

高斯混合模型（GMM）的实现

-这是暑假第二期工作内容

1. 什么是高斯混合模型？

百度中的解释：

高斯混合模型就是用高斯概率密度函数精确的量化事物，将事物分解为若干基于此高斯概率密度函数的曲线的模型。

维基百科中的解释：

维基百科中没有关于高斯混合模型的定义，但是存在对于高斯混合模型的典型例子。

2. 我们为什么要使用高斯混合模型？

根据个性化文本语音转换系统的流程，我们需要先对源语音进行分析，提取源语音的特征参数，从而获得映射规则，再根据映射规则将源语音的特征参数转换为目标语音的特征参数，最后合成出转换语音。而我们的高斯混合模型就是获得映射规则的工具

3. 高斯混合模型是如何工作的？

设有一组特征参数，参数的个数为 n ，假设每个点由一个单高斯分布生成，而这一组参数共有 m 个单高斯模型生成，具体哪个参数属于哪个单高斯模型未知，且每个单高斯模型在混合模型中占的比例未知，将所有不同分布的数据点混在一起所构成的分布称为高斯混合分布。

这些数据的概率分布密度函数可以通过加权函数表示： $p(x_i) = \sum_{j=1}^m a_j N_j(x_i; \mu_j; \Sigma_j)$ ，这就是 GMM。

其中， $\sum_{j=1}^m a_j = 1$ ， $N_j(x_i; \mu_j; \Sigma_j) = \frac{1}{\sqrt{(2\pi)^m |\Sigma_j|}} \exp\left[-\frac{1}{2}(x_i - \mu_j)^T \Sigma_j^{-1}(x_i - \mu_j)\right]$

目标语音参数透过由本语音参数生成的高斯概率模型生成另外属于目标的概率模型。然后建立目标语音。

在这之前还要使用 EM 估计 GMM 的参数，论文中采用的是 k-means 算法。

K-means 算法见论文第 37 页。参考资料 <http://www.csdn.net/article/2012-07-03/2807073-k-means>

有关 k-means 的 C++ 代码：

<http://blog.csdn.net/qll125596718/article/details/8243404>

K-means 聚类算法流程：

1. 在具有无数点的图中随意放入 K 个种子点
2. 计算每个点至每个种子点的距离
3. 规定每个点到其中一个种子点的距离最小，则规定其属于某个种子点
4. 将每个种子点移动到属于其的点群中间
5. 回到第二部分，直到种子点不需要再移动为止，称之为收敛
6. E-Step：估计未知参数的期望值，给出当前的参数估计

7. S-Step: 重新估计分布参数, 以使得数据的似然性最大, 给出未知变量的期望估计