

基于高斯混合模型的自然环境声音的识别

余清清, 李 应, 李 勇

YU Qingqing, LI Ying, LI Yong

福州大学 数学与计算机科学学院, 福州 350108

College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China

YU Qingqing, LI Ying, LI Yong. Natural sounds recognition using GMM distribution. Computer Engineering and Applications, 2011, 47(25): 152-155.

Abstract: A recognition method for natural sounds based on Gaussian Mixture Model (GMM) distribution is proposed. Mel-Frequency Cepstral Coefficients (MFCCs) are used to analyze natural sounds for their feature extraction. The expectation maximization algorithm is used to learn a Gaussian mixture model distribution of MFCCs for the set of audio feature vectors that describe each sound. Minimum classification error criterion and vote rule are used to yield higher recognition accuracy for natural sounds. Experimentally, compared with K -Nearest Neighbor (KNN) method, GMM is able to achieve a higher accuracy rate for discriminating 36 classes of natural sounds. The classified accuracy rate of GMM reaches to 95.83%.

Key words: Mel-frequency cepstral coefficients; Gaussian mixture model; natural sounds recognition; vote rule

摘 要: 提出了一种基于高斯混合模型(GMM)的自然环境声音的识别方法。提取Mel频率倒谱系数(MFCCs)来分析声音信号; 对于每种声音使用期望最大化算法基于MFCC特征集建立高斯混合模型; 使用最小错误率判决规则和投票裁决的方法进行识别。使用GMM对36种自然环境的声音进行识别的正确率可达95.83%, 且识别效果优于 K 最近邻(KNN)。

关键词: Mel频率倒谱系数; 高斯混合模型; 自然环境声音的识别; 投票裁决

DOI: 10.3778/j.issn.1002-8331.2011.25.040 文章编号: 1002-8331(2011)25-0152-04 文献标识码: A 中图分类号: TP391.4

1 引言

声音的识别是声音信号处理的一个基本问题。它在众多领域有着重要的应用, 它可以用于导航、辅助机器人技术、移动设备服务、监督和安全等领域。声音数据里含有丰富的信息。通过对声音的识别和分析可以获取丰富的环境特征, 如气候、地域、时间、物种等等, 它可以获取到视觉所捕捉不到的信息。再者, 声音数据可以随时随地获取, 它不受限于光线, 同时也无需在视野之内, 它所需的存储空间也比视频信号小, 有诸多优势。

虽然目前音频分类和识别的研究大部分集中在语音识别和音乐、语音的分类上, 但是也有关于环境声音识别和分类的研究。在老年人卫生保健上, 使用二进制等级分类器对环境声音进行识别^[1]; 使用匹配追踪算法提高受噪声影响的环境声音的识别正确率^[2]; 高噪声情境下使用小波-JFK对环境声音进行识别^[3]; 利用不受噪声影响的非线性赫宾学习算法进行车辆声音的识别^[4]; 使用SVM算法实现公共交通工具上声音事件的侦查^[5]; 利用HMM和MFCC进行浴室声音事件的分类, 实现浴室活动自动监控系统^[6]; 使用独立成分分析的MFCC和基于帧的SVM算法来实现家庭环境声音的自动辨识^[7]; 这些研究

从不同侧面揭示了环境音频数据与动物、人类生活的密切关系, 同时也预示着这些相关技术在现实中的潜在应用。

通过在森林、原野等地使用自动监听和识别物种声音的技术对濒临灭绝的动物进行侦察, 有利及时发现它们并采取相应的保护措施。地震给人类带来了巨大的灾难, 在地震带分布地区使用自动监听和识别动物声音异常的设备可作为地震预警系统的辅助技术。可见, 自然环境声音的识别和分类与人类息息相关。然而, 本文研究的正是自然环境声音的分类问题。

在声音分类问题中, 选用合适的特征表征信号是决定分类效果的关键。有很多特征可以用于表征声音信号。常用的有线性预测模型(LPC)、线性预测倒谱系数(LPCC)、Mel频率倒谱参数(MFCC)、波动模型、短时能量、过零率(ZCR)、带宽、频率中心等等。在众多声音识别系统中最常用的是Mel频率倒谱系数(MFCC)。MFCC充分考虑了人的听觉特性, 具有良好的识别性能。因此使用MFCC特征来描述自然环境音频数据。

