

文章编号: 2096-1618(2018)04-0359-06

一种基于超混沌的个性化耳鸣康复音合成方法

金宁敏, 何培宇, 潘帆, 陈杰梅

(四川大学电子信息学院, 四川 成都 610065)

摘要:耳鸣主要表现为非外部声音产生的听觉感知。因其发病率较高、严重影响人的正常生活而越来越受到关注。研究表明,相似而不重复的音乐可作为耳鸣康复音来缓解耳鸣症状。但现有的合成音乐存在无法与试听者个性化匹配和试听效果较差等问题。针对这些问题,基于超混沌算法与主旋律提取,提出一种个性化合成耳鸣康复音的新方法。结果表明,此方法合成的康复音在保证整体相似不重复的前提下,试听效果更舒缓自然,且在耳鸣模型中对耳鸣有一定抑制作用,对耳鸣的治疗具有一定的参考价值。

关键词:信号与信息处理;语音合成;耳鸣康复音;个性化;超混沌;主旋律提取

中图分类号: TN912.3

文献标志码: A

doi: 10.16836/j.cnki.jcuit.2018.04.002

0 引言

耳鸣,美国国家标准协会将其定义为非外部声音的听觉感知。研究结果表明成人中大约有 10% ~ 15% 的人有耳鸣的经历^[1]。耳鸣的并发症严重影响人的正常生活,其病因复杂,机理尚不明确^[2],因此探寻耳鸣康复方法一直都是耳科学界的热门话题之一。

2014 年,德国音乐治疗研究中心通过研究正式确认音乐治疗耳鸣疗效明显,听音乐对治疗慢性和急性耳鸣都有效果。该研究中心通过核磁共振试验得出结论:音乐疗法的确在耳鸣患者脑部带来可测量的变化^[3]。然而,Hann 等^[4]研究表明,音乐对短期内缓解耳鸣有效,放松心情,但重复而单调的音乐会唤起耳鸣患者潜意识的记忆,不能长时间供耳鸣患者作为治疗音使用。为此,陈杰梅等^[5-6]研究了基于分形或混沌算法合成耳鸣康复音。但这些合成的耳鸣康复音多由混沌或分形算法产生的序列直接映射成音调值及音符时值,虽然满足相似不重复的特点,但并不符合音乐作曲的规律。此外,这些方法没有专门针对试听者进行匹配,混沌分形序列存在不可控等因素,故合成康复音存在无法与试听者个性化匹配和试听效果较差等问题。

针对上述问题,从谱曲的旋律发展流程^[7-8](即先确定主旋律,再发展主旋律,通过重复或者改变主旋律来体现歌曲主题)得到启发,提出一种基于超混沌的个性化耳鸣康复音合成方法。此方法先令试听者对样本音乐进行打分,经过筛选后选择合适的样本音乐进

行主旋律提取,并采用不同发展方法进行变换得到旋律片段组。同时用混沌算法生成相似不重复的序列,将序列映射成旋律片段组中对应的编号,最后借助乐器数字接口(musical instrument digital interface, MIDI)技术合成与试听者相匹配的耳鸣康复音。考虑到此方法是将混沌序列映射成不同的旋律变换片段,可能存在局部相似性较大的缺点,而超混沌系统在局部上比混沌系统具有更强的不稳定性^[9],故采用超混沌系统产生的序列来映射成旋律变换片段组的编号。

实验分析表明,这种方法产生的耳鸣康复音在满足相似不重复的前提下,试听效果大幅度提高。此外,由于该方法针对不同试听者有试听、评分、选优的过程,故能到达个性化匹配生成耳鸣康复音的效果,并验证了合成音乐在耳鸣模型对耳鸣有抑制作用。

1 个性化合成耳鸣康复音

合成耳鸣康复音的整体思路为:首先,让试听者试听样本音乐(包括古典音乐、乡村音乐、流行音乐等不同风格音乐)。让试听者对所听到的音乐作品进行评分,选出试听者偏好的音乐作品。然后,对所选出的样本音乐进行主旋律提取,对提取出的旋律采用重复、变化等方法进行变换,对变换后的片段组进行编号。同时采用超混沌算法产生相似不重复的序列。最后,把混沌序列映射成旋律变换组中对应的编号,最终根据编号顺序利用 MIDI 合成机制合成耳鸣康复音。

