

HMM算法在语音识别中的应用——算法学习

原创

2016年06月15日 14:04:37

1942



0

总体框架



yzbx

原创

97

粉丝

20

喜欢

9

评论

14

等级：博客5

访问量：11万+

积分：2045

排名：2万+



少儿编程



广告

他的最新文章

更多文章

Densely Connected Convolutional Networks (DenseNet 2017 CVPR)

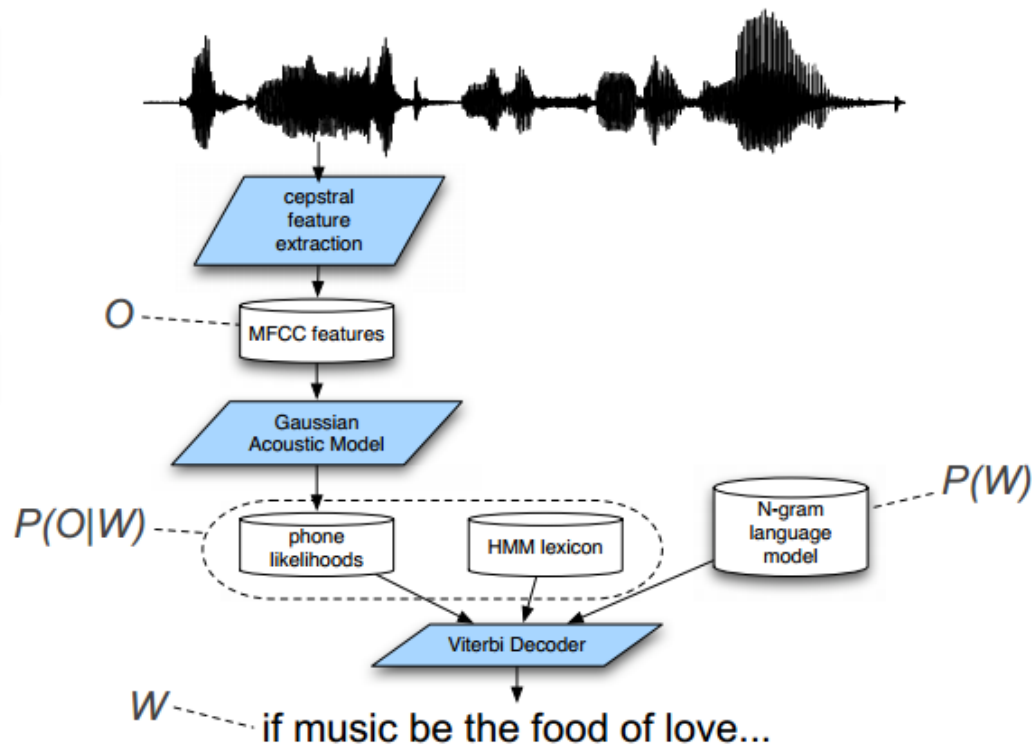
加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

注册



Speech Recognition Architecture



输入

Treat acoustic(听觉的) input O as sequence of individual observations

$$O = o_1, o_2, \dots, o_t$$

ubuntu (linux) 开发者初始化要做的第K件事

Linux 设备信息登记

Qt 开源版安装与下载

文章分类

IDE使用经验	7篇
反防盗链	2篇
php	1篇
文本编辑器	1篇
工具使用记录	23篇
算法学习	27篇

展开

文章存档

2018年1月	2篇
2017年11月	1篇
2017年9月	2篇
2017年5月	1篇
2017年4月	1篇
2017年3月	2篇

展开

Define sentence as a sequence of words

$$W = w_1, w_2, \dots, w_n$$

判断模型

- 最大概率: $W = \arg \max_{W \in L} P(W|O)$
- 贝叶斯: $W = \arg \max_{W \in L} \frac{P(O|W)P(W)}{P(O)}$
- 化简: 由于 $P(O)$ 对所有 W 一样, $W = \arg \max_{W \in L} P(O|W)P(W)$

