

基于 PSOLA 与 DCT 的情感语音合成方法

李 勇,魏 珧,王柳渝

(重庆邮电大学 自动化学院,重庆 400065)

摘 要: 情感语音合成可以增强语音的表现力,为使合成的情感语音更自然,提出一种结合时域基音同步叠加(PSOLA)和离散余弦变换(DCT)的情感语音合成方法。根据情感语音数据库中的高兴、悲伤、中性语音进行韵律参数分析归纳情感规则,调整中性语音各音节的基音频率、能量和时长。使用 DCT 方法对基音标记过的语音段进行基音频率的调整,并利用 PSOLA 算法修改基音频率使其逼近目标情感语音的基频。实验结果表明,该方法比单独使用 PSOLA 算法合成的情感语音更具情感色彩,其主观情感的识别率更高,合成的情感语音质量更好。
关键词: 情感语音合成;离散余弦变换;基音同步叠加;基频;时长;能量

中文引用格式:李 勇,魏 珧,王柳渝. 基于 PSOLA 与 DCT 的情感语音合成方法[J]. 计算机工程,2017,43(12):278-282,291.
英文引用格式:LI Yong,WEI Dang,WANG Liuyu. Emotional Speech Synthesis Method Based on PSOLA and DCT[J]. Computer Engineering,2017,43(12):278-282,291.

Emotional Speech Synthesis Method Based on PSOLA and DCT

LI Yong,WEI Dang,WANG Liuyu

(School of Automation,Chongqing University of Posts and Telecommunications,Chongqing 400065,China)

[Abstract] Emotional speech synthesis is expected to make the synthesized speech more expressive. In order to synthesis more natural emotional speech signals, this paper proposes a new emotional speech synthesis method combining Pitch Synchronous Overlap Add (PSOLA) and Discrete Cosine Transform (DCT). The research builds up emotional rules for happy, sad, neutral speech. Through analyzing the prosody parameters, it can modify the each syllable of neutral speech's fundamental frequency, energy and duration based on the emotional rules. The combination method adjusts pitch frequency for which marked pitch through DCT method, and then adjusts the pitch frequency to approach the target emotional fundamental frequency by the PSOLA algorithm. Experimental results show that the proposed method is more sensitive than the PSOLA algorithm. The subjective emotion recognition rate is higher, and the synthesized emotion speech quality is better.
[Key words] emotional speech synthesis; Discrete Cosine Transform (DCT); Pitch Synchronous Overlap Add (PSOLA); fundamental frequency; duration; energy
DOI:10.3969/j.issn.1000-3428.2017.12.050

0 概述

随着人工智能的快速发展,语音已成为人机交互的媒介,情感语音是人工智能的一个重要分支,也是语音信号的重要组成部分^[1]。情感语音能较好地表达人们的意愿,相同的文本内容用语音能表达出不同的情感。随着语音合成技术的不断成熟,人们越来越希望能合成像人类一样有情感的语音。在众多的情感语音合成研究中,文献[2-4]通过修改韵律参数的基频、时长、能量等韵律参数来合成高兴、悲伤和生气等情感语音。基音同步叠加(Pitch Synchronous Overlap Add, PSOLA)算法^[5-6]可以灵活地调整波形信号的基频、时长和短时能量,需要注意的是 PSOLA 算法

对于基频修改范围为 10%,超过该范围修改后的语音会出现明显的机器音,音质会出现失真,这是由 PSOLA 算法本身的局限性造成的。而离散余弦变换(Discrete Cosine Transform, DCT)调整基音频率得到的语音质量较高^[7],如果两者按照固定比例进行调整,那么可以使基频、时长、能量参数达到目标情感语音所需的参数,在调整过程中准确的基音周期标注非常关键。常见的基音周期标注算法有倒谱法、自相关函数法、短时平均幅度差函数法、线性预测的基音检测法等。本文基于 PSOLA 算法对基频修改的局限性,提出一种基于 PSOLA 和 DCT 的情感语音合成方法。该方法使用自相关函数法估计基音周期,在中性语音的基础上对基频、时长和能量语

作者简介:李 勇(1976—),男,副教授、博士,主研方向为情感语音合成、认知网络;魏 珧、王柳渝,硕士研究生。
收稿日期:2016-12-06 修回日期:2017-01-13 E-mail:617519640@qq.com