此方法是根据试听者试听后筛选出样本音乐,之后对样本音乐进行提取、变换,所以能够保证合成的康复音满足试听者的需求,达到个性化匹配的效果。此

外,由于是通过超混沌序列映射成变换后的旋律片段,能保证合成的康复音在满足相似不重复的前提下,符合音乐旋律的发展规律,达到良好的试听效果。具体流程图如图1所示。

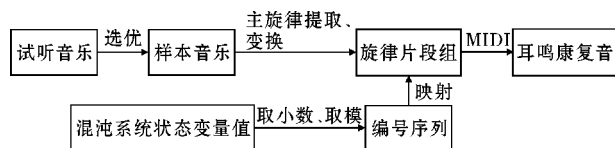


图1 个性化合成耳鸣康复音流程图

2 具体实现过程

2.1 判优

先提供古典、乡村、流行3种类型的音乐各3个样本音乐,试听者对9个音乐作品进行试听,从舒缓度、流畅度、接受度3方面进行打分。选出综合得分最高的音乐片段,即视为与试听者相匹配的样本音乐。

2.2 提取主旋律

主旋律,在一部音乐作品或一个乐章行进过程中再现或变奏的主要乐句或音型^[10]。对于一部音乐作品或者一个乐章,通常都有一个主题,其表达是通过主旋律的反复重申^[11]。因此,提取音乐作品的主旋律,即视为判断出该作品中重复出现最多的音乐片段。提取主旋律第一步就是要对一个音乐作品进行分段。分段的方法有小节划分、标志划分、智能划分等,而本文采用的是对音乐作品的小节进行分段。根据基本乐理定义,1个乐段分为2个乐句,1个乐句由2个乐节构成,1个乐节由2个动机(小节)构成^[7]。一个完整的乐句对应4个小节,故本文采用4个小节对音乐进行分段。

对音乐作品分段后,统计各音乐片段重复出现的次数,出现最多的即判为音乐作品的主旋律,记录下主旋律片段的音调值与音符时值。

2.3 旋律发展方法

一个音乐作品的发展方法多种多样,但主要分为重复和变化这两大类^[7]。主要采用的重复方法有严格重复、非严格重复、严格模进、非严格模进4种方法;变化方法有逆行、倒影、产生随机片段3种方法。

严格重复:对应所提取主旋律的音调值与音符时值的完全重写。

非严格重复:分为3种模式,分别在主旋律的头

部、中间及尾部进行音调值与音符时值的替换,试验中随机选择一种模式。

严格模进:音调值随机选择上行或者下行,音符时值保持不变。如音调值为78的音符进行2度上行模进,音调值就变为79。

非严格模进:对主旋律的前一半音符值进行3度模进,模进方向随机选择,后一半音符值不变。片段的音符时值保持不变。如音调值为78的音符进行3度下行模进,值就变为76。

逆行:对主旋律的音调值与音符时值分别进行前后倒置。如音符值序列72、74、75、79、84将变成84、79、75、74、72。

倒影:对主旋律的音调值进行轴对称倒映,如音符值序列72、74、75、79、84的对称轴为最小值72与最大值84的均值78,轴对称变换后序列变为84、82、81、77、72,音符时值也进行轴对称变化。

随机片段:根据混沌算法产生随机序列,按照随机序列的顺序对主旋律的音调值与音符时值进行重新排序。

根据上述7种变换方法依次对变换后的旋律片段进行编号,产生以供合成所需的旋律片段组。

2.4 生成音乐

采用超混沌算法产生相似不重复的序列,将序列进行取小数部分,取模映射等步骤产生与旋律片段组对应的编号序列。根据编号序列,利用MIDI合成机制生成耳鸣康复音。

2.4.1 超混沌算法生成序列

具有两个或者两个以上的正的Lyapunov指数的混沌系统称为超混沌系统。2004年,Li等^[12]通过设计的非线性状态控制器从Chen系统中得到了超混沌Chen系统,并对其动力学行为进行研究。Chen系统模型的表达式为:

$$\begin{cases} \dot{x} = a(y - x) + w \\ \dot{y} = dx - xz + cy \\ \dot{z} = xy - bz \\ \dot{w} = yz + rw \end{cases} \quad (1)$$

