1. 时频分析环节在识别类应用中对应特征提取，在编码合成类应用中对应压缩表示机理
2. 语音是非平稳信号，但可以认为在一个较短的时间内是平稳的，因为在较短的时间内声管形状不会有大的改变。一般用短时窗截取一段信号进行分析，并在信号中以一定步长向前滑动。这一过程也被称为分帧，帧长一般取10~30ms，滑动步长也被称为帧移，每次滑动前一窗与后一窗信号间重叠部分的长度被称为帧叠。帧叠是帧长的一半是一种常见的分帧设计，可以保证相邻帧间分析结果的连续性。时域分析常用方窗，保证采样值的真实性。频域分析常用Hamming窗和Hanning窗（窗函数见课件）。DFT变换需要对信号做周期性延拓，一帧信号通常不是周期性的，因而频率分析时很容易产生吉布斯效应，导致频谱失真。分帧时加Hamming窗或Hanning窗可以缓解这一现象。
3. 短时能量和短时过零率是两个常用的时域特征，计算公式见课件。短时能量刻画了声音的强弱，过零率是在时域上对声波的振动情况进行了粗略的刻画，反映的还是频率信息，为了避免直流噪声、背景噪声等影响过零率的计算结果，常采用门限过零率。在语音识别算法中，能量经常是声学特征向量中的一维特征，它对区分当前帧是语音还是背景音，是元音还是辅音有重要的作用。能量的变化轨迹也对识别语音的韵律模式有一定的参考意义。过零率统计特征对区分语音和音乐有帮助。
4. 在现实世界中背景声音持续存在，识别算法需要从连续的声音数据中找到语音（或其它目标声音）的起始点和结束点，这一任务被称为“端点检测”。也常被称为“语音活动检测”（Voice Activity Detection, VAD）。分析可知，元音的能量高于辅音，辅音的能量与背景音的能量没有太大区别。但辅音的过零率明显高于背景音。利用这一特性，可以设计基于短时能量和短时过零率的双门限端点检测方法。实际上算法中用了三个门限，用了一个较高的能量门限确保当前信号一定是语音，而不会是某种能量较大的环境声音。再用一个较低能量门限找到元音的起点和终点。最后通过一个过零率门限找到辅音的起始点和终点。算法过程见课件。算法中的门限值一般根据当前信号的情况自适应的确定。还有很多其它的端点检测方法。
5. 掌握基频和基音周期的概念，掌握基于短时自相关函数的基音频率检测方法。元（浊）音有周期性，辅（清）音没有周期性。如果一个信号有周期性，其自相关函数一般也有周期性。元音的自相关函数会在基音周期的整数倍处出现峰值，设第一个峰值处为，它就对应基音周期，所以语音的基频等于采样频率除以。计算语音信号的自相关函数，然后在一定范围内搜索最大值，该最大值的位置就是，限定搜索范围是为了避免把倍频误当作基频，比如可以在0~900的范围内搜索基频。清音的自相关函数没有明显的周期性，表现为最大值处的自相关函数值小于等于能量的四分之一，也就是*R*(0)的四分之一。基于短时自相关函数的基频检测算法见课件。还有很多其它的基频检测算法，如平均幅度差法。
6. 语音内容的可分性主要蕴含在其频域表示中。较长语音段的声学特性的可视化频域分析方法是语谱图。它是较早出现的基于专用硬件设备实现的频域分析方法，可以看做是由每帧频谱拼接而成。它是一种包含三个维度信息的可视化表示，横坐标是时间，纵坐标是频率，通过灰度值或颜色值表示强度（特定频带能量的大小）。窄带语谱图和宽带语谱图的帧长不同，窄带语谱的帧长值较大，从而傅里叶分析后频线较多，带宽较窄。窄带语谱图的频域分辨率高、语谱图上横的线条明显。宽带语谱图的时域分辨率高、可以看见语谱图上纵的线条。
7. 分析人类的发声器官和发声机理，声道对声门波进行了频率调制，从而产生了不同语音内容。研究者提出基于同态解卷来得到声道系统的单位冲激响应的前提是，假定声门脉冲已知，一般假定其为准周期三角波。在之后的研究工作中证明这一假定过于粗糙了。
8. 掌握倒谱域、复倒谱、倒谱的概念，倒谱域是一种伪时域，无实际的物理意义。对声音信号。

复倒谱计算为：



倒谱计算为：



复倒谱是可逆变换，倒谱不是可逆变换。倒谱具有很好的压缩特性和鲁棒性， 因而在语音声学特征中常被采用。

1. 掌握语音信号产生的系统模型，见课件。将声道简化为若干节均匀界面无损声管的级联，给出了声道滤波器为全极点模型的假设。声门激励对元音和辅音不相同，元音是以基音周期为周期的冲激序列，再经过声门波整形得到的。辅音则是随机噪声。声道滤波器的控制参数就是全极点模型的系数。辐射模型则仿真了口唇的辐射作用。
2. 求取全极点声道滤波器系数的思路：假定其是AR模型这种特殊的全极点模型，然后通过求取最佳线性预测器来求AR模型系数。掌握求解线性预测系数的方程组的推导方法。在对分析窗进行假定的条件下，将该方程组简化为yule-walker方程组，并给出快速求解算法：Levinson—Durbin递推算法。掌握其推衍参数线性预测倒谱系数的快速计算方法。详细内容见课件。
3. 掌握人类的听觉感知机理，以及 Mel尺度下的听觉感知模型，掌握 Mel频域倒谱系数的快速计算方法。详细内容见课件。