音参数进行修改。

1 情感语料库选择

1.1 情感类型选择

由文献[8]可知,情感按作用来分可以分为积极情绪、消极情绪和中间情绪。积极情绪包括愉快、眷恋、骄傲、满意等,消极情感包括悲伤、恼怒、嫉妒、厌恶、懊悔等,中间情感即人在平静状态下表达的话语,听不出任何积极和消极的情感。本文各选取一个积极、消极和中间情感词,即将高兴、悲伤和中性作为本文情感语音类型的研究对象。

1.2 情感语音修改规则

本文引用中国社科院发布的情感语音库,其中高兴、悲伤和中性的句子各 50 句。为了分析情感语音规则特征,本文借助 Praat 语音分析软件对语音进行分析。对音库中每一句语音样本的每一个音节,从基频均值、时长、强度等角度进行参数提取与比较。文献[9]研究表明情感语音对音节的修改比对句子的修改效果要好,所以本文着重对音节的情感语音做出总结。分析中性、高兴和悲伤 3 种语气下的波形,提取其中一句语音的分析结果。“老师教授知识”的 3 种情感语音的每个音节的语音特征参数总结如表 1~表 3 所示。

表 1 中性语音的韵律参数

中性语音	基频/Hz	时长/s	能量/dB
老	98.66	0.178	79.50
师	158.00	0.295	72.65
教	156.10	0.202	74.48
授	93.25	0.148	67.40
知	134.80	0.217	72.72
识	79.01	0.190	61.34

表 2 高兴语音的韵律参数

高兴语音	基频/Hz	时长/s	能量/dB
老	122.60	0.165	88.90
师	212.90	0.279	81.34
教	204.60	0.159	79.93
授	137.70	0.165	78.31
知	201.70	0.202	83.30
识	97.48	0.253	66.82

表 3 悲伤语音的韵律参数

悲伤语音	基频/Hz	时长/s	能量/dB
老	88.31	0.187	76.80
师	133.20	0.299	69.33
教	139.80	0.205	70.92
授	79.90	0.156	64.02
知	120.90	0.221	69.25
识	71.92	0.209	58.87

总结情感规则在情感语音合成中有重要的意义。本文是在中性语音的基础上修改每个音节的韵律参数,达到目标情感语音的韵律特征参数值。通过大量的情感语料库分析发现,对于高兴情感语音,除了重音以外的音节,它们的基频要比中性语音的基频高出 20%~40%,而重音音节的基频要比中性语音基频高 40%~55%,时长略微偏短,强度也明显偏大,能量比中性语音能量高 8%~16%。对于悲伤语音而言,各个音节变化比较有规律,悲伤语音的基频比中性语音的基频要低 10%~18%,没有明显的重音词,所以对悲伤语音不做重音词归纳,它们的时长略微偏长,强度比中性语音强度要低 4%~6%。本文将根据这些情感语音规则进行情感语音合成。

2 情感语音合成方法

2.1 基音检测与标注

基音周期是声带(发浊音时)的振动周期,基频的倒数就是基音周期,基音周期是语音信号最重要的参数之一^[10],也是本文的重点研究对象。本文分析的基础是对音频信号进行基音标注。首先需对语音信号进行基音周期估算,由于语音信号不具有周期性,但是在语音处理中通常可以认为语音信号在 10 ms~30 ms 中处于稳态,因此称为短时分析技术。短时分析技术将语音信号分为一帧一帧的数据进行处理。本文使用短时自相关函数算法进行基音检测^[11]。

设语音信号加窗分帧后得到的语音信号为 $x_i(n)$,则短时自相关函数 $R_i(k)$ 为:

$$R_i(k) = \sum_{n=1}^{N-k} x_i(n)x_i(n+k)$$

(1)

其中, i 表示第 i 帧语音信号, N 代表帧长, k 代表时间延迟量。

短时自相关函数 $R_i(k)$ 具体如下性质:

- 1) $R_i(k)$ 是偶函数,有 $R_i(k) = R_i(-k)$ 。
- 2) 若短时自相关信号 $x_i(n)$ 为周期信号,则自相关函数 $R_i(k)$ 也为周期信号,且它们的周期相同。
- 3) $R_i(k)$ 在基音周期的整数倍点上达到最大值。