高斯混合模型(GMM)是比较成熟的语音识别技术, 可以很好地描述说话人在不同环境和生理条件下的声音特征, 已广泛地应用于说话人辨认系统, 它是用多个高斯分布的概率

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.61075022); 福建省教育厅A类科技项目(No.JA09021)。

作者简介: 余清清(1986—), 女, 硕士研究生, 主要研究领域为音频数据检索; 李应(1964—), 男, 博士, 副教授; 李勇(1986—), 男, 硕士研究生。

E-mail: yqing_77@126.com

收稿日期: 2010-04-06; **修回日期:** 2010-08-11

密度函数的组合来描述特征矢量在概率空间的分布状况。高斯混合模型(GMM)具有简单高效的特点,因此本文利用高斯混合模型来识别自然环境中的声音信号。为每类自然环境中的声音信号基于MFCC特征向量集建立高斯混合模型。

2 前端处理

2.1 预处理

如图1所示前端处理的第一个阶段是预处理,首先将所获取到的自然环境声音文件转换成以脉冲编码调制(PCM)格式表示的波形文件;其次,所有的声音文件取样时使用的采样频率都是11 025 Hz,多声道转为单声道;最后,由于所做的实验是考虑噪音很小的情况,所以将声音波形文件中明显的噪音杂质和静音部分删去,得到能保证真实体现该类声音特性的声音波形文件。将声音文件整理成两部分,作为训练样本的声音取用时间长度长的声音文件(时间长度大于10秒),而作为测试样本的声音文件较短(时间长度约为2秒)。

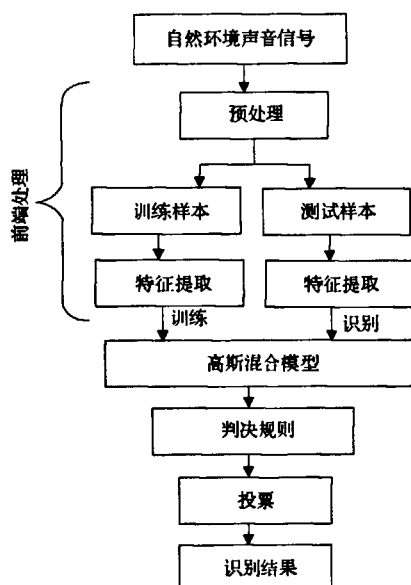


图1 基于GMM的自然环境声音识别系统图

2.2 计算MFCCs

MFCC是从人耳对频率高低的非线性心理感觉角度进行分析,采用了一种非线性的Mel频率尺度来模拟人耳的听觉系统。Mel频率尺度的值大体上与线性频率呈对数关系,更符合人耳的听觉特性。

MFCC参数的计算是以“bark”为其频率基准的,它和线性频率的转换关系是:

$$mel = 2595 \log_{10} \left(1 + \frac{f}{700} \right)$$

这里,线性频率 f 的单位是Hz。

计算MFCCs的具体步骤如下:

(1)将自然环境音频信号划分成一系列连续的帧,这里加汉宁窗分帧,每帧包含 $N=512$ 个样本,相邻帧有256个样本重叠。设音频数据时域信号为 $x(n)$,第 i 帧音频信号 $x_i(n)$ 可以表示成:

$$x_i(n) = x(i \cdot N + n)w(n) \quad (0 \leq n \leq N-1)$$

这里 $w(n)$ 是汉宁窗

$$w(n) = \begin{cases} 0.5 - 0.5 \cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & \text{其他} \end{cases}$$

对每帧信号 $x(n)$ 做快速傅里叶变换(FFT)后,取模的平方得到离散功率谱 $P(n)$ 。

(2)设计一个有36个频带的滤波器,即 $L=36$,频率范围0到5 512 Hz。这些频带在Mel轴上近乎均匀分布。使用重叠的三角形频带,第 i 个频带的起始频率 $mel(idx(i))$ 的对应频率是:

$$freq(i) = (10^{mel(idx(i))/2595} - 1) \times 700$$

其中, $idx(i)$ 是第 i 个频带的起始位置,第 i 个频带的起始位置刚好是第 $i-2$ 个频带的终止位置。第 i 个三角形的高度 $h_i = 2/(freq(i+2) - freq(i))$,它的中心频率是 $freq(i+1)$ 。

