

文章编号: 2096-1618(2021)03-0275-05

一种具有个性化频段的舒缓耳鸣康复音 数字合成新方法

漆蕾宇, 何培宇, 陈杰梅, 蔡丽

(四川大学电子信息学院, 四川 成都 610065)

摘要:耳鸣是临床上一种没有外界声源刺激产生的主观听觉感知。研究表明相似不重复的分形音或是与耳鸣相近频段的声音能治疗或缓解耳鸣。现有合成的耳鸣音乐存在频率单一和试听效果差等问题,因此提出一种既有舒缓的分形音乐又有特定频段自然音的个性化耳鸣康复音的数字合成新方法。该方法首先基于迭代函数系统算法(iterative function system, IFS)、调内音符、和弦伴奏生成音乐的第一、二音轨,再基于迭代函数系统算法、线性比例重叠相加法、含有特定频段的自然音片段生成音乐的第三音轨。实验结果表明,此方法合成的音乐符合分形音乐的基本特征,试听效果舒缓,在模拟耳鸣模型中也能对同频段的耳鸣声起到抑制作用,对耳鸣治疗具有一定的参考价值。

关键词:线性比例重叠相加法;音频拼接;迭代函数系统;耳鸣康复音

中图分类号: TN912.3

文献标志码: A

doi: 10.16836/j.cnki.jcui.2021.03.006

0 引言

耳鸣能在脑内或者颅内感知到非外界的声音,从而影响到患者的日常生活,使患者焦虑、烦躁、重则抑郁等^[1-3]。因此,探寻耳鸣康复方法具有重要意义。

研究表明具有相似不重复特性的分形音乐^[4]或是与患者耳鸣频率相近的声音能满足耳鸣治疗音的特点^[5]。目前国内的相关研究^[6-8]合成的音乐虽然能满足相似不重复的特点,但由于都是音符序列产生的音乐使音频频率单一,不能与患者进行适配,而且还存在着试听效果不佳等问题。除此之外,国外海德堡神经音乐疗法^[9-10]提出了一种基于患者的耳鸣频率与音乐结合的方法用于治疗耳鸣,并已取得一定成效,这种方法是治疗师将与耳鸣匹配的正弦波信号产生的声音添加到背景音乐中但没有考虑患者对该声音的接受度。

针对上述问题,本文将利用各类算法从音乐的曲调、节奏以及和声三个方面进行个性化合成,提出一种个性化频段的舒缓耳鸣康复音数字合成方法。此方法先将乐器数字接口(musical instrument digital interface, MIDI)的中音区分为24个常用的大小调式音符组合,选取一个调式音符组合后,利用IFS算法产生的分形图像的层次特性将组合中的各个音符映射为主音轨音符序列,再利用动态规划算法为每个特征音符搭配调式内和弦,生成两轨分形音乐。音符的多样性以及伴奏和弦的加入使得分形音乐更加舒缓。然后再利用自然音,自然音不仅是多数耳鸣患者的掩蔽声源,也是一种可以调节消极情绪使人放松的声音^[11]。先将自然音进行多样化处理产生与患者耳鸣频段匹配的自然音片段,再将各个

片段用IFS算法产生的分形序列进行映射得到连续音频,然后将其作为第三音轨加入到分形音乐中。

实验结果分析表明,本文方法产生的康复音既具有分形音乐的相似不重复特性且试听效果更加舒缓,也含有特定频率特性,使得合成的康复音能与患者个性化适配,达到耳鸣康复音的效果,并在耳鸣模型中对特定频段的耳鸣有抑制作用。

1 耳鸣康复音合成方法

音乐的主要特征要素是曲调、节奏、和弦、力度、速度^[12]。从音乐的调式、节奏以及和弦三个方面进行个性化合成,在形成旋律时,为满足分形音乐的相似不重复特性,采用IFS算法产生分形序列映射音符的组成,并最终利用MIDI技术生成康复音的第一、第二音轨。