2.2 离散余弦变换

离散余弦变换算法是一种信号扩展和压缩的变换算法,在语音领域应用广泛。在反离散余弦变换回原信号时能较好地保持信号的能量集中性^[12]。文献[7]使用 DCT 算法对语音信号基音周期进行修改。首先从基音同步帧中提取线性预测的残差信号,然后使用 DCT 变换算法进行压缩和拉伸。设某一帧源语音的残差信号为 $\{s(n): 0 \leq n \leq N_1 - 1\}$,基音周期为 p_1 ,目标基音周期为 p_2 。

1) 使用 DCT 对源残差信号 $\{s(n); 0 \leq n \leq N_1 - 1\}$ 进行变换, 如式(2)所示。

$$S(0) = \frac{1}{\sqrt{N_1}} \sum_{n=0}^{N_1-1} s(n)$$

$$S(k) = \sqrt{\frac{2}{N_1}} \sum_{n=0}^{N_1-1} s(n) \cos \frac{(2n+1)k\pi}{2N_1}$$

$$k = 1, 2, \dots, N_1 - 1 \quad (2)$$

2) 对残差信号 $\{s(n); 0 \leq n \leq N_1 - 1\}$ 修改基音周期, 需要调整 IDCT 的点数 N_2 来实现, 其中 $N_2 = N_1 \cdot p_2/p_1$ 。在进行 IDCT 变换前, 需要对 DCT 系数做标准化处理, 以补偿信号能量的变化。然后, 比较 p_1 和 p_2 , 当 $p_1 < p_2$ 时, 信号的基音周期变长, 应该在 DCT 系数 $S(k)$ 后面补充 $N_2 - N_1$ 个 0; 反之, 截取 DCT 系数 $S(k)$ 的前 N_2 个系数, 舍弃后面的 $N_1 - N_2$ 个系数^[13]。在截取过程中要保证有用的数据部分不被截取, 尽量截掉 DCT 系数的高位部分, 因为高位部分的能量集中性不如低位部分好。

3) 经过截取和补零后的系数进行 N_2 点的反余弦变换, 获得目标基音周期的残差信号 $s(n)'$ 。

$$s(n)' = \frac{1}{\sqrt{N_2}} S(0) + \sqrt{\frac{2}{N_2}} \sum_{n=0}^{N_2-1} S(k) \cos \frac{(2n+1)k\pi}{2N_2}$$

$$n = 0, 1, \dots, N_2 - 1 \quad (3)$$

上述生成的目标残差信号生成算法只是对于浊音信号进行转换, 而没有对清音做改变^[14]。需要注意的是在进行反离散余弦变换 (Inverse Discrete Cosine Transform, IDCT) 变换前, 必须对 DCT 系数进行标准化处理, 补偿能量损失, 改变语音帧间叠加段的长度来补偿对语音时长造成的影响^[15]。

2.3 PSOLA 算法

20 世纪 80 年代末, 由 Moulines 和 Charpentier 提出的基音同步叠加技术^[5]被广泛应用于语音合成技术中, 能有效解决波形拼接问题, 可用于语音信号基频和时长的修改, 在合成中还能保持原始发音的主要音段特征^[16]。PSOLA 算法有时域基音同步叠加 (TD-PSOLA)、频域基音同步叠加 (FD-PSOLA)、线性预测基音同步叠加 (LPC-PSOLA) 3 种实现方式。本文使用时域基音同步叠加方式, 实现步骤具体如下:

1) 基音同步叠加分析

对原始语音信号进行连续的基音同步标记, 基音同步标记的准确性非常关键, 将原始语音信号与一系列基音同步的窗进行加窗处理, 基音同步分析是 TD-PSOLA 方法的核心, 其主要内容为: 基频检测和间隔标记。令 $s_m(n)$ 表示加窗的短时间信号, 本文使用的是汉明窗, 长度为基音周期较短者长度的 2 倍, 得到一组有重叠的短时分析信号, 如式(4)所示。

