

基于 GMM 的 AMR-NB 与 G.729A 之间的 LSP 参数转码方法

刘张宇, 鲍长春, 邱建伟, 徐昊

(北京工业大学 电子信息与控制工程学院 语音与音频信号处理实验室, 北京 100124)

摘要: 通过对高斯混合模型的数学分析, 建立并实现了一种基于高斯混合模型的、AMR-NB 与 G.729A 之间的 LSP 参数转码算法, 并通过实验对高斯混合模型在转码中的一些实际应用问题进行了分析和讨论。实验结果表明, 提出的基于高斯混合模型的 LSP 参数转码算法在合成语音质量不次于 DTE 方法的同时, LSP 参数转码复杂度下降了 78% 以上。

关键词: 语音转码; GMM; AMR-NB; G.729A; LSP

中图分类号: TN912.3

文献标识码: A

文章编号: 1000-436X(2010)02-0044-07

LSP transcoding method between 3GPP AMR-NB and ITU-T G.729A based on GMM

LIU Zhang-yu, BAO Chang-chun, QIU Jian-wei, XU Hao

(Speech and Audio Signal Processing Lab, School of Electronic Information and Control Engineering,

Beijing University of Technology, Beijing 100124, China)

Abstract: An efficient transcoding algorithm of LSP (line spectrum pairs) parameters based on GMM (Gaussian mixture model) was proposed and implemented through the mathematical analysis of GMM. Meanwhile, several practical issues in details were analyzed under experiments. The experiments indicate that the computational complexity of the proposed LSP transcoding algorithm is reduced by above 78% and the transcoded speech quality is not worse than DTE method.

Key words: speech transcoding; GMM; AMR-NB; G.729A; LSP

1 引言

AMR-NB^[1] (adaptive multiple rate-narrow band) 是 3GPP (3rd generation partnership project) 制订的语音压缩标准, 现广泛应用于 WCDMA (wideband code division multiple access) 和 TDS-CDMA (time division-synchronous code division multiple access) 等第 3 代移动通信系统中。G.729^[2] 是 ITU (international telecommunication union) 于 1995 年制定的语音编码标准, G.729A^[3] 是 G.729 的低复杂度版本, 现主要应用于语音压缩与 VoIP (voice over internet protocol) 等通信系统。这 2 种基于 CELP^[4] (code-excited liner prediction) 的语音编码技

术在当今移动通信系统和网络通信系统中的作用越来越重要。为了实现不同供应商之间通信设备的兼容与互通, 需要在这 2 种不同语音编码标准之间进行转码工作。

在基于 CELP 的 AMR-NB 和 G.729A 语音转码中, LSP 参数转码是转码算法流程的第一步, 也是整个语音转码算法的核心部分之一, 基音参数, 代数码数以及增益的转码均建立在 LSP 参数转码的基础之上, 因此, LSP 参数转码效果将直接影响最终合成的语音质量。目前在 LSP 参数转码中应用最为广泛的方法是直接转码 (DTE, decode then encode) 模式, 这种传统的码流转换方案具有 2 个缺点: 1) 运算量大; 2) 由于二次压缩造成语音失真

度加大,降低了合成语音质量^[5]。为了解决 DTE 方法带来的弊端,国外学者提出了利用直接参数转换的方法^[6,7]来实现这 2 种编码标准之间的 LSP 参数转码,然而,LSP 参数转码虽然能够有效降低计算复杂度,但仍然存在二次量化失真,因此,需要找到一种更加有效的 LSP 参数转码算法以提高转码语音质量。

本文对高斯混合模型(GMM, Goussian mixture model)进行了分析研究,并将其应用到了 AMR-NB 与 G.729A 之间的 LSP 参数转码算法中。该方法利用大量训练语音数据,通过 EM 迭代算法进行高斯混合模型参数的估计,最后得到 LSP 参数转码函数。通过大量实验,本文分析了训练数据量、GMM 数量、不同初始化方法的选取、收敛门限的限定和协方差矩阵限定与转码算法性能的关系,并得出了相应的结论。本算法在保证语音质量的情况下,极大地降低了计算复杂度和存储空间。

