

# 数字语音处理

王晗 wanghan@bjfu.edu.cn





# 本节目标和任务

A

学习理解语音发声和听觉机理及其数字模型；掌握语音信号处理的基本概念、原理和方法，为以后的学习提供理论依据；掌握语音信号的短时域分析和频域分析方法，并且能够用Matlab或Python语言编程实现对实际语音信号进行的处理；熟悉语音信号处理的应用领域，并了解该领域的新理论、新技术和发展趋势，为今后的进一步学习和工作打下坚实的基础



# CONTENTS

A { 语音信号处理的发展

B { 语音信号处理的应用

C { 语音信号处理的过程

D { 语音信号处理编程



# 1. 语音信号处理的发展





# 概述



语音信号处理是一门新兴的边缘学科，是语音学与数字信号处理两个学科相结合的产物。它和认知科学、心理学、语言学、计算机科学、模式识别和人工智能等学科有着紧密的联系。

语音信号处理的目的：

得到某些语音特征参数以便高效地传输或存储；

通过某种处理运算以达到某种用途的要求。

# 概述



语音信息的交换大致上可以分为三大类：

(1) 人与人之间的语言通信：

包括语音压缩与编码、语音增强等。

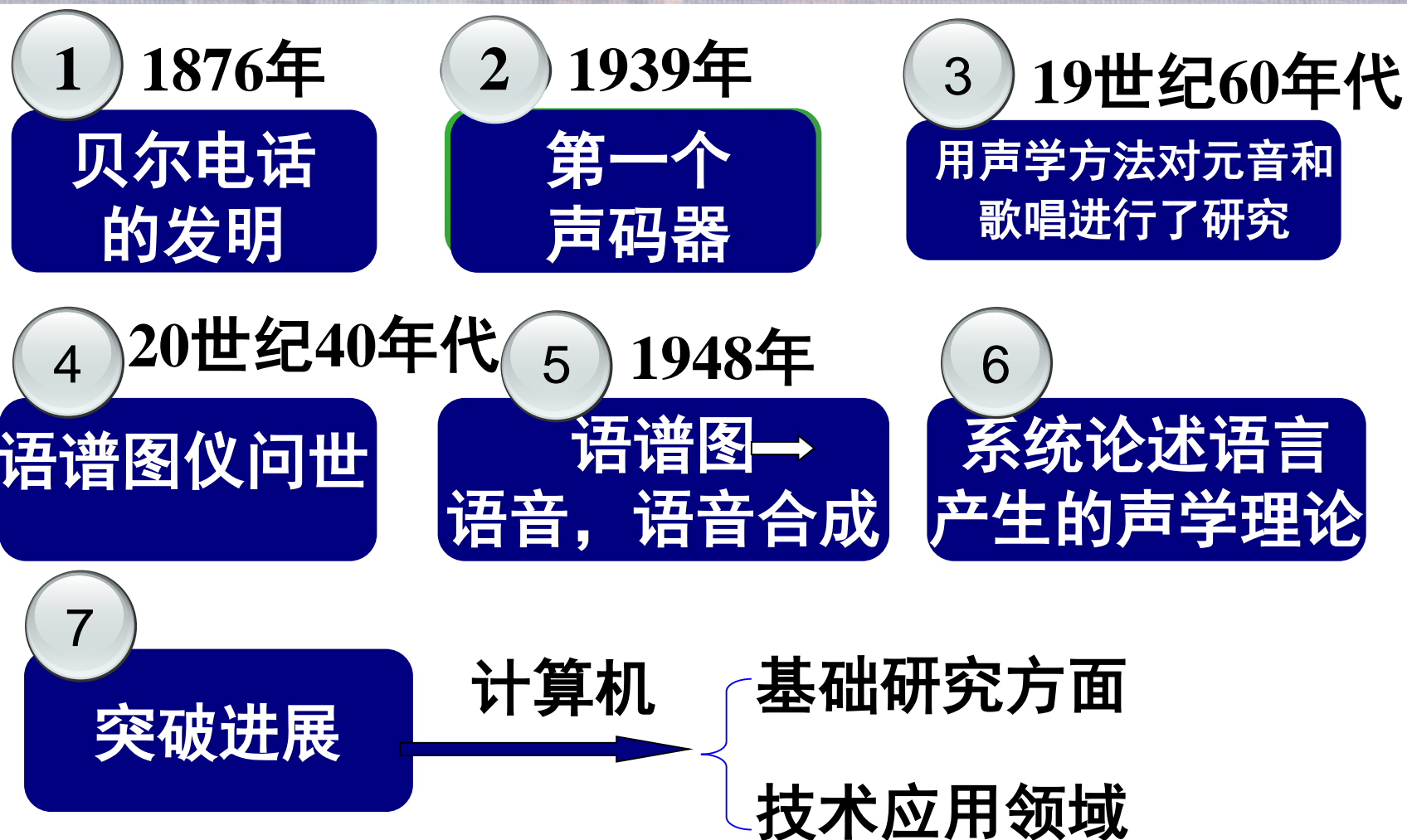
(2) 第一类人机语言通信问题：

语音合成，机器讲话、人听话的研究。

(3) 第二类人机语言通信问题：

语音识别和理解：人讲话、机器听话的情况。

# 时间轴





# 语音信号处理的发展

A

语音合成

B

语音编码

C

语音识别



# 语音合成

## 语音合成的目的：

**让计算机能够产生高清晰度、高自然度的连续语音。**

以语言信息压缩、存储为主要目的对语音信号数字模型进行研究，同时研究音素、音节、词组与句子的发音规则。最终恢复出自然流畅的语音来。  