$$s_m(n) = h_m(t_m - n) s(n) \quad (4)$$

其中, t_m 是基音标记点, h_m 是窗函数序列, $s(n)$ 为源语音信号。

2) 基音同步修改

首先根据源语音信号和目标语音波形之间的基频和时长的修正要求, 建立源语音波形和合成语音波形之间的映射关系, 再由映射关系合成序列短时信号, 则基音同步修改完成。

3) 基音同步合成

将待合成的短时信号序列在目标基音标注处排列, 并重叠相加得到合成语音波形。基频的修改通过改变基音标记序列的间隔来实现, 时长的修改通过增加或者删除基音标记序列来实现。由源基音标注点得到最后合成信号的基音同步标注点 t_q 后, 根据原始波形与合成波形误差最小的原则, 得到:

$$\bar{s}(n) = \frac{\sum_q a_q \bar{s}_q(n) h_q(t_q - n)}{\sum_q h_q^2(t_q - n)} \quad (5)$$

其中, $\bar{s}(n)$ 代表修改后的合成波形, a_q 代表用来合成的能量补偿变化因子, 可以调整合成语音的幅值, 实现合成语音音强的改变, $\bar{s}_q(n)$ 代表合成语音的短时信号, 由短时信号 $s_m(n)$ 经过变换得到的, h_q 代表合成的窗函数序列, t_q 代表合成语音的基音同步标志。

3 情感语音合成实验

为了验证本文提出的基于时域基音同步叠加和离散余弦变换的情感语音合成方法的有效性, 以及分析各参数对情感合成的作用, 将进行 2 组实验。实验内容包括: 1) 单纯通过使用 PSOLA 算法来合成情感语音; 2) 同时使用 PSOLA 和 DCT 算法来合成情感语音。韵律参数修改规则根据 1.2 节总结的规则进行修改。实验分别用 2 种方法合成 30 句情感语音, 对合成的情感语音进行听力测试, 通过平均意见得分 (Mean Opinion Score, MOS) 结果比较 2 组实验的结果。从听力测试的结果可以看出, 同时使用 PSOLA 和 DCT 合成的情感评分更高, 说明本文提出方法是切实可行的。

3.1 基音标注

基音周期的检测是依靠短时自相关函数的这些性质进行估计, 通过比较原始信号及其延迟后的信号的相似性可以确定基音周期, 还可以根据短时自相关函数的 2 个最大值间的距离来估计基音周期。检测出基音周期后, 需要进行基音脉冲标注。基音标注一般都是对浊音进行标注, 清音一般不做考虑。浊音段的标注都是在基音脉冲的短时能量的波峰或者波谷时刻上。“老师教授知识”语音的部分脉冲标注如图 1 所示。

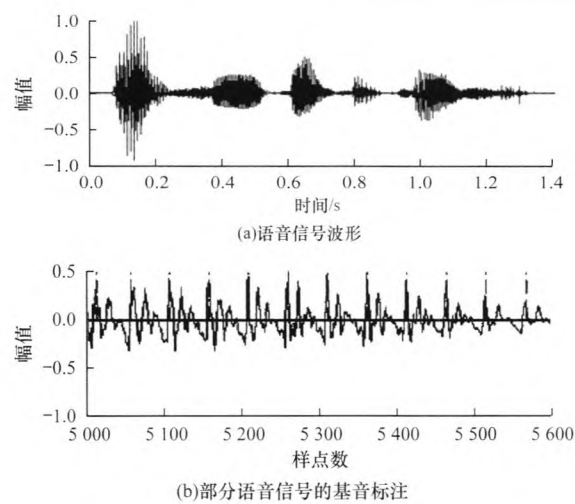


图 1 “老师教授知识”语音的部分脉冲标注

3.2 基于 PSOLA 和 DCT 的语音信号修改

图 2 是本文情感语音合成方法的具体流程。在合成目标语音之前,首先分析语音信号各个音节的基频调整比例(P_ratio)、能量调整比例(E_ratio)和时长调整比例(T_ratio),之后对于源语音按照 DCT 方法调整基频(P_ratio_1)。由文献[7]可知,基频修改幅度在 0.8 ~ 1.3 范围内的语音质量最好。所以,本文使用 DCT 修改基频时的基频修改幅度保持在该范围内,当超出该范围时使用 PSOLA 算法对语音基频再做修改。修改方法是对 DCT 处理后的结果进行基音标注,然后按照基频调整比例($P_ratio_2 = P_ratio / P_ratio_1$)、时长调整比例(T_ratio)、能量调整比例(E_ratio)进行调整,实现 3 个参量各音节的调整,得到最终的情感语音合成信号波形。基于 PSOLA 和 DCT 的语音信号修改流程如图 3 所示。