2 GMM 概述

GMM 是一种多维概率密度函数,常用来表示未知概率分布数据的分布函数,它在本质上是单状态的 HMM 模型,其核心思想是用多个高斯分布的概率密度函数的组合来描述特征矢量在概率空间的分布状况^[8]。根据统计理论,若干个高斯概率密度的线性组合可以逼近任意分布,因此 GMM 能够很好地描述各种形式的语音特征统计分布及其特性。下面介绍 GMM 的基本原理。

2.1 GMM 的参数描述

GMM 是由 M 个服从高斯分布的概率密度函数的加权组合而成的,其中每个高斯概率密度函数可以看作一个类,如图 1 所示。

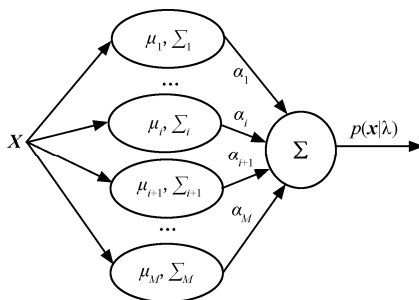


图 1 GMM 组成示意图

其表达式为

$$P_{\text{GMM}}(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{i=1}^M \alpha_i N(\mathbf{x}; \boldsymbol{\mu}_i; \boldsymbol{\Sigma}_i), \alpha_i \geq 0 \quad (1)$$

式(1)中 \mathbf{x} 是 D 维的特征向量, α_i 是各高斯函数的混合权重,必须满足 $\sum_{i=1}^M \alpha_i = 1$ 的限制, $\boldsymbol{\mu}$ 是高斯分布的均值向量, $\boldsymbol{\Sigma}$ 是高斯分布的协方差矩阵, M 是混合高斯模型中高斯函数的数量。 $N(\mathbf{x}; \boldsymbol{\mu}_i; \boldsymbol{\Sigma}_i)$ 为 M 个 D 维的高斯概率密度函数,计算公式如下所示:

$$N(\mathbf{x}; \boldsymbol{\mu}_i; \boldsymbol{\Sigma}_i) = \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}_i|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right], \quad i=1, \dots, M \quad (2)$$

一个 GMM 可以由均值矢量、协方差矩阵和混合权重等参数进行描述,通常用 $\boldsymbol{\lambda}$ 来表示这些参数的集合,如式(3)所示:

$$\boldsymbol{\lambda} = \{\alpha_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\} \quad i=1, \dots, M \quad (3)$$

应用式(3),可将式(1)改写为

$$p(\mathbf{x} | \boldsymbol{\lambda}) = \sum_{i=1}^M \alpha_i p_i(\mathbf{x}_i), \quad i=1, \dots, M \quad (4)$$

其中

$$p_i(\mathbf{x}_i) = N(\mathbf{x}; \boldsymbol{\mu}_i; \boldsymbol{\Sigma}_i) \quad (5)$$

2.2 GMM 的参数估计(EM 算法)

在语音的 LSP 参数转码中应用 GMM 需要解决一个问题,即通过输入码流 \mathbf{x} 来求得模型参数 $\boldsymbol{\lambda}$,使得 $p(\mathbf{x} | \boldsymbol{\lambda})$ 达到最大值,从而求得 LSP 转码函数。这种优化准则即为最大似然估计准则(MLE),而如何调整模型参数 $\boldsymbol{\lambda}$,使 $p(\mathbf{x} | \boldsymbol{\lambda})$ 达到最大值,也就是 GMM 的训练问题。