模型

html 利用 frameset 进行简单的框架布局

📖 17663

firefox + foxypoxy + shadowsock in ubuntu

📖 6950

算法学习 -- Staple: Complementary Learners for Real-Time Tracking

📖 4787

按键精灵自动发QQ消息

📖 4364

Dot 与 GraphViz 经验总结

📖 4329

前景提取--PBAS算法

📖 3206

前景提取--SuBSENSE算法

📖 2878

2016 mingw 5.3.0 + Qt 5.7.0 + opencv 3 under windows7

📖 2412

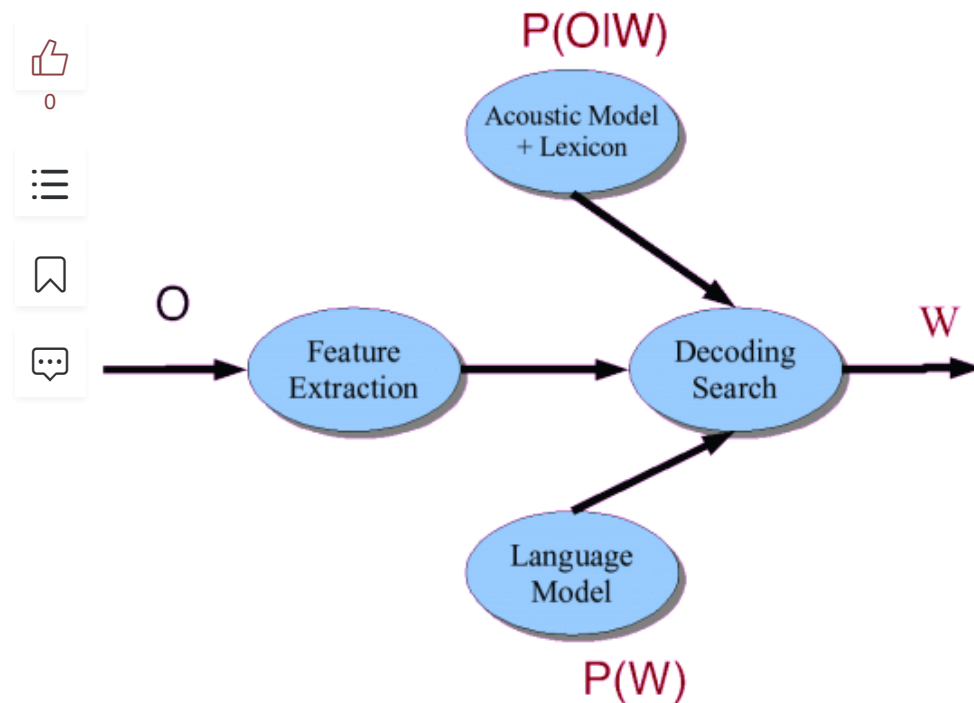
opencv simpleBlobDetector入坑

📖 2229

前景提取--color 算法

📖 2192

Speech Architecture meets Noisy Channel



- Feature Extraction: 39 “MFCC” features
- Acoustic Model: Gaussians for computing $p(o|q)$