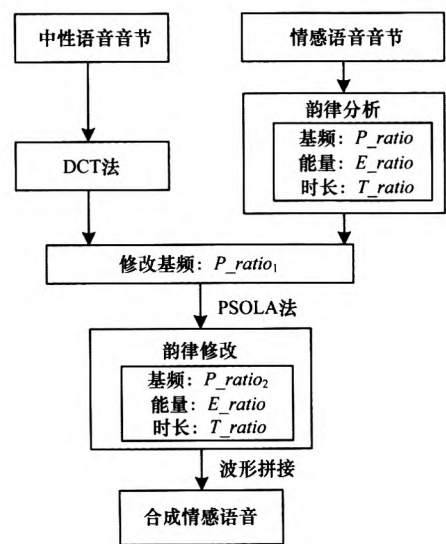


图 2 情感语音合成流程

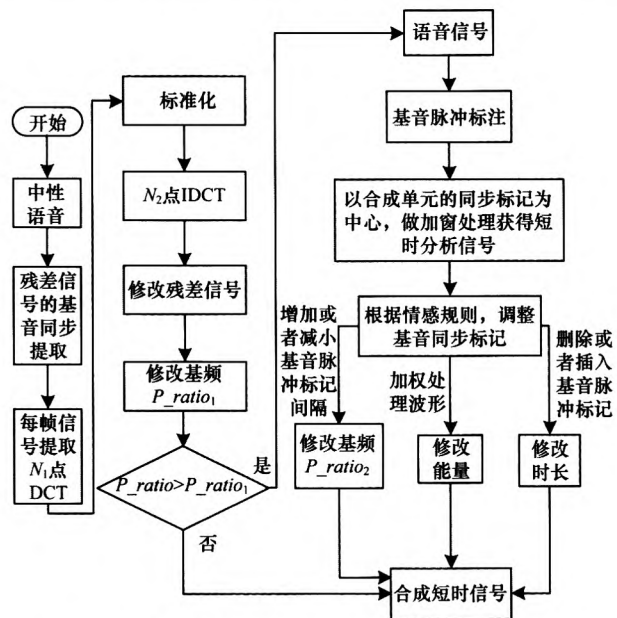


图 3 基于 PSOLA 和 DCT 的语音信号修改流程

对语句“老师教授知识”分别用 PSOLA 和 PSOLA + DCT 算法进行修改,合成高兴和悲伤的情感语音,人工朗读的情感语音波形和修改后合成的波形如图 4 ~ 图 6 所示。由于本文基于每个音节对语音进行修改,因此能更好地体现出情感语音的特点。图 4 ~ 图 6 分别是由真人录制的情感语音波形、PSOLA 算法合成的语音波形和 PSOLA + DCT 算法合成的语音波形。在参数修改后的波形对比图中,连续的曲线代表语音的能量曲线,其中基频曲线是不连续的,因为清音的基频为 0。可以看出,2 种情感语音合成方法在音节的韵律上与目标情感基本一致,也就是说 2 种合成方法对信号的修改都是理想的,在该前提下,通过听觉实验对语音进行评价,发现使用 PSOLA 合成的情感语音在某些音节上有失真的情况,而使用 PSOLA + DCT 对于语音修改的句子不会有失真的情况。结果表明,使用 DCT 和 PSOLA 算法相结合的情感语音合成方法更能使中性语音逼近目标情感语音,提高语音合成质量。