本文采用 EM 迭代算法进行 GMM 参数训练。该算法主要分为下面 2 个步骤。

1) E 步,即预估参数。根据所有训练数据来估计高斯混合模型的混合权重、均值向量和协方差矩阵等参数。

2) M 步,即最大化。从上一步得到的估计结果中,根据最大似然准则重新估算模型参数值,直到参数值达到最佳为止。

其中 EM 迭代算法中使用的公式如下所示^[9]:

$$\alpha_i = \frac{1}{T} \sum_{i=1}^M h(i | \mathbf{x}_i) \quad (6)$$

$$\boldsymbol{\mu}_i = \frac{\sum_{i=1}^M h(i | \mathbf{x}_i) \mathbf{x}_i}{\sum_{i=1}^M h(i | \mathbf{x}_i)} \quad (7)$$

$$\sigma_i = \frac{\sum_{i=1}^M h(i|x_i)x_i^2}{\sum_{i=1}^M h(i|x_i)} - \mu_i^2 \quad (8)$$

其中, $h(i|x_i)$ 是特征矢量 x_i 在第 i 个高斯密度函数的后验概率,

$$h(i|x_i) = \frac{\alpha_i p_i(x_i)}{\sum_{i=1}^M \alpha_i p_i(x_i)} \quad (9)$$

下面给出 GMM 的训练流程, 如图 2 所示。

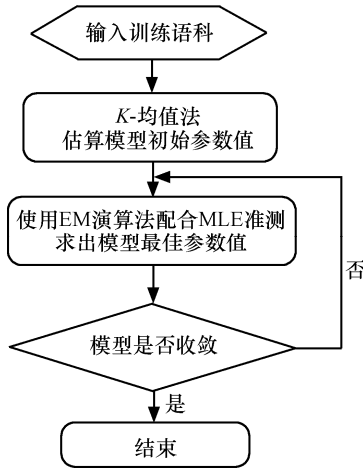


图 2 GMM 训练流程图

在进行 EM 算法之前, 首先需要对参数集 $\lambda = \{\alpha_i, \mu_i, \Sigma_i\}$ 进行初始化, 本文采用的是 K 均值方法进行参数值初始化, 即对训练数据中的所有特征矢量求均值和方差, 作为初始均值和方差, 初始权重设为相等权重, 即 $\alpha = 1/M$ 。利用 EM 迭代算法求出新的 $h(i|x_i)$, 并与前一次得到的 $h(i|x_i)$ 进行比较, 如果比较得到的差值小于一个设定的门限 δ , 则迭代结束, 即可求得相对应的 GMM 参数集 $\lambda = \{\alpha_i, \mu_i, \Sigma_i\}$ 。

3 基于 GMM 的 LSP 参数转码

3.1 LSP 参数转码函数的建立和求取

如何建立和求取基于 GMM 的 LSP 参数转码函数是 LSP 转码中的核心问题。图 3 给出了转码函数 $F(x)$ 的建模框架。

首先将语音训练数据分别通过 AMR-NB 与 G.729A 的编码端, 得到 2 组 LSP 参数向量, 定义为 X 和 Y , 以最小均方误差准则, 对 X 和 Y 根据最大似然准则进行联合高斯混合模型训练, 即得到转

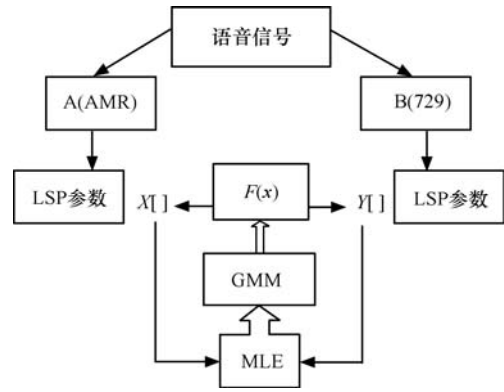


图 3 基于 GMM 的 LSP 转码函数训练过程

码函数 $F(x)$ 。将 $F(x)$ 引入到转码算法中, 当源端的 LSP 参数码流通过 $F(x)$ 后, 即得到目标端的 LSP 参数, 从而完成 LSP 转码。接下来利用联合高斯分布的条件期望预测方法^[10]对转码函数 $F(x)$ 进行数学建模, 如图 4 所示。