(3)通过每个三角形滤波器的输出:

$$m(l) = \sum_{k=freq(l)}^{freq(l+2)} W_l(k) P_l(k) \quad (l=1, 2, \dots, L)$$

$$W_l(k) = \begin{cases} h(l) \frac{k - freq(l)}{freq(l+1) - freq(l)}, & freq(l) \leq k \leq freq(l+1) \\ h(l) \frac{freq(l+2) - k}{freq(l+2) - freq(l+1)}, & freq(l+1) < k \leq freq(l+2) \end{cases}$$

这里, $W_l(k)$ 是第 l 个滤波器, $h(l) = 2/(freq(l+2) - freq(l))$, $l=1, 2, \dots, L$ 。

(4)对所有滤波器输出做对数运算,再进一步做离散余弦变换(DCT)即可得到一组MFCC参数:

$$mfcc(j) = \sum_{l=1}^L \log m(l) \cos\left\{\left(l - \frac{1}{2}\right) \frac{j\pi}{L}\right\}$$

其中, $j=0, 1, \dots, R$, $1 \leq R \leq L$ 。这里的 $mfcc$ 为一帧音频数据所产生的MFCC参数,一帧共产生 $R+1$ 个参数,一般只取用后面的 R 个。

对于一个音频数据文件的每帧的MFCC,计算前17个MFCC参数即 $R=16$,每帧的 R 个MFCC参数作为矩阵的一行,按顺序组成的矩阵就是MFCC矩阵,它可以表示为:

$$mfccm(i, j) = \sum_{l=1}^L \log m(l) \cos\left\{\left(l - \frac{1}{2}\right) \frac{j\pi}{L}\right\}$$

其中, $i=1, 2, \dots, I$, $j=0, 1, \dots, R$, $1 \leq R \leq L$, I 为一个音频数据文件的帧数, L 为三角频带的数量。

3 建立高斯混合模型与决策

3.1 判决规则

如图1所示,前端处理的目的是提取自然环境声音信号的特征。这个阶段的输出是一组表征声音信号段的特征矢量序列, $X = \{x_1, x_2, \dots, x_I\}$, 其中 $x_i (1 \leq i \leq I)$ 表示的是第 i 帧的特征矢量。

假设自然环境中共有 c 类声音,用 H_1, H_2, \dots, H_c 表示类型,其中第 i 类声音用 $H_i (1 \leq i \leq c)$ 表示。给定的声音信号帧特征矢量 x 是 D 维的, $P(H_i)$ 是第 i 类声音的先验概率。用 $\rho(x/H_i) (1 \leq i \leq c)$ 表示类条件概率密度,根据全概率公式,模式样本 x 出现的全概率密度为:

$$\rho(x) = \sum_{i=1}^c \rho\left(\frac{x}{H_i}\right) \cdot P(H_i) \quad (1 \leq i \leq c)$$

根据贝叶斯定理,在模式样本 x 出现的条件下,第 i 类的声音出现的后验概率为:

$$P\left(\frac{H_i}{x}\right) = \frac{\rho(x/H_i)P(H_i)}{\rho(x)} \quad (1 \leq i \leq c) \quad (1)$$

根据最小错误率判决规则,若

$$P\left(\frac{H_k}{x}\right) = \max_{i=1}^c \{P\left(\frac{H_i}{x}\right)\} \quad (2)$$

则 $x \in H_k$ 。

在式(1)中, i 无论取何值分母的值都是固定的,所以式(2)就可等价于:若

$$\rho\left(\frac{x}{H_k}\right) \cdot P(H_k) = \max_{i=1}^c \{\rho\left(\frac{x}{H_i}\right)P(H_i)\} \quad (3)$$

则 $x \in H_k$ 。

假设所有的先验概率 $P(H_i)$ ($1 \leq i \leq c$) 都相等,则式(3)可以简化为:若

$$\rho\left(\frac{x}{H_k}\right) = \max_{i=1}^c \{\rho\left(\frac{x}{H_i}\right)\} \quad (4)$$

则 $x \in H_k$ 。

3.2 高斯混合模型

在上面的判决规则的实现中很重要的一步是似然函数 $\rho(x/H)$ 的选择,在这里选择高斯混合模型。因为自然环境声音信号的 MFCC 特征矢量空间符合高斯混合分布,可以为自然环境中各种类型的声音都基于 MFCC 特征空间建立高斯混合模型。声音类型 H 用模型参数符号 θ 来表示,假设每类自然环境声音的高斯混合模型都由 M 个 D 维的高斯分量组成,可以用 M 个高斯分量加权求和的形式来表示:

$$\rho\left(\frac{x}{H}\right) = P\left(\frac{x}{\theta}\right) = \sum_{j=1}^M W_j P\left(\frac{x}{\lambda_j}\right) \quad (5)$$