4万的二手车



联系我们



请扫描二维码联系客服

✉ webmaster@csdn.net

☎ 400-660-0108

🗣 QQ客服 🗣 客服论坛

关于 招聘 广告服务 百度

©1999-2018 CSDN版权所有

京ICP证09002463号

经营性网站备案信息

网络110报警服务

中国互联网举报中心

北京互联网违法和不良信息举报中心

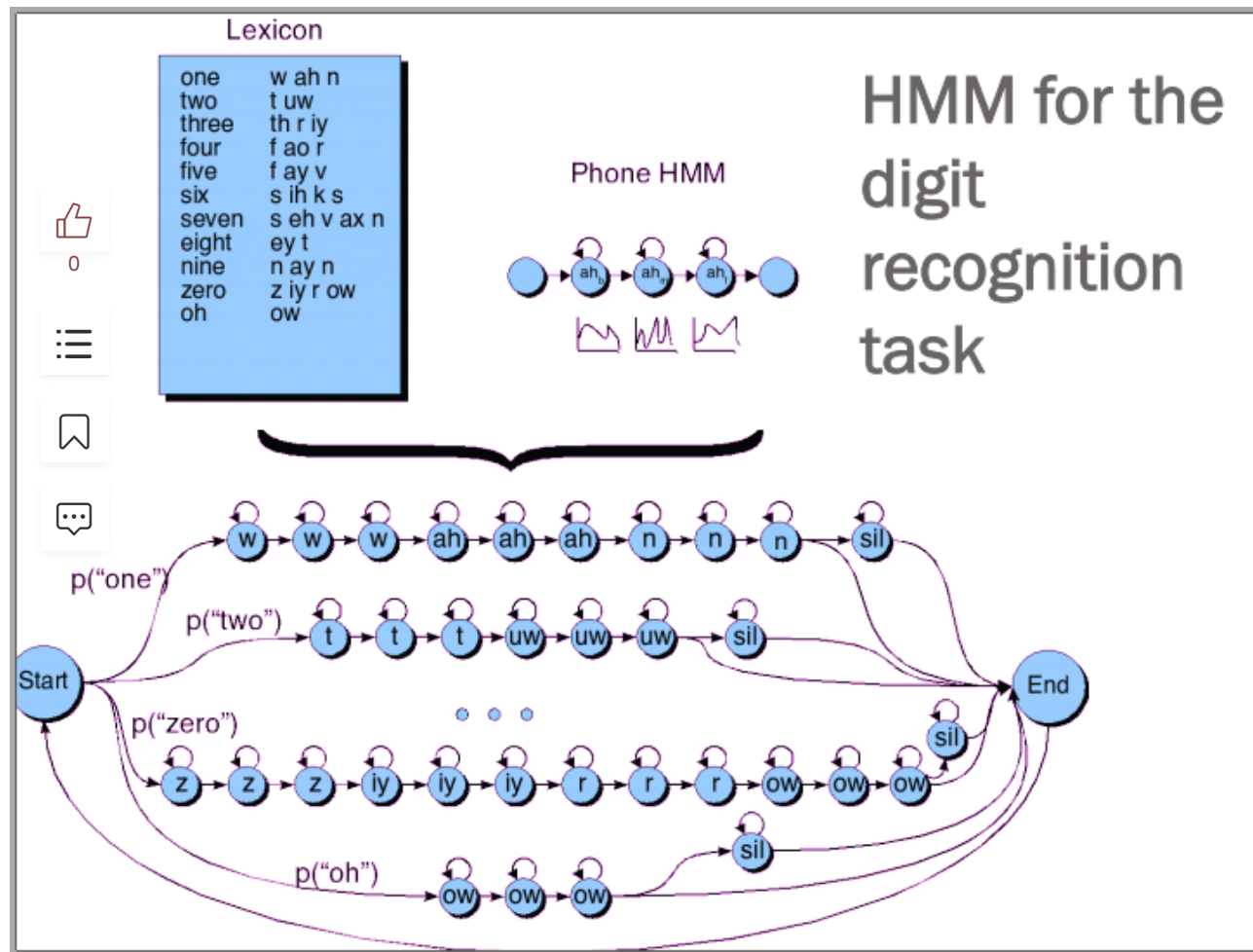
加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