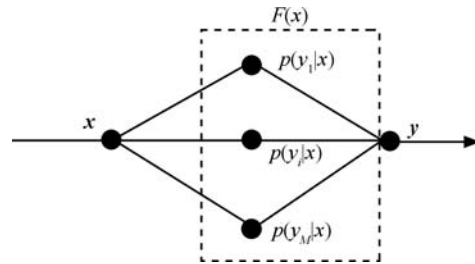


图 4 基于加权后验概率的转码函数建模

源端的一组 LSP 参数 X 经过转码函数得到目标端的一组 LSP 参数 Y , 由于 LSP 参数具有独立性, 因此 X 与 Y 是按序一一对应的, 根据后验概率的思想, 转码函数是 M 个加权后验概率的组合, 其数学表达式如下:

$$\begin{aligned} F(x) &= E(y|x) \\ &= \int yp(y|x)dy \\ &= \sum_i^M p_i(x)[\mu_i^y + \Sigma_i^{yx}(\Sigma_i^{xx})^{-1}(x - \mu_i^x)] \end{aligned} \quad (10)$$

根据贝叶斯公式, 得到

$$p_i(x) = \frac{\alpha_i N(x; \mu_i^x, \Sigma_i^{xx})}{\sum_{j=1}^M \alpha_j N(x; \mu_j^x, \Sigma_j^{xx})} \quad (11)$$

为了求解式 (10) 中的未知参数, 需要对 2 端 LSP 参数进行联合高斯混合模型训练^[11]。首先把按时间对齐的 AMR-NB 端 LSP 参数和 G.729A 端 LSP 参数合在一起, 如式 (12) 所示:

$$\mathbf{Z}_{2N \times D} = \begin{bmatrix} \mathbf{X}_{N \times D} \\ \mathbf{Y}_{N \times D} \end{bmatrix} \quad (12)$$

式(12)中, N 是训练数据的数量, D 是 LSP 参数的维数。然后利用 EM 迭代算法对矢量集 $\mathbf{Z}_{2N \times D}$ 进行 GMM 训练, 得到 GMM 参数集 $\lambda \{ \alpha_i, \mu_i, \Sigma_i \}$, 其中协方差矩阵和均值分别表示为

$$\Sigma_i = \begin{bmatrix} \Sigma_i^{XX} & \Sigma_i^{YX} \\ \Sigma_i^{XY} & \Sigma_i^{YY} \end{bmatrix}, \quad i = 1, \dots, M \quad (13)$$

$$\mu_i = \begin{bmatrix} \mu_i^X \\ \mu_i^Y \end{bmatrix}, \quad i = 1, \dots, M \quad (14)$$

将式(2)、式(11)、式(13)、式(14)代入式(1)即可求得 $F(\mathbf{x})$, 即求得目标端 LSP 参数, 完成 LSP 参数转码。

3.2 GMM 在 LSP 参数转码应用中的问题

GMM 参数在 LSP 参数转码中的应用中最关键的部分是参数的迭代估计, 而在 EM 迭代算法中需要注意训练数据的选择和模型参数的选择等问题。下面通过一系列实验对这些问题进行详细的分析。本文实验均以 AMR-NB 10.2kbit/s 模式与 G.729A 转码为例。

3.2.1 训练数据量对转码结果的影响

1) 不同训练数据量对合成语音质量的影响。

由于 GMM 是一种概率统计模型, 因此训练数据量的大小对建模效果有较大的影响, 从而间接影响转码效果。首先取 GMM 数为 32, 分别采用从 8s 到约 26min 不同时长的 NTT 数据库标准语音作为测试数据进行 GMM 训练, 并对 6 句中文语音(男女声各 3 句)进行转码实验, 得到了平均的客观 MOS 分值^[12, 13]。

如图 5 所示, 当高斯混合函数个数一定时, 转码语音质量随着训练数据集的增大而提高, 但在训练数据量超过 2 万帧之后, 继续加大训练数据对语音质量的提高不再有明显作用。另外, 在基于 GMM 的 LSP 参数转码算法中, 在测试语料相同的条件下, AMR-NB 向 G.729A 转码的 MOS 分要略高于 G.729A 向 AMR-NB, 这种情况与 DTE 以及传统参数转码算法相似。