其中, W_j ($j=1, 2, \dots, M$) 为混合权值,相当于第 j 个高斯分量出现的概率,满足 $\sum_{j=1}^M W_j = 1$, 用 $\lambda_j = \{\mu_j, \Sigma_j\}$ 表示模型的第 j 个高斯分量的参数, μ_j 为均值向量, Σ_j 为协方差矩阵,整个声音信号的模型参数 $\theta = \{W_j, \mu_j, \Sigma_j\}$ ($j=1, 2, \dots, M$), 其中高斯概率密度函数为:

$$P\left(\frac{x}{\lambda_j}\right) = \frac{1}{(2\pi)^{D/2} |\Sigma_j|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu_j)^T \Sigma_j^{-1}(x-\mu_j)\right\} \quad (6)$$

由前端处理的特征提取可以得到每种自然环境声音的 MFCC, 每帧都会舍去第一个 Mel 频率倒谱系数,取 16 个系数,因此建立高斯混合模型的特征矢量维数是 16 维。在前面的预处理中,已得到作为训练样本时间长度较长的声音文件,假设每个声音都可被分成 I 帧。因此每个声音可以用于训练的特征矢量共有 I 个,特征矢量序列 $X = \{x_1, x_2, \dots, x_I\}$, 其中 x_i 表示的是第 i 帧的特征矢量。为每个声音建立有 10 个高斯分量的高斯混合模型,即 $M=10$ 。为了计算的简便和更有效,使用对角协方差矩阵,即令 $\Sigma_j = \sigma_j^2 E$ ($j=1, 2, \dots, 10$) E 是单位矩阵,因此对于每个模型所要求的参数 $\theta = \{W_1, \mu_1, \sigma_1, W_2, \mu_2, \sigma_2, \dots, W_{10}, \mu_{10}, \sigma_{10}\}$ 。

以下是模型训练的具体步骤。首先初始化参数,对于给定的一组训练样本的特征向量序列 $X = \{x_1, x_2, \dots, x_I\}$, 使用 K -means 聚类算法计算群聚中心点,以作为 μ_j ($j=1, 2, \dots, 10$) 的初始参数, $W_j = 1/10$ 。然后,使用期望极大化(EM)算法迭代估计最大似然模型参数 θ , 即通过特征矢量序列 $X =$

$\{x_1, x_2, \dots, x_I\}$ 的训练,使得 $P(X/\theta)$ 最大, $P(X/\theta) = \prod_{i=1}^I P(x_i/\theta)$ 。

使用 EM 算法迭代改进 GMM 参数 θ , 使得基于观测特征向量集的估计模型似然值单调递增,即对于第 k 次和第 $k+1$ 次迭代, $P(X/\theta^{(k+1)}) > P(X/\theta^{(k)})$ 。对于每个声音模型都使用 20 次迭代,即最终的模型参数 $\theta = \theta^{(20)}$ 。

通常假设特征向量集 X 中的各个矢量间是相互独立的,所以基于一组特征向量序列 $X = \{x_1, x_2, \dots, x_I\}$ 的模型 θ 的对数似然函数的计算如下:

$$\log P\left(\frac{X}{\theta}\right) = \sum_{i=1}^I \log P\left(\frac{x_i}{\theta}\right) \quad (7)$$

其中 $P(x_i/\theta)$ 可用式(5)计算得到,通常使用平均对数似然值来标准化对数似然值以排除声音信号时间长度不同的影响,即将 $\log P(X/\theta)$ 除以 I 可得到平均对数似然值。同时因为实际中不能总是保证所有的特征向量间都是相互独立的,所以这个平均对数似然值可以作为式(7)的一个粗略的补偿因子^[9]。

3.3 识别决策

自然环境声音识别的流程如图 1 所示,经过特征向量集的训练可以得到自然环境中每类声音的模型参数 $\theta = \{W_1, \mu_1, \sigma_1, W_2, \mu_2, \sigma_2, \dots, W_{10}, \mu_{10}, \sigma_{10}\}$, 即建好了每种声音的模型。下面就是如何对给出的声音进行识别,这里使用 3.1 节所讲的判决规则,对声音信号的每一帧识别步骤如下:

(1) 将所要识别的声音先经过预处理、特征提取得到 MFCC 特征向量集,设该声音的 MFCC 特征向量集是 $X_{\text{identify}} = \{x_1, x_2, \dots, x_I\}$, 这里 I 是帧数。

(2) 将 X_{identify} 的每一帧特征向量 x_i ($1 \leq i \leq I$) 代入所有声音模型中计算在第 i 类的声音出现的条件下,模式样本 x_i 出现的后验概率值,即求式(5),

$$\rho\left(\frac{x_i}{H_i}\right) = P\left(\frac{x_i}{\theta_i}\right) = \sum_{j=1}^M W_j P\left(\frac{x_i}{\lambda_j}\right) \quad (i=1, 2, \dots, c)$$

(3) 根据最小错误率判决规则若 $\rho(x_i/H_k) = \max_{i=1}^c \{\rho(x_i/H_i)\}$, 则 $x_i \in H_k$, $\text{Class}_i = k$, $\text{Postprob}_i = \rho(x_i/H_k)$, 这里 Class_i 表示第 i 帧识别出来的声音类型是第 Class_i 类, Postprob_i 表示第 i 帧识别出来的对应的最大后验概率值。

重复上面三个步骤将待识别声音信号的特征矢量序列 $X_{\text{identify}} = \{x_1, x_2, \dots, x_I\}$ 中的每帧逐个进行识别,最后可得到 I 个 Class_i ($1 \leq i \leq I$) 值和 I 个最大概率值 Postprob_i 。尝试使用了多种方法来融合各帧概率值^[9-10], 最终得到最佳效果的是多数投票的方法。因此最后采用多数投票裁决得出所要识别的声音信号所属的类型。投票裁决的步骤如下:

(1) 初始化 $\text{Vote}_i = 0$ ($i=1, 2, \dots, c$), 这里 Vote_i 表示的是第 i 类声音的票数。

(2) for $l=1$ to I ,
if ($\text{class}_l \neq 0$) $\text{Vote}_{\text{Class}_l} = \text{Vote}_{\text{Class}_l} + 1$ ($1 \leq \text{class}_l \leq c$)

(3) 若 $\text{Vote}_a = \max_{i=1}^c \text{Vote}_i$ ($i=1, 2, \dots, c$), 则 $X_{\text{identify}} \in H_a$, 即最终识别出来该声音的类型是 H_a 。

4 实验与结果

使用上述的高斯混合模型对 36 类自然环境声音进行识

别,且也使用KNN在相同的数据集上进行实验。

4.1 实验设置

使用的自然环境声音文件来自文献[11-12]。共使用36种自然环境声音,有包括鸟类、兽类、昆虫、爬虫类等多种类型的动物的声音,还有风雷雨电和河海溪流的声音。由于不能对自然界所有声音都逐一进行实验,所以仅采集各大类自然声音中的部分进行实验。在模型训练时,每一种声音使用一个时间长度较长的声音(时间长度大于10秒)进行训练。在测试时,各种声音使用不同数量的测试例子,具体的测试例子数目如表1所示,每个测试声音文件时间长度约为2秒。共计用于测试的声音文件120个,用于训练的声音36个。自然环境声音识别的流程如图1所示,通过前端处理可以得到自然环境声音的MFCC特征向量集,然后将MFCC特征向量数据集导入高斯混合模型进行训练和测试,这个高斯混合模型的高斯分量为10。KNN的前端处理和GMM方法相同,而KNN模型中的距离尺度使用欧几里德距离,参数k的值取3。

表1 36种自然环境声音的识别结果表

编号	类型	测试个数	GMM 识别错误数	KNN 识别错误数
1	狗奔跑、喘气声	4	0	0
2	猎犬吠叫	3	0	0
3	大象	4	0	0
4	猴子	3	0	2
5	牛羚	4	0	0
6	山羊	3	0	2
7	海豹	3	0	1
8	驴	3	0	0
9	蛇	3	1	0
10	鸽子	3	0	0
11	云雀	4	0	0
12	知更鸟	3	0	0
13	麻雀	3	0	0
14	啄木鸟	3	0	0
15	燕子	3	0	0
16	白颊鸟类	4	0	0
17	小鸡	3	0	1
18	鸭子	3	0	0
19	蟋蟀	4	0	0
20	猫头鹰	3	0	0
21	猫发脾气声	3	0	0
22	猫的哭声	3	1	1
23	嚎叫的风声	4	1	0
24	低音风声	4	0	0
25	高音风声	4	0	0
26	狂风吹树	4	0	0
27	闪电雷鸣	4	1	0
28	雷雨交加	4	0	1
29	蜜蜂	3	1	1
30	蚊子	3	0	0
31	火燃烧	3	0	0
32	海洋	3	0	0
33	河流	3	0	0
34	细流声	3	0	0
35	小溪	3	0	0
36	荡水声	3	0	0
总计		120	5	9