本文方法的第三音轨是利用含有耳鸣频段的自然音频生成的。先将自然音片段从音色、音调、响度和速度方面进行单独或者联合的变换处理。处理后的音频片段通过拼接形成第三音轨,但如果直接将音频拼接,会在片段连接处出现声音突变的问题。所以利用线性比例重叠相加算法拼接音频,减少声音突变的问题。

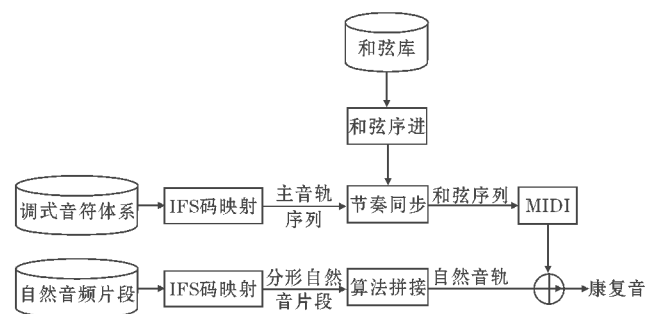


图1 耳鸣康复音合成过程

收稿日期: 2020-12-15

基金项目: 四川省科技支撑资助项目(2011SZ0123、2013GZ1043)

2 分形音乐合成

2.1 IFS 算法生成分形序列

分形是一类组成部分与整体以某种方式相似的对象^[6]。IFS 算法在分形图像处理领域有着广泛的应用,分形图像的实质是将与整体相似的图像经过仿射变换后拼接而成,仿射变换的数学表达式为

$$\omega: \begin{cases} x' = ax + by + e \\ y' = cx + dy + f \end{cases} \quad (1)$$

其中, ω 表示仿射变换, a, b, c, d, e, f 为仿射变换系数。不同的仿射变换系数与该仿射变换被调用的不同概率会产生多变的分形图形,分形图形中各个点的位置在可控范围内具有随机性。所以利用图形的 x, y 变量序列进行取整和线性变换后会得到整体相似局部不重复的序列值,利用变换后的序列值来映射 MIDI 音符的生成可以得到相似不重复的音符序列值。

2.2 动态规划算法编配和弦

和弦编配过程是一个不断寻找最佳路径的过程,本文编配和弦采用基于动态规划的和弦编配算法^[13]。对每一阶段的和弦选择要使得全局目标函数为最优,全局最优目标函数定义为

$$V_{k,n}^* = \text{OPT}_{U_k, U_{k+1}, L, U_n} \sum_{k=1}^n V_k(S_k, U_k) \quad (2)$$

其中, S_k 定义为第 k 阶段的所有状态变量的总和, U_k 为第 k 阶段的决策变量。从起始点开始逐步计算最优函数,同时记录前驱路径,到达终点后回溯出最优的路径,就可得到每一阶段最优匹配的和弦。

2.3 数字合成过程

将中音区音符分为不同调式组合,用调式组合内的音符映射成主音轨旋律,避免了音调变换不和谐现象。生成主音轨音符时,先用 IFS 算法产生分形图,记录坐标值形成分形序列,再将该序列进行线性变化,然后将处理后的序列映射为选择调式的音符值,形成了第一音轨的音符值序列。

$$y(n) = \begin{cases} y(n) & n \leq (i-1)L \\ y(n)\omega_2 + y_i(m)\omega_1 & (i-1)\Delta L + 1 \leq n \leq (i-1)\Delta L + M, 1 \leq m \leq M \\ y_i(m) & (i-1)\Delta L + M + 1 \leq n \leq (i-1)\Delta L + L, M + 1 \leq m \leq L \end{cases} \quad (3)$$

式中两个斜三角的窗函数为式4,其中 M 为重叠部分长度。

$$\begin{aligned} \omega_1(n) &= (n-1)/M \\ \omega_2(n) &= (M-n)/M \\ n &= 1, 2, \dots, M \end{aligned} \quad (4)$$