例如文语转换系统(Text-to-speech)。语音分析与合成将赋予计算机说话的功能。  
也是进行语音编码、语音识别研究的基础。



# 语音编码

## 语音编码的目的：

在保证一定语音质量的前提下，尽可能降低编码比特率，以节省频率资源。



# 语音编码

## 语音编码与文本到语音转换两个主要区别

一、前者是人与人之间的话音交流，要保留说话人的声音特征。后者是文本到声音的转换即计算机发声。它可以是标准播音员或其它声音。

二、前者不仅对压缩率和音质有要求，而且要求较低的编、解码延迟。而后者对处理帧长没有什么太严格的限制。



# 语音识别

研究如何使**计算机能够听懂人类的语言**。以汉语语音为例：汉语约有400个音节，加上声调约1200个音调节，把这些语音信号的特征存储到计算机内，并与计算机接收到的汉语发音进行比较，找到特征相同的音节或音调节，这个过程就是语音识别。将识别出的音节序列转换成文字，就是语言理解。许多算法将理解过程溶入到识别中来提高识别的准确性。因此可以将语音识别与理解归入同一类应用。



# 语音识别

## 语音识别研究领域包括：

- (1) 根据对说话人说话方式的要求，可以分为孤立字语音识别系统，连接字语音识别系统以及连续语音识别系统。
- (2) 根据对说话人的依赖程度可以分为特定人和非特定人语音识别系统。
- (3) 根据词汇量大小，可以分为小词汇量、中等词汇量、大词汇量以及无限词汇量语音识别系统。

[Demo1](#)

[Demo2](#)

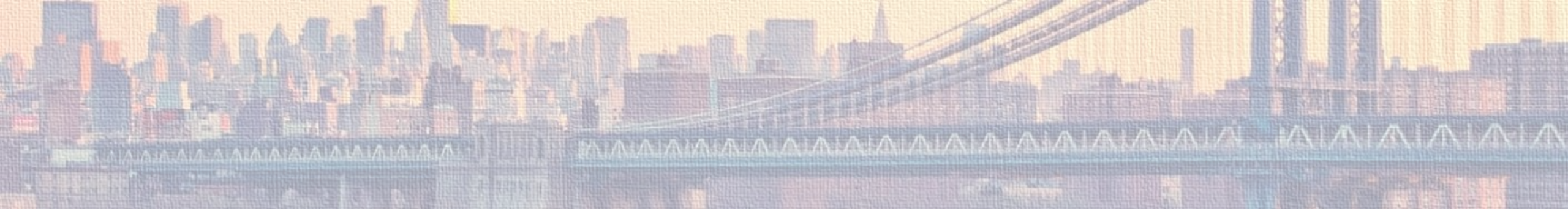


# 语音信号处理的应用及新方向

## 一：语音信号处理的应用

- 语音技术的应用领域
- 包括语音识别、说话人的鉴别和确认、语种的鉴别和确认、关键词检测和确认、语音合成、语音编码等。
- 但其中最具有挑战性和最富有应用前景的为语音识别技术。