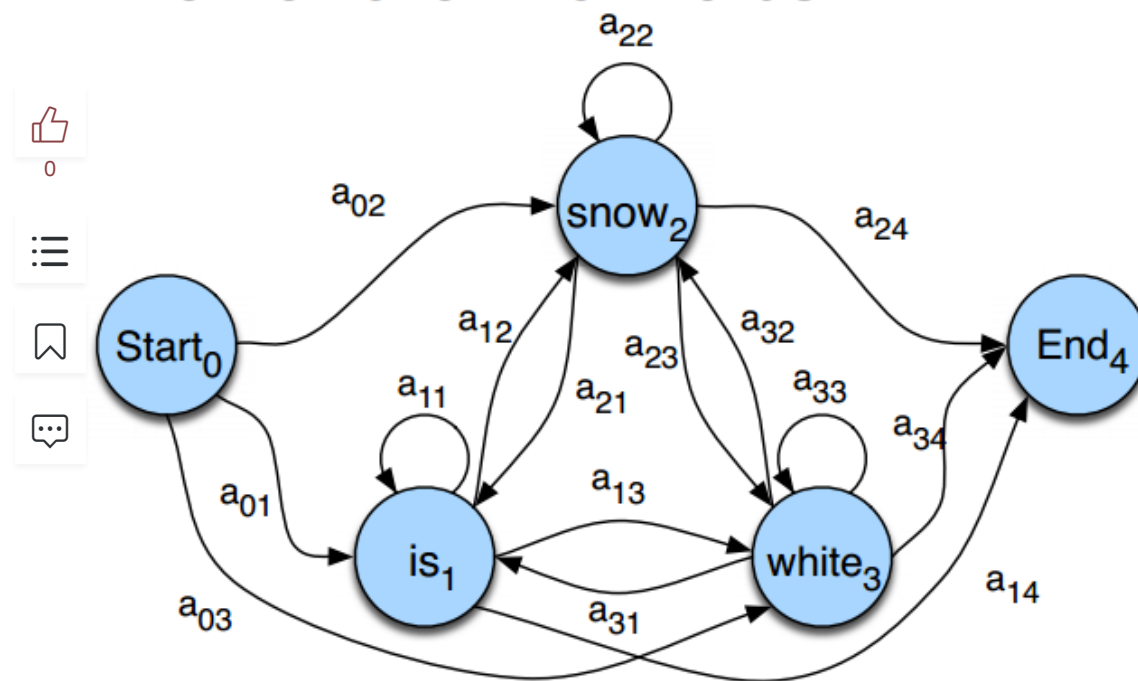
注册



- Lexicon(词典)/Pronunciation(发音) Model: HMM, what phones can follow each other



Markov chain for words



- Language Model: N-grams for computing $p(w_i | w_{i-1})$

markov chian

- states: $Q = q_1, q_2, \dots, q_N, q_t$ is the state at time t.
- transition probability: $A = [a_{11}, a_{12}, \dots, a_{NN}]$

加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

注册



- $a_{ij} = P(q_{t-1} = i | q_t = j)$
- $\sum_{j=1}^N a_{ij} = 1, \text{ for } i \in [1, N]$

- markov assumption:

- $P(q_i | q_1 q_2 \dots q_{i-1}) = P(q_i | q_{i-1})$

- initial status

- $\pi_i = P(q_1 = i)$

- $\sum_{i=1}^N \pi_i = 1$

hidden markov chain



Hidden Markov Models

$Q = q_1 q_2 \dots q_N$

a set of N **states**

$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$

a **transition probability matrix** A , each a_{ij} representing the probability of moving from state i to state j , s.t. $\sum_{j=1}^n a_{ij} = 1 \quad \forall i$

$O = o_1 o_2 \dots o_T$

a sequence of T **observations**, each one drawn from a vocabulary $V = v_1, v_2, \dots, v_V$

$B = (b_{ij}(o_t))$

a sequence of **observation likelihoods**, also called **emission probabilities**, each expressing the probability of an observation o_t being generated from a state i

q_0, q_F

a special **start state** and **end (final) state** that are not associated with observations, together with transition probabilities $a_{01} a_{02} \dots a_{0n}$ out of the start state and $a_{1F} a_{2F} \dots a_{nF}$ into the end state

three problem

The Three Basic Problems for HMMs

Jack Ferguson at IDA in the 1960s



Problem 1 (Evaluation): Given the observation sequence $O=(o_1o_2...o_T)$, and an HMM model $\Phi = (A,B)$, **how do we efficiently compute $P(O | \Phi)$** , the probability of the observation sequence, given the model?