式中,参数 $a=35, b=3, c=12, d=13, r=0.79$ 。求解微分方程组,得到4个维度状态变量 x, y, z, w 的数值解。

2.4.2 混沌序列映射成旋律片段组编号

状态变量是微分方程的数值解,存在局部递增的问题,故只取数值解的小数部分。然后将数值解的小数部分进行模为7的取模映射,生成与旋律片段组对应的1至7的整数序列。例如,整数1就对应经严格

重复变换的旋律片段;整数 4 对应经非严格模进变换的旋律片段;整数 7 对应随机序列产生的旋律片段。

2.4.3 MIDI 合成机理生成主旋律轨道音频

每个旋律片段都有其对应的音调值和音符时值序列。根据映射后的整数序列,对应成旋律片段的音调值和音符时值。将组合后的音调值和音符时值保存到 MIDI 轨道信息中,利用 MIDI 合成机理生成主旋律音频。

2.4.4 和声轨道的处理

MIDI 文件由多个音符轨道构成。采用同主旋律轨道相同的处理方法,对音乐的其他和声轨道进行相同方式处理,合成和声轨道音频,最终生成多轨道的耳鸣康复音。

3 仿真结果分析

3.1 Lorenz 混沌系统与 Chen 超混沌系统比较

前文提到相对混沌系统、超混沌系统在局部具有更大的不稳定性。为了研究两种混沌系统的稳定性,分别研究两种系统对初始值的敏感度及状态变量的波动范围。

Lorenz 系统模型的表达式为:

{ x-dot = a(y - x) y-dot = cx - y - xz z-dot = xy - bz (2)

式中,参数 a=10,b=2,c=28。

比较式(1)与(2),可以看出 Chen 超混沌系统比 Lorenz 多了一个维度 w,该维度变量 w 对维度变量 x 有影响。

3.1.1 初始值敏感度比较

混沌运动均有一个基本特征,即对初始值的极端敏感性。对混沌系统的状态变量加上一个微小的增量,对混沌的状态变量有运动趋势有强烈的影响,并且导致对系统长时间的不可预测性。

为了比较两个混沌系统对初始值的敏感度,状态变量取相同的初始值 x0=1.2,y0=1.3,z0=1.6。对 x0 加上一个很小的增量 0.01,得到变换后的初始值:x1=1.21,y1=1.3,z1=1.6。分别观察两个系统在两组不同状态变量初始值下的运动状态。图 2、图 3 分别表示 Lorenz 系统和 Chen 超混沌系统在两组不同状态初始值下状态变量 x 在迭代过程的变化情况,其中‘.’表示 x0 组初始值下的状态变量,‘—’表示 x1 组初始值下的状态变量,可以看出,Chen 超混沌系统对初值的变化更为敏感。

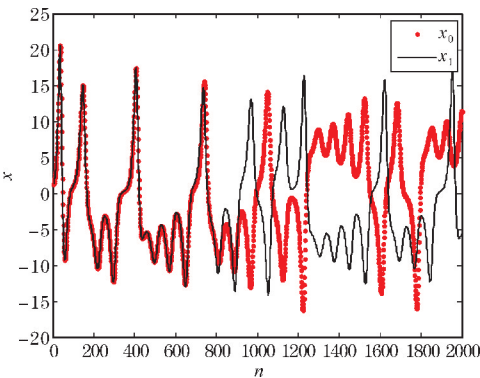


图 2 Lorenz 系统初始值敏感度

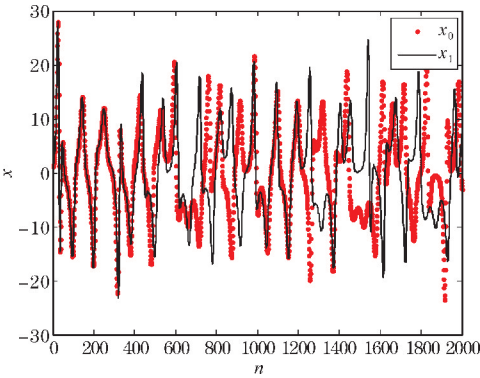


图 3 Chen 系统初始值敏感度

分别计算两种系统在两组不同初始值下状态变量 x 差值的均方误差,其中 Lorenz 系统均方误差为 117.8364,Chen 超混沌系统为 140.7795。对实验结果进行 100 次统计平均,即初始值分别从 1.2 到 1.21,1.21 到 1.22...进行 100 次增量得到 100 组初始值。分别求解在 100 组初始值下两种混沌系统状态变量差值的均方误差,最后平均得到 Lorenz 系统均方误差均值为 92.7739,Chen 超混沌系统为 135.0309。从结果可以看出,相对于 Lorenz 系统,Chen 超混沌系统对初始值的变化更为敏感,初始值的变化,对混沌状态的运动趋势有更大的影响,导致系统更大的不可预测性。