2) 不同训练数据量对计算复杂度的影响。

LSP 参数是通过转码函数 $F(x)$ 进行转码的, 转码函数 $F(x)$ 的构建是通过 GMM 参数进行训练完成的, 而 GMM 的 EM 参数估计并不依赖于源

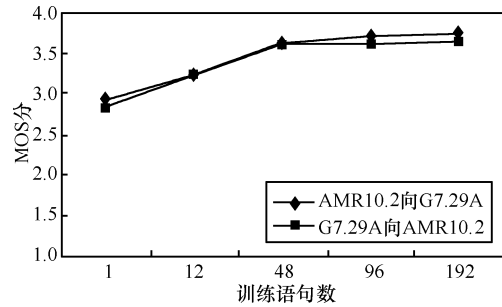


图 5 训练数据量的大小对转码客观 MOS 分的影响

LSP 码流, 是预先完成、独立于转码算法之外的。因此, 整个 GMM 训练过程对转码算法的计算复杂度没有影响, 训练数据集的增加也与转码计算复杂度无关。

3.2.2 GMM 数选取对转码结果的影响

1) 不同 GMM 个数对平均谱失真(SD, spectral distortion)的影响。

谱失真的定义如下:

$$SD_n = \frac{1}{\pi} \left\{ \int_0^\pi \left[20 \lg \frac{|H(\omega)|}{|\tilde{H}(\omega)|} \right] d\omega \right\}^{1/2} \quad (15)$$

其中 $H(\omega)$ 和 $\tilde{H}(\omega)$ 分别为转码之前和转码之后的线性预测分析谱包络。以 AMR-NB10.2kbit/s 模式向 G.729A 转码为例, 测试语料为 3 男声 3 女声共 6 句中文语音, 平均谱失真与 GMM 模型个数的关系如图 6 所示。

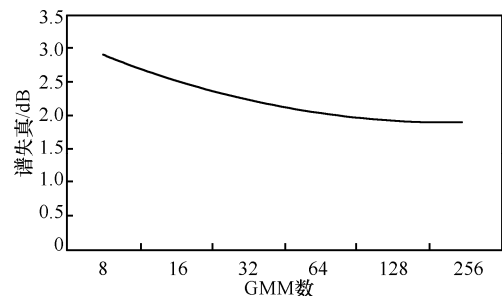


图 6 LSP 参数转码随 GMM 数增加的谱失真变化曲线

从图 6 中可以看到随着 GMM 个数的增加, LSP 参数转码谱失真逐渐下降, 最后趋向于一个极值。另外, 基于 DTE 方法的 AMR-NB10.2kbit/s 模式向 G.729A 转码的谱失真是 2.683dB, 因此, 当 GMM 个数大于等于 16 时, 本 LSP 参数转码方法产生的谱失真小于 DTE 转码方法。

2) 不同 GMM 个数对合成语音质量的影响。

GMM 是由具有 M 个混合成分的高斯密度函数来进行线性叠加的, 因此高斯模型的阶数, 即高斯密度函数的个数的大小与基于 GMM 的 LSP

参数转码效果直接相关。基于 GMM 的 LSP 参数转码算法是利用 GMM 对线谱频率参数进行拟合，因此，在理论上 M 越大，声道谱参数包络就拟合得越精确，转码性能也就越好。本实验以 AMR10.2kbit/s 转码速率为例，首先利用 76 800 帧的 NTT 数据库标准语音作为训练数据，分别采用 5 种从小到大的不同的混合数进行 GMM 训练，并对 6 句中文语音（男女声各 3 句）进行转码实验，得到了平均的客观 MOS 分值，GMM 个数对 LSP 参数转码后合成语音质量的影响实验结果图 7 所示。

