称音高，表示人耳对声调高低的主观感受。客观上音高大小主要取决于声波基频的高低。
- 基频越低，给人的感觉越低沉。基频频率增加一倍，音乐上称提高了一个八度。



## 2.2.1 声音心理学

---

- 在任一时刻  $t$ ，声波可以分解成一系列正弦波的线性叠加：

$$f(t) = \sum_{n=0}^{\infty} A_n \sin(n\omega t + \varphi_n)$$

- $A_n$  是振幅，表示声音的强弱
- $\omega$  是声波的基频，表示声音音调的高低
- $\varphi_n$  是  $n$  次谐波的初相位

## 2.2.1 声音心理学

- 基频

- 根据傅里叶变换的理论，声音可以分解成若干个不同频率纯音的叠加。
- 这些频率都是某一频率的倍数，这一频率就称作基频，也就决定了这个音的音高。

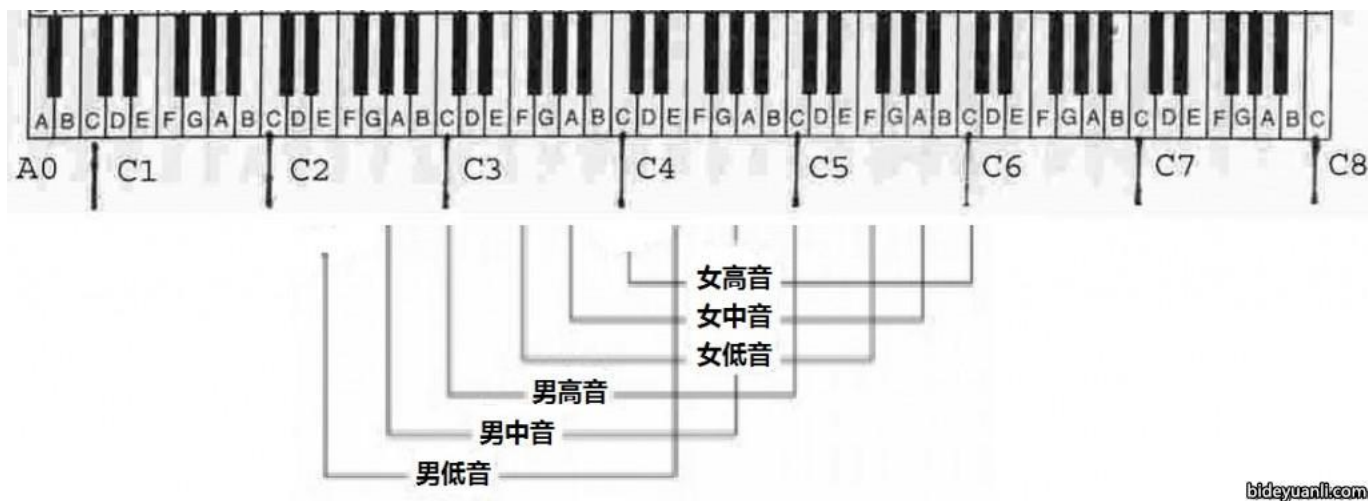
- 基音

- 基频所产生的音称为基音。（频率最低）

- 泛音

- 基频为 $f$ ，则频率为 $2f$ 的音称为第一泛音，频率为 $3f$ 的音称为第二泛音。
- 频率为基频整数倍的正弦，振荡为谐波。基波频率3倍的波称之为三次谐波。

## 2.2.1 声音心理学



- 一个人第一泛音可以和另一人的基频叠合一起;
- 女声比男声高八度;

## 2.2.1 声音心理学

---

- c. 音色

- 又称音品，由声音波形的谐波频谱分布决定。
- 在相同响度及音高情况下，辨别不同的乐器，或者不同的说话的人，就是因为这些声源产生的频谱的差异。

## 2.2.1 声音心理学

### • 频谱

- 对一个声源发出的声音的**频率**成份和**强度**的分析，叫频谱分析，并且可用**频谱图**表示，其**纵轴**为**幅度**，**横轴**为**频率**。
- 将声音的**基频**和各次**谐波**按**频率**分别用一竖线画在横轴上，它的长短表示幅度的大小。