3.1.2 状态变量波动范围比较

为了研究两种混沌系统的状态变量的波动范围,对两个系统也设置了相同的状态变量初始值:x0=1.2,y0=1.3,z0=1.6。x0 进行 100 次增量,每次增量 0.01,得到 100 组初始值。对实验结果进行 100 次统计平均,得到状态变量 x 的统计均方误差。Lorenz 系统为 51.2015,Chen 超混沌系统为 72.0309。实验结果表明在相同初始值下,Chen 超混沌系统产生的序列的波动范围更大。

由这两个混沌特征的统计值可以看出,相对 Lorenz 混沌系统,Chen 超混沌系统对初始值的变化更为敏感,状态变量的波动更大,具有更强的不稳定性。而这些特征正好满足所需混沌序列的要求。故采用 Chen 超混沌系统产生所需的混沌序列。

3.2 1/f 波动分析

“1/f”波动介于白噪声与布朗噪声之间,在局部呈无序状态,而在宏观上具有一定的相关性。它的功率谱密度与频率成反比,是一种能令人感到舒适的波动。满足1/f波动的音乐满足相似不重复的特点^[13-14]。

功率谱 $S(f)$ 表示在频率 f 处的功率谱密度,它表示信号时间的相关性。 $S(f)$ 的关系可表示为:

$$S(f) \propto f^{-\beta} \tag{3}$$

其中 β 表示频率 f 的指数。当 $0.5 \leq \beta \leq 1.5$ 时,可以视信号符合 1/f 波动^[2]。仿真实验中对式(3)左右两边取对数处理,得到:

$$\lg(S(f)) \propto -\beta \lg(f) \tag{4}$$

先求取序列的功率谱密度,之后功率谱密度与频率分别取对数,描绘出取对数后频率对应的功率谱密度,并画出拟合曲线。具体实验中分别对合成康复音的音调值、音符时值序列进行分析。

求取音调值序列的功率谱密度 $S(f)$,将 $S(f)$ 与频率 f 之间的关系在取对数后进行分析,拟合曲线如图 4 所示。拟合曲线斜率为-0.9364,可以看出,音调值序列满足“1/f 波动”。