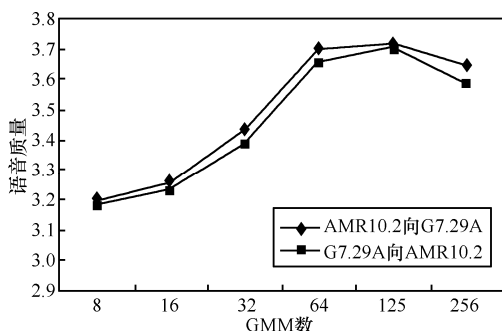


图 7 GMM 数大小对转码语音质量的影响

从图 7 可以看出，在训练数据量一定的条件下，一般地，转码语音质量随着 GMM 数的增加而提高，但在混合模型达到 256 时，转码性能低于 128 个混合模型数，也就是说，在 LSP 参数转码的实际应用中，GMM 训练出现了过训练现象。因此单从 GMM 数对 LSP 参数转码质量的影响来看，存在一个 GMM 的个数能够对应最优的转码语音质量。经过实验比较，取 GMM 个数为 128。

3.2.3 EM 算法中迭代次数的分析与收敛门限的确定

前文中已经讨论了 EM 迭代算法的流程，应用最大似然法来获得所要最大化的目标 GMM 参数集 λ 需要预先设定一个门限值 δ ，而这个门限值与 EM 迭代算法的迭代速度和收敛精度密切相关，因此，选择一个合适的 δ 是比较重要的。由于 GMM 个数同样影响迭代速度，因此本实验对门限值 and GMM 数进行联合分析。在本实验中，GMM 数分别取 8, 16, 32, 64, 训练数据为 24 句 NTT 标准语音库语音，每句时长为 8s。实验结果如图 8 所示。

从图 8 中可以看到，迭代次数是由收敛门限值和 GMM 数 2 个因素共同决定的。当 GMM 数

较少时，迭代次数主要取决于收敛门限的取值，这是因为少量的 GMM 不足以准确描述谱参数特征，需要更高的精度来保证 GMM 的准确性。当 GMM 数足够多时，在本实验中即 GMM 数达到 64 时，大量的 GMM 只需较少次数的迭代便可以满足收敛门限的要求。在实际应用中，出于对计算复杂度的考量，GMM 数不能取值过大，因此，需要通过确定较高的收敛门限来保证精度，在本文中取门限值 δ 为 10^{-6} 。

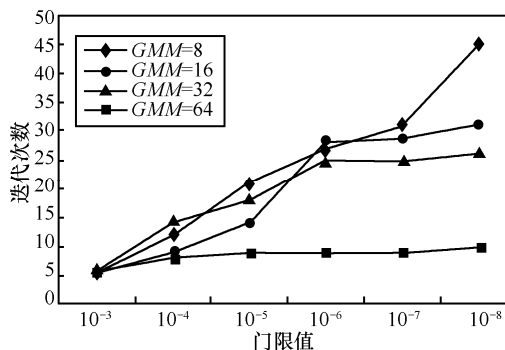


图 8 门限值和 GMM 数与迭代次数的关系

4 实验结果

为了验证本文所提算法的有效性，在主观 A/B 听力测试与客观 MOS 分测试中，LSP 参数采用基于 GMM 的 LSP 参数转码技术进行转码，而基音参数、代数码数以及增益采用 DTE 转码方式进行转码。

4.1 主观 A/B 听力测试

本实验采用 A/B 听力测试方法对转码语音进行了主观测试。测试语音由 24 句组成，每句长 8s，分别由 2 男 2 女 4 个说话人发音。6 名测听人员分别对 LSP 经 DTE 转码的语音和经 GMM 转码的语音进行主观测听，并得到以下主观偏好结果，如表 1 所示。

表 1 LSP 转码语音的主观 A/B 听力对比

转码方向	偏好百分比 (共 144 票)		
	DTE	GMM	No_Ref
AMR-NB10.2 to G729A	28.2%	32.9%	38.9%
G729A to AMR-NB10.2	26.4%	27.8%	45.8%

从表 1 中可以看到，在 AMR-NB 10.2kbit/s 与 G729A 之间的转码中，基于 GMM 的 LSP 转码主观听力质量不次于 DTE 的 LSP 转码方法。