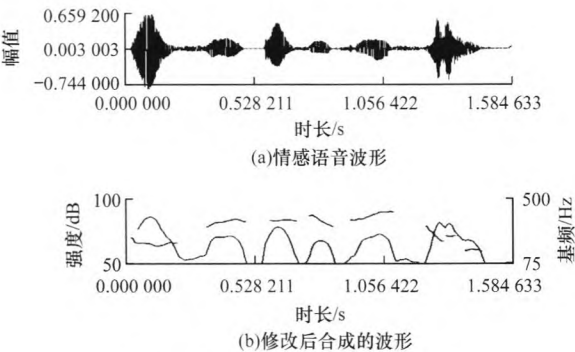


图 4 人工朗读的高兴情感语音

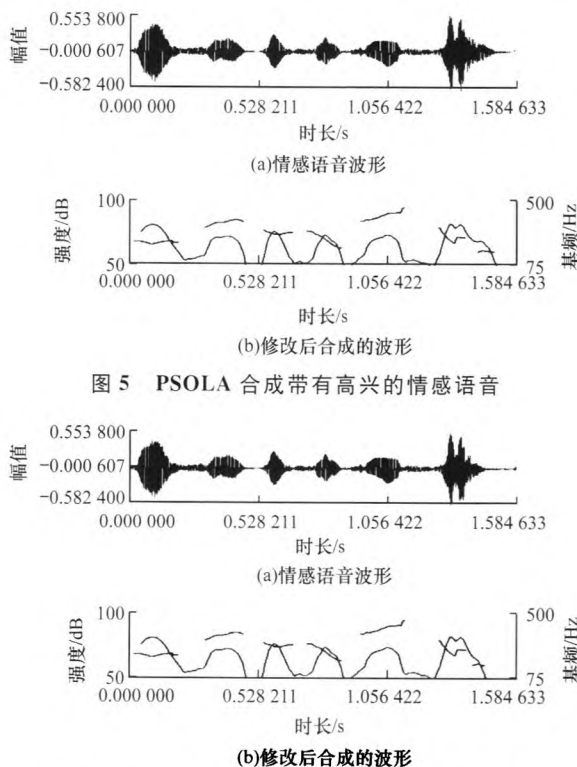


图 5 PSOLA 合成带有高兴的情感语音

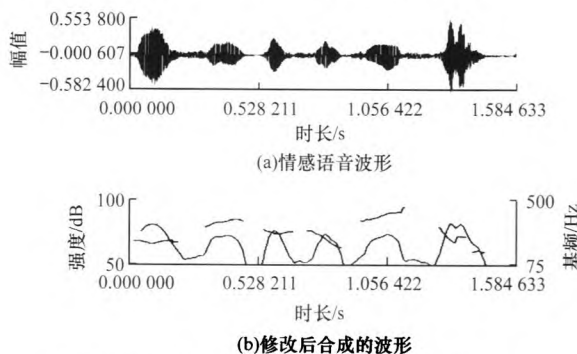


图 6 PSOLA + DCT 合成带有高兴的情感语音

4 合成效果评价

为了比较语音合成效果,分别对 2 种方法合成的情感语音进行人工评价并且打分。对比实验中情感语音合成方法的有效性以及效果,本文采用语音合成领域较常规的性能评价方法——听者实验,评分按照 MOS 分进行评价。在语料库中选取了 30 句没有带任何情感色彩词的中性语音,按照以上 2 种方法进行高兴和悲伤的情感语音合成。将合成的语音顺序进行打乱并且没有标注任何情感,然后邀请 10 位未经任何训练的同学(5 男、5 女,平均年龄为 24 岁)进行试听并且打分。在评估过程中,受试者可以反复试听然后打分,打分规则是 1 分~5 分,依次被附注为明显机器音、不太自然、较自然、自然、流畅。结果取 10 位同学打分的平均分,如图 7 所示。从实验结果可以看出,使用 PSOLA + DCT 方法的情感语音合成识别率要比单独使用 PSOLA 高,所以本文通过 PSOLA 算法来改进合成情感语音合成质量的方法是有效的。

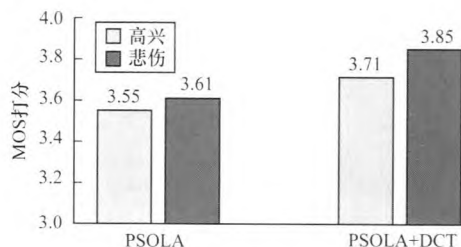


图 7 情感语音合成效果

5 结束语

本文分析情感语音合成中需要调整的语音参量,提出基于 DCT 和 PSOLA 的情感语音合成方法。实验结果表明,同时使用 DCT 和 PSOLA 合成的语音情感识别率较高,证实了该方法合成的语音情感表达要优于单独使用 PSOLA 的方法,并且情感语音表达更加接近目标语音,使测试者感受到更加自然的语音,提高了人机交互的能力。下一步将研究频谱参数对情感语音的影响,以便能合成更自然的情感语音,从而应用于人机对话中。