Problem 2 (Decoding): Given the observation sequence $O=(o_1o_2...o_T)$, and an HMM model $\Phi = (A,B)$, **how do we choose a corresponding state sequence $Q=(q_1q_2...q_T)$** that is optimal in some sense (i.e., best explains the observations)?

Problem 3 (Learning): **How do we adjust the model parameters $\Phi = (A,B)$ to maximize $P(O | \Phi)$?**

third problem



0



加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

注册



Decoding



0



- One possibility:
 - For each hidden state sequence Q
 - HHH, HHC, HCH,
 - Compute $P(O|Q)$
 - Pick the highest one
- Why not?
 - N^T
- Instead:
 - The Viterbi algorithm
 - Is again a **dynamic programming** algorithm
 - Uses a similar trellis to the Forward algorithm

Viterbi intuition

加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

注册



observation sequence together with the best state sequence

$$v_t(j) = \max_{q_0, q_1, \dots, q_{t-1}} P(q_0, q_1 \dots q_{t-1}, o_1, o_2 \dots o_t, q_t = j | \lambda)$$



0



$$v_t(j) = \max_{i=1}^N v_{t-1}(i) a_{ij} b_j(o_t)$$

Viterbi Recursion

1. Initialization:



0

$$v_1(j) = a_{0j}b_j(o_1) \quad 1 \leq j \leq N$$

$$bt_1(j) = 0$$



2. Recursion (recall that states 0 and q_F are non-emitting):



$$v_t(j) = \max_{i=1}^N v_{t-1}(i) a_{ij} b_j(o_t); \quad 1 \leq j \leq N, 1 < t \leq T$$



$$bt_t(j) = \operatorname{argmax}_{i=1}^N v_{t-1}(i) a_{ij} b_j(o_t); \quad 1 \leq j \leq N, 1 < t \leq T$$

3. Termination:

$$\text{The best score: } P^* = v_T(q_F) = \max_{i=1}^N v_T(i) * a_{i,F}$$

$$\text{The start of backtrace: } q_T^* = bt_T(q_F) = \operatorname{argmax}_{i=1}^N v_T(i) * a_{i,F}$$

以上所有内容从以下附件中抽取

- [lec3.pdf](#)
- [斯坦福大学关于语音识别的全部课件](#)

加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

[登录](#)
[注册](#)


版权声明：本文为博主原创文章，未经博主允许不得转载。 <http://blog.csdn.net/u010598445/article/details/51681413>



目前您尚未登录，请 [登录](#) 或 [注册](#) 后进行评论



0

 **从声学模型到HMM算法总结 2016 年语音识别的重大进步**  charleslei 2017年02月27日 16:57  5393

从声学模型到HMM算法总结 2016 年语音识别的重大进步

 **语音识别-解码过程**  quheDiegooo 2017年02月23日 17:49  3430

语音识别-解码过程

从PHP菜鸟到高手，我是如何脱颖而出的！

从入门到精通，你必须熟练的知识点。



HMM算法在语音识别中的应用——算法学习  u010598445 2016年06月15日 14:04  1942

总体框架输入Treat acoustic(听觉的) input O as sequence of individual observations $O=o_1,o_2,\dots,o_t,O=o_1,o_2,\dots$

DTW算法（语音识别）  fzxy002763 2013年01月25日 01:54  29005

加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

[登录](#)

[注册](#)