4.2 客观 MOS 分测试

本实验使用 ITU-T P.862.1 所规定的 MOS_LQO^[14]为客观语音质量的衡量标准。实验选取 NTT 标准语音数据库的 96 句中文语音作为测试数据源, 一共 4 男 4 女 8 位说话人, 每人讲 12 句话, 每句 8s 时长, 对于 AMR-NB 来说是每句话 400 帧, 对于 G.729A 来说是每句话 800 帧。以 AMR-NB 10.2kbit/s 模式向 G.729A 转码为例, 与 DTE 转码模式进行比较, 实验测得的 MOS 分结果如表 2 所示。

表 2 LSP 转码语音的 MOS 分比较

测试语音	MOS_LQO 分	
	DTE	GMM
48 句男声	3.593	3.594
48 句女声	3.293	3.285
96 句语音	3.445	3.440

从表 2 中的 MOS 分值中可以看到, 在男声测试语音中, LSP 2 种转码算法的质量非常接近; 在女声测试语音中, 基于 GMM 的 LSP 转码语音质量略低于 DTE 方法; 在所有语句中, GMM 方法的平均 MOS 分与 DTE 方法比较接近, 表明转码语音质量在可接受的范围之内。

4.3 复杂度结果

在本实验中预设高斯混合模型值为 128, 基于 GMM 的 LSP 转码算法与基于 DTE 的 LSP 转码算法的计算复杂度与空间复杂度对比如表 3 所示。

表 3 LSP 转码方法的复杂度比较

转码方法	复杂度		
	运算量	WMOPS	存储量/B
DTE	594 500	1.235 5	898 650×32
GMM	129 110	0.260 69	96 816×32
下降百分比	78.28%	78.90%	89.23%

从表 3 中可以看到, 基于 GMM 的 LSP 转码方法与 DTE 方法相比, 极大地降低了计算复杂度和空间复杂度, 取得了比较好的结果。

5 结束语

本文主要概述了高斯混合模型的定义和基本

原理, 详细介绍了高斯混合模型的参数估计算法——EM 迭代算法, 提出并实现了基于 GMM 的 AMR-NB 与 G.729A 之间的 LSP 参数转码算法, 以 10.2kbit/s 转码速率为例, 分析了 GMM 在 LSP 参数转码中的几个实际应用问题, 最后给出了实验结果。实验结果表明, 基于 GMM 的 LSP 转码方法能够在保证合成语音质量的前提下, 极大地降低计算复杂度和空间复杂度。GMM 在 LSP 参数转码应用中的有效性, 为 GMM 在其他参数转码中的应用提供了重要的借鉴意义。

参考文献:

- [1] ETSI EN 301 704 V7.2.1 Adaptive Multi-Rate(AMR)Speech Transcoding[S]. 2000.
- [2] ITU-T G.729: Coding of Speech at 8kbit/s Using Conjugate Structure Algebraic Code Excited Linear Prediction(CS-ACELP)[S]. 1996.
- [3] ITU-T G.729A: Educued Complexity 8kbit/s CS-ACELP Speech Codec[S]. 1996.
- [4] 鲍长春. 数字语音编码原理[M]. 西安: 西安电子科技大学出版社, 2007.
- [5] BAO C C. Principles of Digital Speech Coding[M]. Xi'an: Xidian University Press, 2007.
- [6] 邱建伟, 鲍长春, 徐昊等. 基于 CELP 编码模型参数转码技术[J]. 电声技术, 2009,(4):84-87.
- [7] QIU J W, BAO C C, XU H, *et al.* Parameter transcoding techniques based on CELP speech coding[J]. Audio Engineering. 2009(4): 84-87.
- [8] OTA Y, SUZUKI M, TSUCHINAGA Y, *et al.* Speech coding translation for IP and 3G mobile integrated network[A]. IEEE International Conference on Communications[C]. New York: IEEE Press, 2002. 114-118.
- [9] GHENANIA M, LAMBLIN C. Low-cost smart transcoding algorithm between ITU-T G.729 (8kbit/s) and 3GPPNB-AMR (12.2kbit/s)[A]. European Signal Processing Conference[C]. Vienna: EUSIPCO Press, 2004, (3): 1681-1684.
- [10] 赵永刚, 唐昆, 崔慧娟. 基于 Gaussian 混合模型的 LSF 参数量化方法[J]. 清华大学学报(自然科学版), 2006, 46(10): 1727-1730.
- [11] ZHAO Y G, TANG K, CUI H J. Quantization of LSF parameters using a Gaussian mixture model[J]. J Tsinghua University (Sci & Tech), 2006, 46(10): 1727-1730.
- [12] 吴金池. 语音辨识系统之研究[D]. 台湾国立中央大学, 2003.9-17.
- [13] WU J C. Research on Speech Recognition System[D]. Taiwan, China: National Central University, 2003.
- [14] KAIN A B. High Resolution Voice Transformation[D]. Oregon Health and Science University, 2001.36-54.