采用文献[15]的方法产生自然音频库。自然音轨根据耳鸣患者的耳鸣匹配,选择合适的自然音类型,如蝉鸣、水声、风声等。然后将自然音片段分别进行响

度和弦伴奏选择常用的调式内正副三和弦,按照三和弦进行的基本方向,用动态规划算法在全局最优的情况下为每小节匹配适宜的三和弦。分解和弦是指各个组成音按照根音、三音、五音等顺序依次发声,常用于抒情、委婉的音乐。本文为合成舒缓分形音乐选用分解和弦。

本文的节奏序列通过先确定合成音乐的拍号,再将该拍号下可能出现的每小节的节奏组合进行编号处理,然后将 IFS 序列进行线性变换,最后将小节节奏按照 IFS 序列进行映射排列即构成了音符节奏序列即音符时值序列。将音符值与对应持续时值确定后,利用 MIDI 技术实现分形康复音的合成。

3 个性化频段音轨合成过程

个性化频段音轨根据不同患者的耳鸣频率情况利用自然音片段分别进行响度、速度、音调、音色变换后拼接形成。在语音合成算法中常使用数据叠接的方法将一帧帧数据连接成连续、平滑的数据流。常见的方法有重叠相加法、重叠存储法和线性比例重叠相加法^[14]。其中线性比例重叠相加法常用于前后两帧变化较大或者不确定是否会有较大变化的情况。自然音片段的多样性使得在拼接处不确定前后的两段音频是否会有较大的变化,所以在片段拼接处采用线性比例重叠相加法。

线性比例重叠相加法是在重叠相加法的基础上进行了改进。重叠相加法中是将两帧中的重叠部分直接进行相加,线性比例重叠相加法中是将两帧的重叠部分先分别通过两个斜三角的窗函数,然后再相加。假设设置帧长为 L ,帧移为 ΔL ,则两帧中的重叠部分长为 $M = L - \Delta L$,假设 $y_i(m)$ 为要加入到 $y(n)$ 中的一帧, $y_i(m)$ 与其前一帧 $y_{i-1}(m)$ 有重叠部分,两帧的重叠部分在 $y_i(m)$ 中对应于 $1:M$ 的位置,而此时前一帧 $y_{i-1}(m)$ 已经加入到 $y(n)$ 中,所以 $y_i(m)$ 与 $y(n)$ 的重叠部分对应为 $y(n)$ 中的 $((i-1)\Delta L + 1) : ((i-1)\Delta L + M)$,则利用线性比例重叠相加法后为

度、速度、音调、音色变换,并将变换后的各个片段进行编号。为使在耳鸣同频率的声音刺激之余,音乐对患者能起到舒缓的作用,将这些自然音片段用 IFS 算法产生的分形序列作为映射序列来控制分形自然音的顺序合成。然后将各个自然音片段利用线性比例重叠相加法进行拼接合成第三音轨。

4 实验结果与分析

4.1 音频分析

合成方法对治疗音增加了个性化频段特性。为验证此特性,将相同分形序列映射产生的分形音乐与本文方法产生的频段集中在 2000 ~ 3000 Hz 和 3000 ~ 4000 Hz 的音乐进行频域图比较,如图 2 所示。

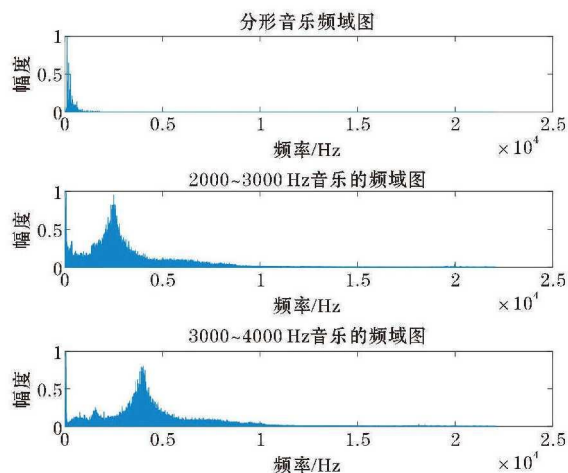


图2 分形音乐与合成音乐的频域图对比

实验结果表明分形音乐的频段主要集中在低频,本文方法产生的音乐可以根据耳鸣患者的个性差异产生不同频段的音乐用于治疗。既能与患者进行个性化适配也能增加音频的频率多样性。