参考文献

- [1] RUSSELL S J, NORVIG P, CANNY J F, et al. Artificial Intelligence: A Modern Approach [M]. Upper Saddle River, USA: Prentice Hall, Inc., 2003.
- [2] 邵艳秋, 韩纪庆, 王卓然, 等. 韵律参数和频谱包络修改相结合的情感语音合成技术研究[J]. 信号处理, 2007, 23(4): 526-530.
- [3] WU C H, HSIA C C, LEE C H, et al. Hierarchical Prosody Conversion Using Regression-based Clustering for Emotional Speech Synthesis[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2010, 18(6): 1394-1405.
- [4] HAMADA Y, ELBAROUGY R, AKAGI M. A Method for Emotional Speech Synthesis Based on the Position of Emotional State in Valence-activation Space[C]//Proceedings of Signal and Information Processing Association Annual Summit and Conference. Washington D. C., USA: IEEE Press, 2014: 1-7.
- [5] MOULINES E, CHARPENTIER F. Pitch-synchronous Waveform Processing Techniques for Text-to-speech Synthesis Using Diphones[J]. Speech Communication, 1990, 9(5/6): 453-467.
- [6] GOVIND D, PRASANNA S R M. Expressive Speech Synthesis: A Review[J]. International Journal of Speech Technology, 2013, 16(2): 237-260.
- [7] MURALISHANKAR R, RAMAKRISHNAN A G, PRATHIBHA P. Modification of Pitch Using DCT in the Source Domain[J]. Speech Communication, 2004, 42(2): 143-154.
- [8] 耿德勤. 医学心理学[M]. 南京: 东南大学出版社, 2003.
- [9] YADAV J, RAO K S. Generation of Emotional Speech by Prosody Imposition on Sentence, Word and Syllable Level Fragments of Neutral Speech[C]//Proceedings of 2015 International Conference on Cognitive Computing and Information Processing. Washington D. C., USA: IEEE Press, 2015: 1-5.
- [10] 郑继明, 王劲松. 语音基音周期检测方法[J]. 计算机工程, 2010, 36(10): 273-275.
- [11] 赵力. 语音信号处理[M]. 北京: 机械工业出版社, 2003.
- [12] RAO K R, YIP P. Discrete Cosine Transform: Algorithms, Advantages, Applications [J]. Discrete Cosine Transform Algorithms Advantages Applications, 1990, 14(6): 507-508.

(下转第 291 页)

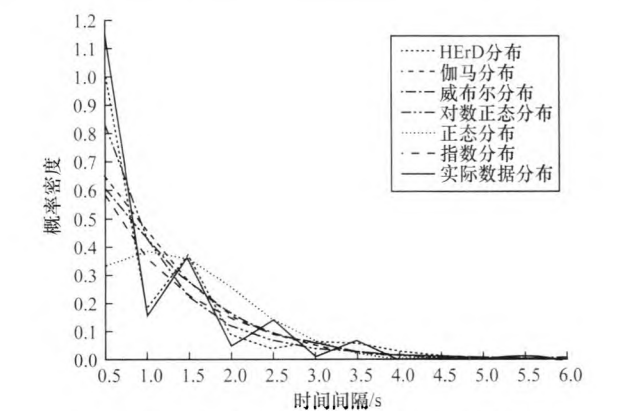


图 16 环球广场周五早高峰电梯客流到达时间规律二次拟合结果

表 11 环球广场 HErD 分布参数约简前后拟合指标对比					
日期	c_x^2	HErD (约简前)	HErD (约简后)	对数正态 分布	伽马分布
星期一	0.847	0.982 20	0.968 9	0.811 40	0.681 0
星期二	0.864	0.983 10	0.938 6	0.805 20	0.672 4
星期三	0.796	0.977 90	0.961 1	0.806 60	0.684 2
星期四	0.805	0.984 70	0.977 5	0.810 90	0.684 8
星期五	0.837	0.980 90	0.968 4	0.812 10	0.682 1
拟合度均值		0.981 76	0.962 9	0.809 16	0.680 9