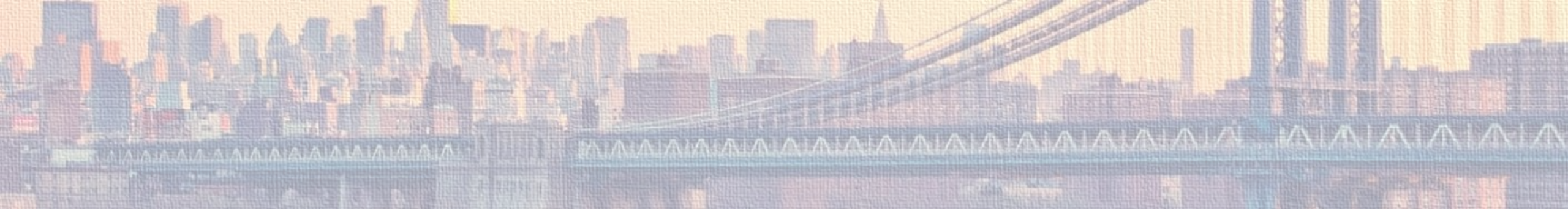




## 1. 说话人识别技术的应用

- 安全加密、银行信息电话查询服务
- 公安机关破案和法庭取证方面
- 在声控应用中，识别输入的语音内容，并根据内容来执行相应的动作，这包括了声控电话转换、声控语音拨号系统、声控智能玩具、信息网络查询、家庭服务、宾馆服务、旅行社服务系统、医疗服务、股票查询服务和工业控制等。





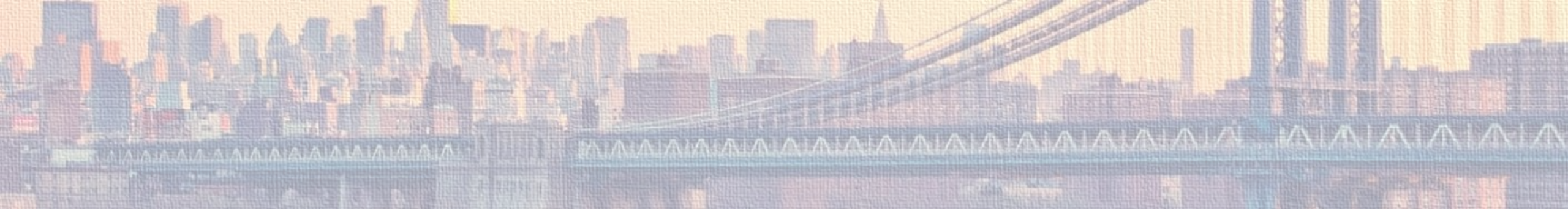
- ## 1. 说话人识别技术的应用
- 在电话与通信系统中，智能语音接口正在把电话机从一个单纯的服务工具变成为一个服务的“提供者”和生活“伙伴”。
- 目前，自动口语分析、用户交换机、电话机、手机已经包含了语音识别拨号功能，还有语音记事本、语音智能玩具等产品也包含了语音识别与语音合成功能。人们可以通过电话网络用语音识别口语对话系统查询有关的机票、旅游、银行信息。



## ➤ 2. 语音合成的应用

- 公交汽车上的自动报站、各种场合的自动报时、自动报警、手机查询服务和各种文本校对中的语音提示、在电信声讯服务：如股票、售后服务、车站查询等信息；也可用于基于微型机的办公、教学、娱乐等智能多媒体软件，例如语言学习、教学软件、语音玩具、语音书籍等；也可与语音合成技术与机器翻译技术结合，实现语音翻译等。





### ➤ 3. 语音编码的应用

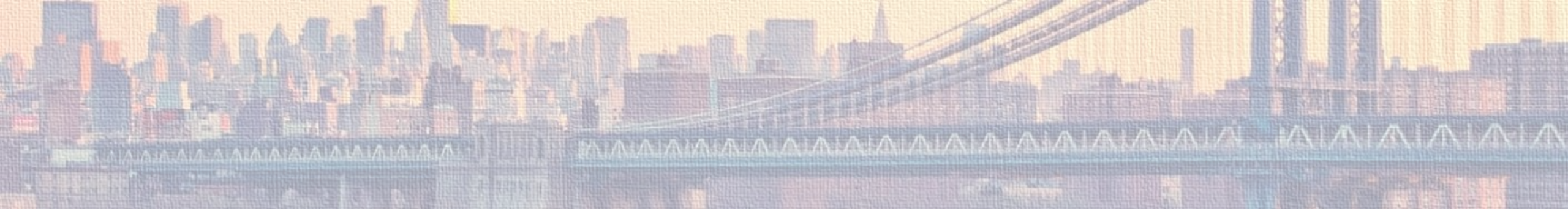
- 在移动通信、卫星通信、军事保密通信、信息高速公路和**IP**电话通信中的应用。
- 在个人移动通信、语音存储、多媒体通信、数字数据网(**DDN**)中的应用。
- 未来的**ISDN**、卫星通信、移动通信、微波接力通信和信息高速公路以及保密电话等中的应用。



## ➤ 二、语音信号处理的新方向

- (1) 基于语音的信息检索。网络技术及数字图书馆 技术的发展，基于语音识别的信息检索技术正成为研究热点。
- (2) 基于语音识别的广播新闻的自动文摘技术的研究。由于广播、电视中的发音较为标准规范，在识别中避免了说话人发音上的不规范，有利于语音识别系统性能的提高。