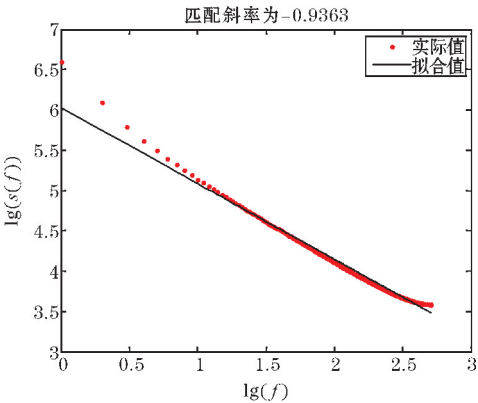


图 4 音调值序列功率谱拟合曲线

对音符时值序列进行相同的分析,拟合曲线如图 5 所示,拟合曲线斜率为-0.9278。音符时值序列也满足“1/f 波动”。

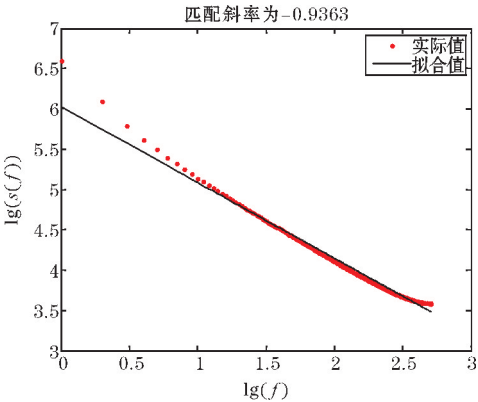


图 5 音符时值序列功率谱拟合曲线

由实验结果可知,由本方法合成的音调值、音符时值满足“1/f波动”,合成的音乐可以使人感到舒适、和谐,且局部与整体具有相似性^[13-14]。

3.3 耳鸣模型激励刺激分析

文献[15]提出了一种基于 LMS 算法的自适应耳鸣模型,根据耳蜗中不同部位对频率响应不同的特性,利用 FIR 带通滤波器将听觉通路分为不同的分路。各分路信号进入后,与该路中受听毛细胞健康程度影响的参数相乘后再与同外部声音输入无关的自发信号混合并进入听觉中枢系统进而模拟人耳系统。由于受听毛细胞健康受损,导致其参数变化引起系统自身的异常反馈,最终形成耳鸣。为了验证合成的耳鸣康复音是否能在一定程度抑制耳鸣,则需要将合成的音频信号作为输入信号激励耳鸣模型,观察系统的输出功率。实验结果如图 6 所示。

20 s前为耳鸣模型的原始输出,视为耳鸣症状;在 20 ~ 60 s加入耳鸣康复音作为刺激信号,由图 6 可以看出输出功率减小,视为耳鸣症状得到逐渐抑制;而60 s撤去刺激信号,输出功率又逐渐上升,视为耳鸣症状复发现象。仿真实验表明由本方法合成的耳鸣康复音可以抑制耳鸣症状。

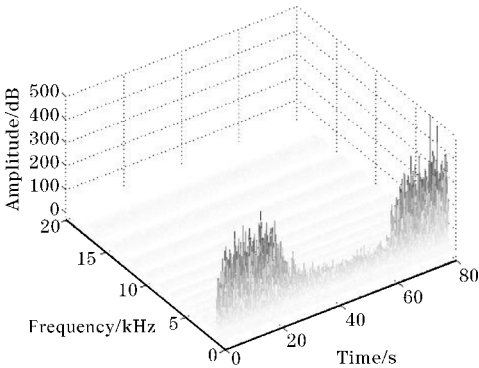


图 6 耳鸣模型输出时频图

3.4 仿真试听

文献[5-6]分别利用分形、混沌算法产生的序列直接映射成音调值及音符时值合成耳鸣康复音 TRS_1、TRS_2。而本文基于超混沌算法与主旋律提取,个性化合成耳鸣康复音 TRS_3。试听实验在普通人群中招募 30 名(15 名男性、15 名女性)不同行业(学生、公司职员为主)志愿者作为试听人员,平均年龄 34.5 岁(22 ~ 61 岁)。试听环境为一个安静、舒适的实验室,使用的播放设备是 IPOD CLASSIC,接收耳机型号为 sennheiser PC310。音频文件时长 5 min,格式为 MIDI,音乐轨道为 3,音源为 Windows 默认的 Roland 公司

128 个音色库的通用合成器。对于每一位志愿者都会进行 3 组重复测试。测试在试听者相对放松情况下进行,每种康复音播放时间有 5 min,3 种康复音试听之间有 1 min 休息时间,每组测试之间有 5 min 休息时间。试听过程中试听者分别从音乐舒缓度、自然度、重复度、突变度等方面在 0 ~ 10 分范围(分数越高表示该项特征越强)对 3 种音乐康复音进行打分。统计平均结果如表 1 所示。

表 1 音乐评分表

耳鸣康复音	TRS_1	TRS_2	TRS_3
舒缓度	3.2	4.0	6.0
自然度	2.0	4.6	8.0
重复度	4.0	3.2	4.1
突变度	6.0	4.6	3.8

试听结果表明,相对于文献[5-6]合成的耳鸣康复音 TRS_1、TRS_2,本文合成的耳鸣康复音 TRS_3 虽然在重复度上略微上涨,但舒缓度有所提升,突变情况也有所减少。最为重要的是在音乐的自然度上有大幅度提升。结果表明,合成的耳鸣康复音在保证相似不重复的前提下,试听效果更舒缓、更自然。

以上试听结果是对 3 种合成康复音的初步主观评价,还有待临床试验数据的考证。

4 结束语

提出了一种基于超混沌的个性化耳鸣康复音合成新方法,从音乐谱曲角度出发,对音乐进行主旋律提取并进行变换,合成音乐。在保证音乐在整体相似不重复的前提下,试听效果更加舒缓自然。同时,由于有个性化匹配的过程,使合成的音乐更符合试听者的需求。仿真实验也验证了合成音乐的音调值及音符时值满足“1/*f*”波动,合成音乐对耳鸣模型有效。因此,耳鸣康复音合成方法对耳鸣的治疗有一定参考价值。

参考文献:

[1] Baguley D,McFerran D,Hall D. Tinnitus[J]. The Lancet,2013,382(9904):1600-1607.