4 结束语

本文提出一种将基于 PH 的分布引入电梯客流分布拟合中,对成都市较具代表性的两幢 CBD 写字楼工作日早高峰电梯客流到达规律进行拟合,并与 5 个常用分布的拟合效果进行对比。实验结果表明,HErD 拟合效果最优($R^2=98.176\%$)且参数约简对其精度影响较小。在后续研究中将扩大研究对象的范围,进一步验证 HErD 的拟合效果,并通过构建电梯规划阶段的客流参数与 HErD 的参数关系,对电梯规划阶段的客流规律进行预测。

参考文献

[1] 张亦辉.组合模型在电梯客流量预测中的仿真研究[J].计算机仿真,2011,28(11):192-195.
[2] SIIKONEN M L. Customer Service in an Elevator System During Up-peak [J]. Transportation Research Part B:Methodological,1997,31(2):127-139.

[3] 孙 东,张惠侨.电梯服务系统的理论建模与分析[J].上海交通大学学报,1996,30(3):113-119.
[4] 宗 群,魏利剑,程义菊,等.基于马尔可夫网络排队论的电梯交通建模及应用[J].天津大学学报(自然科学与工程技术版),2005,38(1):9-13.
[5] LEE Y,TAI S K,CHO H S,et al. Performance Analysis of an Elevator System During Up-peak[J]. Mathe-matical & Computer Modelling,2009,49(3/4):423-431.
[6] AL-SHARIF L, ALQUMSAN A M A, GHANEM W, et al. Modelling of Elevator Traffic Systems Using Queuing Theory [J]. Journal of the Operational Research,2014,41(2):192-199.
[7] 简学琴.电梯服务系统的排队模型[J].安徽师范大学学报(自然科学版),2006,29(5):415-418.
[8] 刘明珠,张国宝.基于排队系统的城市轨道交通进站检票机配置[J].城市轨道交通研究,2004,7(5):34-37.
[9] 李又玲,常致全.基于 Gibbs 采样与概率分布估计的移动云数据存储[J].计算机工程,2017,43(1):13-19.
[10] 黄 卓,潘 晓,郭 波. PH 分布数据拟合的数值加速 EM 算法[J]. 计算机工程,2008,34(14):1-3.
[11] JIANG Y,HU L,ZHU J,et al. PH Fitting of the Arrival Interval Distribution of the Passenger Flow on Urban Rail Transit Stations [J]. Applied Mathematics and Computation,2013,225:158-170.
[12] HU L,JIANG Y,ZHU J,et al. A PH/PH (n)/C/C State-dependent Queuing Model for Metro Station Corridor Width Design[J]. European Journal of Operational Research,2015,240(1):109-126.
[13] ALFA S A,ZHAO. Overload Analysis of the Ph/Ph/1/K Queue and the Queue of the M/G/1/K Type with Very Large K [J]. Asia Pacific Journal of Operational Research,2001,17(2):123-135.
[14] 黄 卓,潘 晓,郭 波.基于 EM 算法的 ACPH 分布数据拟合方法[J]. 计算机工程,2008,34(4):75-78.
[15] THMMLER A, BUCHHOLZ P, TELEK M. A Novel Approach for Phase-Type Fitting with the EM Algorithm[J]. IEEE Transactions on Dependable & Secure Computing, 2006,3(3):245-258.
[16] 田乃硕.休假随机服务系统——数理科学系列[M].北京:北京大学出版社,2001.
[17] 胡 路.考虑稳定性的城市轨道交通车站交通服务设施系统优化配置理论与方法[D].成都:西南交通大学,2014.

编辑 刘 冰

(上接第 282 页)

[13] 简志华,杨 震.一种用于语声转换系统的 LPC 残差信号生成算法[J].信号处理,2008,24(5):762-765.
[14] KAIN A B. High Resolution Voice Transformation[D]. Portland,USA:Oregon Health & Science University,2001.
[15] LEE K S. Statistical Approach for Voice Personality

Transformation[J]. IEEE Transactions on Audio, Speech, and Language Processing,2007,15(2):641-651.
[16] 陈 愉,张宗红,李 炜,等. PSOLA 技术在汉语文-语转换系统中的应用[J]. 计算机工程,2000,26(1):84-86.

编辑 陆燕菲