- [11] 康永国, 双志伟, 陶建华等. 高斯混合模型和码本映射相结合的语音转换算法[A]. 第八届全国人机语音通讯学术会议[C]. 2005. 293-297.
KANG Y G, SHUANG Z W, TAO J H, *et al.* Speech transform algorithm based on Gaussian mixture model and codebook mapping[A]. NCMMSC2005[C]. 2005.293-297.
- [12] 付强. 语音的参数表示和质量客观评价研究[D]. 西安电子科技大学, 2000.46-66.
FU Q. Research on Parameter Representation and Objective Quality Assessment of Speech[D]. Xi'an: Xidian University. 2000.46-66.
- [13] ITU-T P.800.1: Mean Opinion Score(MOS) Terminology[S]. 2003.
- [14] ITU-T P.862.1: Mapping Function for Transforming P.862 Raw Result Scores to MOS-LQO[S]. 2003.



鲍长春 (1965-), 男, 内蒙古赤峰人, 博士, 北京工业大学教授、博士生导师, 主要研究方向为语音与音频编码。

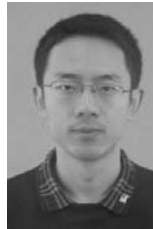


邱建伟 (1983-), 男, 内蒙古巴彦淖尔人, 北京工业大学硕士生, 主要研究方向为语音编码与语音转码。

作者简介:



刘张宇 (1982-), 男, 黑龙江佳木斯人, 北京工业大学硕士生, 主要研究方向为语音编码与语音转码。



徐昊 (1983-), 男, 北京人, 北京工业大学硕士生, 主要研究方向为嵌入式语音编码、语音转码。

(上接第 43 页)

- [14] 邓永录, 梁之舜. 随机点过程及其应用[M]. 北京: 科学出版社, 1992.
DENG Y L, LIANG Z S. Random Point Process and Its Applications[M]. Beijing: Science Press, 1992.
- [15] WEISSTEIN E W. Circle line picking[EB/OL]. <http://mathworld.wolfram.com/CircleLinePicking.html>, 2009.
- [16] JAHNKE E, EMDE F. Tables of Functions with Formulae and Curves[M]. New York: Dover Publications, 1945.
- [17] BOUDEC J L, VOJNOVIC M. The random trip model: stability, stationary regime, and perfect simulation[J]. IEEE/ACM Trans on Networking, 2006, 14(6): 1153-1166
- [18] HUGHES B. Random Walks and Random Environments[M]. Oxford: Clarendon Press, 1995.



卢继华 (1978-), 女, 黑龙江鸡西人, 博士, 北京理工大学讲师, 主要研究方向为扩频通信等。

作者简介:



刘宴涛 (1975-), 男, 吉林蛟河人, 博士, 渤海大学讲师, 主要研究方向为 ad hoc 网络等。



刘珩 (1981-), 女, 湖北恩施人, 博士, 北京理工大学讲师, 主要研究方向为 ad hoc 网络、分布式计算等。