通过前面的处理,KNN和GMM的方法都分别得到所要测试的声音文件(120个)的每帧类型,最后都是使用投票决策

确定每个声音文件的类型。

4.2 结果与分析

GMM、KNN这两种方法分别具体的测试结果如表1所示。由实验结果可得出,GMM的识别正确率可达95.83%,而KNN的识别率是92.5%。基于高斯混合模型的识别效果优于KNN,且GMM所消耗的空间比KNN小。

GMM对于较难区分的声音的识别正确率仍很高。例如表1中的第11类至第16类声音同是鸟类都是较为相似的声音,这7类GMM均能识别正确,GMM在36类声音的识别中识别正确率最低的是66.7%,而KNN的最低识别正确率是33.3%。即GMM最坏的实例的分类正确率仍是高于KNN,可见GMM模型具有较好的“鲁棒”性。

5 总结

使用了基于MFCC特征空间建立高斯混合模型的方法来识别自然环境声音。实验证明了MFCC能有效地描述自然环境声音信号,GMM分布能有效地近似自然环境声音的MFCC特征空间的分布。通过实验,发现了在建高斯混合模型时,要注意选用噪声小能真实体现该类型声音信号的训练样本,如此建立的模型更准确,否则会大大影响识别效果;同时应尽量选用时间长度较长的训练样本,如此能充分体现该类声音信号的特点,建立的模型更全面。总之,模型的训练是影响该系统识别性能最关键的一步。

后续需要进一步研究的是如何提高对于人耳感觉区分度较小的声音的识别。同时由于MFCC的抗干扰小,对于噪音较小的测试例子可得到较好的识别效果,但对于信噪比低的声音识别准确度就下降了。因此,对于信噪比低的声音的识别也是下一步研究的内容。

参考文献:

[1] Peng Y T, Lin C Y, Sun M T. Audio event classification using binary hierarchical classifiers with feature selection for health-care applications[C]//IEEE International Symposium on Circuits and Systems, Seattle, Washington, USA, 2008: 3238-3241.

[2] Chu S, Narayanan S, Kuo C C J, et al. Environmental sound recognition using MP-based features[C]//IEEE International Conference on Acoustics, Speech, and Signal Processing, Las Vegas, Nevada, USA, 2008: 1-4.

[3] Tran H D, Li H Z. Jump function kolmogorov and its application for audio stream segmentation and classification[C]//IEEE International Conference on Acoustics, Speech, and Signal Processing, Las Vegas, Nevada, USA, 2008: 3353-3356.

[4] Lu B, Dibazar A, Berger T W. Nonlinear hebbian learning for noise-independent vehicle sound recognition[C]//IEEE International Joint Conference on Neural Networks, Hong Kong, China, 2008: 1336-1343.

[5] Rouas J L, Louradour J, Ambellouis S. Audio events detection in public transport vehicle[C]//Proceedings of Intelligent Transportation Systems Conference. Toronto, Ontario, Canada: IEEE Computer Society Press, 2006: 733-738.

[6] Chen J, Kam A H, Zhang J, et al. Bathroom activity monitoring based on sound[C]//Proceedings of the Third International Conference on Pervasive Computing. Munich, Germany: Springer-Verlag Berlin Heidelberg, 2005: 47-61.