DTW算法C源码（语音识别算法）

2009年12月31日 10:43 8KB

下载



一个数学公式教你秒懂天下英语

老司机教一个数学公式秒懂天下英语



百度语音识别技术负责人李先刚：如何利用Deep CNN大幅提升识别准确率？

百度语音识别技术负责人李先刚：如何利用Deep CNN大幅提升识别准确率？ 机器之心mp 2016-11-04 14:24:34 技术 百度阅读(440) 0) ...

starzhou 2016年11月24日 13:55 3604

语音识别的算法实现C++

2012年03月01日 01:01 8.12MB

下载



C++语音识别接口快速入门（Microsoft Speech SDK）

C++语音识别接口快速入门（Microsoft Speech SDK）最近毕 MichaelLiang12 2016年05月04日 19:38 11967
业设计用到了微软的C++语音识别接口，查找了很多资料，也碰到了很多问题，走了很多弯路。现在把我自己的经验写下来，一是提升自己...

【VS开发】【智能语音处理】特定人语音识别算法—DTW算法

加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

注册





LG1259156776 2016年07月14日 16:48 1566

语音识别之HMM算法及其源码



c602273091 2015年04月16日 11:14 4916

HMM算法用于语言识别，在这里给出源码，供参考

13.4 非确定人语音识别算法——HMM



joey_su 2013年12月11日 22:09 6530

与DTW相比，HMM一方面用隐含的状态对应于声学层各相对稳定的发音单位，并通过状态转移和状态驻留来描述发音的变化；另一方面，它引入了概率统计模型，不再用动态时间对齐的方法求匹配距离，而是用概率密度函数...

在职研究生有用吗

在职研究生...

百度广告



语音识别的基本方法



xiaoding133 2012年03月16日 15:18 5946

一般来说,语音识别的方法有三种：基于声道模型和语音知识的方法、模板匹配的方法以及利用人工神经网络的方法[1]。

(1) 基于语音学和声学的方法 该方法起步较早，在语音识别技术提出的开始，就有了这...

语音识别算法

2013年05月10日 19:01 442KB

[下载](#)

PDF

语音识别算法matlab 源代码！完整！可运行！

2013年11月14日 16:13 2.05MB

[下载](#)

ZIP

语音识别项目：<http://www.oschina.net/project/tag/203/tts-speech> [sf.net http://www.codes...](http://www.codes...)

走进语音识别中的WFST（一）



l_b_yuan 2016年03月13日 10:58 9758

本人最近在研究语音识别的生成Graph和Lattice的模块，其中用到了WFST这个概念，惊叹于它的神奇也被它的复杂搞得晕头转向。于是决定静下心来仔细阅读了Mohri大牛的Speech Recogni...



教你一招搞定背单词难题！

巧记单词 3 isy!



FIR滤波器设计（三）



silenceee 2006年10月31日 16:37 4972

复卷积给 $n(w)$ 带来过冲积波，所以加窗函数后，对滤波器的理想特性的影响有以下几点:1.Hd(w)在截止频率的间断点变成了连续的曲线，使得H(w)出现了一个过渡带，它的宽度等于窗函数的主瓣的宽度。由...

对各位语音识别新手的几句建议



u010384318 2016年03月14日 22:55 6399

对各位语音识别新手的建议由于工作的原因，很长时间不能更新博客和管理kaldi群，每天看着kaldi群的人数不断增长，由衷的为从事语音感到自豪，希望在我博客和群里能得到你们想要的，但我同时拒绝伸手党。这...

一个语音识别的例子



qp120291570 2012年09月09日 16:27 3136

一、提要 昨天搞定了SDK中的那个例子，接着想按照文档自己来写一个，但是.....直接用文档里的代码能调出来东西的话，简直就是奇迹。所以就花了一些时间，完成了这个语音...

Python3.4 语音控制电脑（基于win10语音识别）

import speech import win32api import os import sys import time import (119) u012511642 2015年08月16日 10:11 4050

加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

注册





0



加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

注册