[2] Jiemei C,Peiyu H,Fan P. A new method of synthesizing chaotic music for tinnitus sound therapy[C]. Digital Signal Processing (DSP),2015 IEEE International Conference on. IEEE,2015:278-282.

[3] Krick C M,Grapp M,Daneshvar-Talebi J, et al. Cortical reorganization in recent-onset tinnitus pa-

tients by the Heidelberg Model of Music Therapy [J]. Frontiers in neuroscience,2015,9:49.

[4] Hann D,Searchfield G D,Sanders M,et al. Strategies for the selection of music in the short-term management of mild tinnitus [J]. Australian and New Zealand Journal of Audiology,2008,30(2):129-140.

[5] 陈杰梅,何培宇,潘帆. 一种基于五声音阶和混沌的耳鸣康复音生成方法[J]. 生物医学工程学杂志,2015,32(6):1329-1334.

[6] 汪璐,何培宇,潘帆. 基于音乐数字设备接口技术的耳鸣康复分形声生成方法研究[J]. 生物医学工程学杂志,2014,31(4):888-893.

[7] 李重光. 基本乐理通用教材[M]. 北京:高等教育出版社,2004:159-164.

[8] Aoki E. Apparatus and method for automatically composing music according to a user-inputted theme melody:U. S. Patent 6,124,543[P]. 2000-9-26.

[9] Chen C H,Sheu L J,Chen H K,et al. A new hyper-chaotic system and its synchronization [J]. Nonlinear Analysis: Real World Applications,2009,10(4):2088-2096.

[10] Salamon J,Gómez E. Melody extraction from polyphonic music signals using pitch contour characteristics [J]. IEEE Transactions on Audio, Speech, and Language Processing,2012,20(6):1759-1770.

[11] 刘丹,张乃尧,朱汉城. 音乐特征识别的研究综述[J]. 计算机工程与应用,2002,38(24):74-77.

[12] 王兴元,王明军. 超混沌 Lorenz 系统[J]. 物理学报,2007,56(9):5136-5141.

[13] Aizawa Y,Murakami C,Kohyama T. Statistical Mechanics of Intermittent Chaos *f*-vSpectral Behaviors of the Semi-Markovian Class[J]. Progress of Theoretical Physics Supplement,1984,79:96-124.

[14] Beauvois M W. Quantifying aesthetic preference and perceived complexity for fractal melodies [J]. Music Perception,2007,24(3):247-264.

[15] Li Shangwen,He Peiyu,Chen Jiemei. An Adaptive Tinnitus Model Based on Least Mean Square Algorithm[C]. in IEEE 13th International Conference on Signal Processing Proceedings,2016:1368-1373.

A Method of Personalized Synthesizing Tinnitus
Rehabilitation Sound based on Hyper-chaos

JIN Ning-min, HE Pei-yu, PAN Fan, CHEN Jie-mei

(School of Electronic Information and Engineering, Sichuan University, Chengdu 610065, China)

Abstract: Tinnitus is mainly manifested as auditory perception without external sounds. It gets more and more attention because of its high incidence and serious impact on people’s normal life. Scientific researches show that similar but none-repetitive music can be used as tinnitus rehabilitation to relieve tinnitus symptoms. However, the current synthetic music can’t personalizedly match with the audio test and audition effect is poor. To solve the problems, this paper proposes a method of personalized synthesizing tinnitus rehabilitation sound based on hyper-chaos and main melody extraction. Results show that the tinnitus rehabilitation sound is not only self-similar but also more soothing and natural. It has a certain inhibitory effect on tinnitus in the tinnitus model. Thus, it has a referential significance for tinnitus treatment.

Keywords: signal and information processing; speech synthesis; tinnitus rehabilitation sound; personalized; hyper-chaos; main melody extraction