(下转164页)

(5)由于是单线程执行,虽然流水线式多线程执行过程中,线程之间可能会出现相互等待,但由于单线程模式对CMP利用率极低,因此性能最差。而模式(3)和(4)虽然也是多线程执行,但由于PIDT线程的计算量较大,而BTT计算量相对较小,如果为BTT分配较多线程则会严重浪费处理器的计算资源,模式(2)则有所好转。模式(1)则充分利用了处理器的计算资源,而且线程之间因不能获得足够数据造成的等待时间较少,从而获得了最优的性能。

实验4 测试FABT-MAM与传统多线程执行模式的性能。传统多线程执行模式中,将投影-选择访问表中的数据页面均匀分配给 N 个连接执行线程。如图6所示,传统的多线程执行模式由于Cache访问性能较流水线式多线程执行模式差,而且流水线式多线程通过合理分配处理器资源和设置合理的缓存参数提高了其执行效率。因此流水线式多线程执行模式的性能优于传统的多线程执行模式。

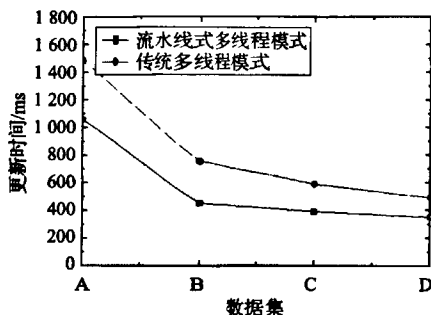


图6 不同线程执行模式性能比较

5 总结和展望

针对传统自顶向下B⁺-Tree访问模式的缺点,提出了一种支持频繁访问的索引结构FAB⁺-Tree。与传统的B⁺-Tree相比,

增加了Hash辅助索引结构用于支持自底向上访问,以及直接访问表和位矢量内存列表改进索引更新性能。实验结果表明FAB⁺-Tree的更新性能远远高于B⁺-Tree,而查询性能也较B⁺-Tree有较大提高。在需要频繁访问B⁺-Tree的应用领域,FAB⁺-Tree具有更高的价值。此外针对目前主流的多核处理器,提出了基于流水线的多线程FAB⁺-Tree访问模块,用于优化访问FAB⁺-Tree的QEP执行。并通过数据访问分析设置合理的线程执行参数,优化了线程的Cache访问性能,充分利用了处理器的计算资源。

参考文献:

- [1] Hardavellas N, Pandis I, Johnson R. Database servers on chip multiprocessors limitations and opportunities[C]//3rd Biennial Conference on Innovative Data Systems Research(CIDR),2007.
- [2] Cieslewicz J, Ross K A, Giannakakis I. Parallel buffer for chip multiprocessors[C]//DaMoN,2007.
- [3] Garcia P, Korth H F. Pipelined Hash-Join on multithreaded architectures[C]//DaMoN,2007.
- [4] Cieslewicz J, Ross K A. Adaptive aggregation on chip multiprocessors[C]//VLDB,2007.
- [5] Cieslewicz J, Ross K A. Data partitioning on chip multiprocessors[C]//DaMoN,2008.
- [6] Lee M L, Hsu W, Jensen C S, et al. Supporting frequent updates in R-trees: a bottom-up approach[C]//VLDB,2003.
- [7] 廖巍,熊伟,景宁,等.支持频繁更新的移动对象混合索引方法[J].计算机研究与发展,2006,43(5).
- [8] Rao J, Ross K R. Cache conscious indexing for decision-support in main memory[C]//VLDB,1999.
- [9] Rao J, Ross K R. Making B⁺-trees cache conscious in main memory[C]//SIGMOD,2000.

(上接155页)

- [7] Wang J C, Lee H P, Wang J F, et al. Robust environmental sound recognition for home automation[J]. IEEE Transactions on Automation Science and Engineering, 2008, 5(1): 25-31.
- [8] Reynolds D A, Quatieri T F, Dunn R B. Speaker verification using adapted Gaussian mixture models[J]. Digital Signal Processing, 2000, 10(1): 19-41.
- [9] Kittler J, Hatef M, Matas J, et al. On combining classifiers[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(3): 226-239.
- [10] Kimura F, Shridhar M. Handwritten numerical recognition based on multiple algorithms[J]. IEEE Transactions on Pattern Recognition, 1991, 24(10): 969-983.

- [11] Sound net of Princeton University[EB/OL]. <http://soundnet.cs.princeton.edu/OMLA/>.
- [12] The freesound project[EB/OL]. <http://www.freesound.org/index.php>.
- [13] Li Y. A quick classification for area environmental audio data based on local search tree[C]//IEEE International Conference about Environmental Science and Information Application Technology, Wuhan, China, 2009: 569-574.
- [14] Reynolds D A. Speaker identification and verification using Gaussian mixture speaker models[J]. Speech Commun, 1995, 17(1): 91-108.