- 二、语音信号处理的新方向
- (3) VoIP技术。是通过TCP/IP网络，而不是传统的电话网络来传输语音的新的通信方式，通常称之为IP电话技术。在网络上对压缩的语音数据以数据包的形式进行传输和识别。
- (4) 语音训练与校正技术。现在越来越多的人希望掌握其他非母语语言，以便方便的进行交流。因此语言学习机已成为当今外语学习者的有利工具。

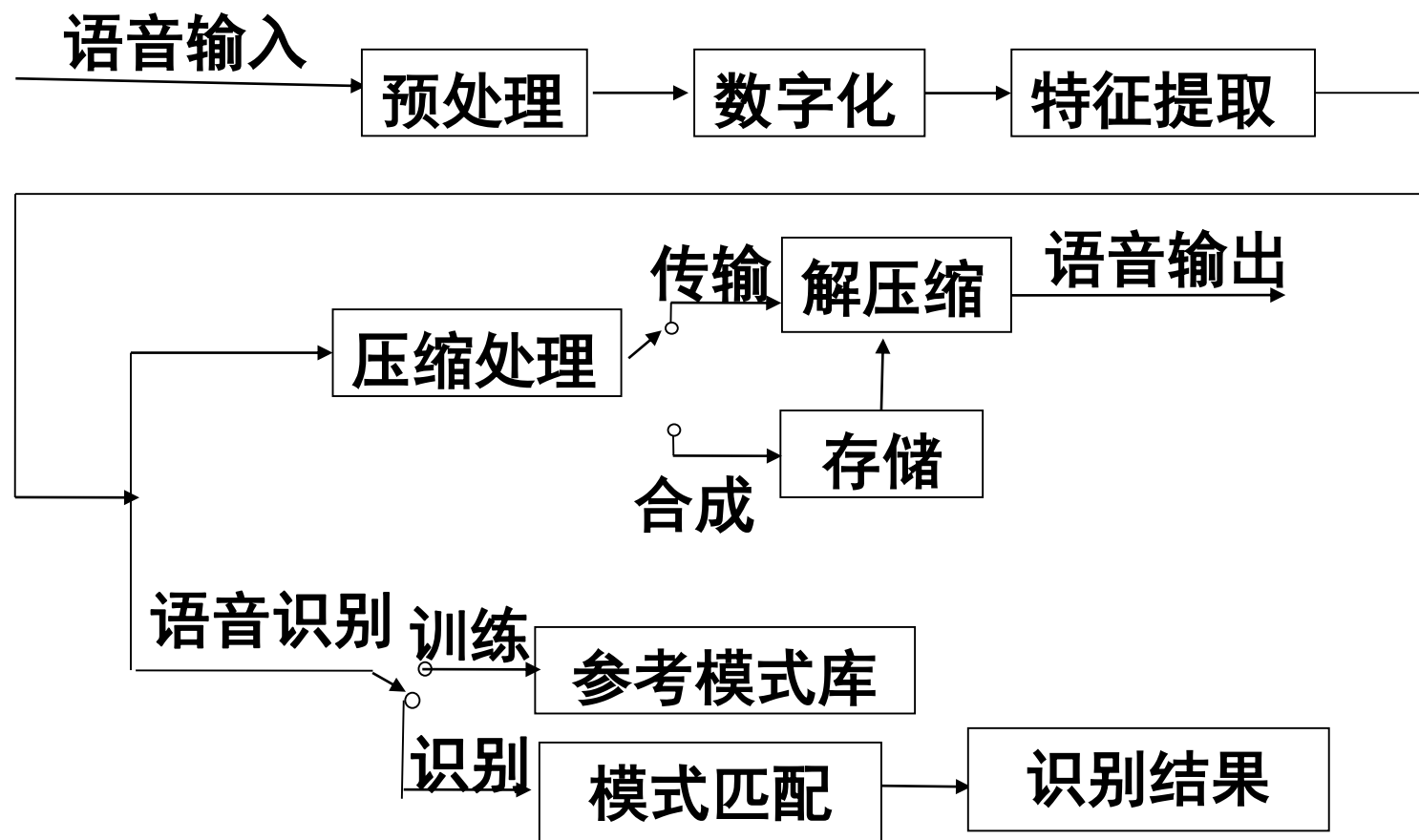


## ➤ 二、语音信号处理的新方向

- (5) 语种识别。是近年来新出现的研究方向，通过分析处理一个语音片断来判别其所属语音的种类，本质上属于语音识别的研究范畴。
- (6) 基于语音的情感处理研究。在人与人的交流中，除了语音信息外，非语言信息也起着重要的作用。为了使人机交流更自然、更人性化，基于语音的情感处理研究也是非常必要的。



# 语音处理过程的结构框图





**THANK  
YOU**