为验证线性比例重叠相加法对音频拼接的连续性,将两段拼接处相差较大的水声音频进行试验。在线性比例重叠相加法拼接时将前段音频的后端 2 倍采样率的点和后段音频的前端 2 倍采样率的点作为重叠部分进行拼接。直接拼接和利用线性比例重叠相加法拼接后音频的时域图和时频图如图 3、4 所示。

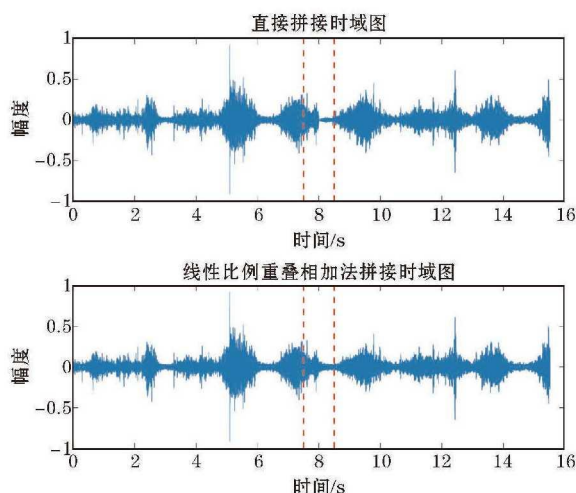


图3 直接拼接与线性比例相加法拼接的时域图对比

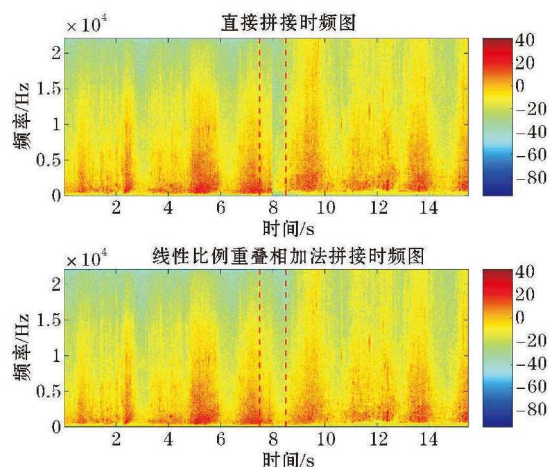


图4 直接拼接与线性比例相加法拼接的时频图对比

由图 3、4 可知当拼接处前后的变换较大时直接拼接会在拼接点数据不连续,且在时频图中也会出现明显的不连续间隔,会造成拼接处声音的突变。而利用线性比例重叠相加法后会在拼接处过渡顺滑,且没有明显间隔。实验结果表明线性比例重叠相加算法能使拼接处连续、平滑。

4.2 波动分析

利用 $1/f$ 波动来验证合成的治疗音的相似不重复特性。 $1/f$ 波动是一种能让人感到舒适的波动,且有该特性的音乐也满足相似不重复性^[16-17]。大部分让人舒缓的音乐都满足 $1/f$ 波动规律,所以要产生舒适的耳鸣音乐也需要满足 $1/f$ 波动规律。该波动在功率谱中表现为波动的功率谱密度与频率的负次幂成正比,利用 $S(f)$ 表示信号在频率 f 处的功率谱密度,其与频率的关系在式子两边同时取对数后可以表示为

$$\lg(S(f)) \propto -\beta \lg(f) \quad (5)$$

研究表明当指数 β 在 $[0.5, 1.5]$, 可以看作是 $1/f$ 波动,并且该波动会给人舒适的感觉。对主音轨,和弦音轨以及音符时值分别进行 $1/f$ 波动分析,由图 5 可知, $\beta=0.9315, \beta=0.9317, \beta=0.7739$ 均在 $[0.5, 1.5]$, 符合 $1/f$ 波动特性。

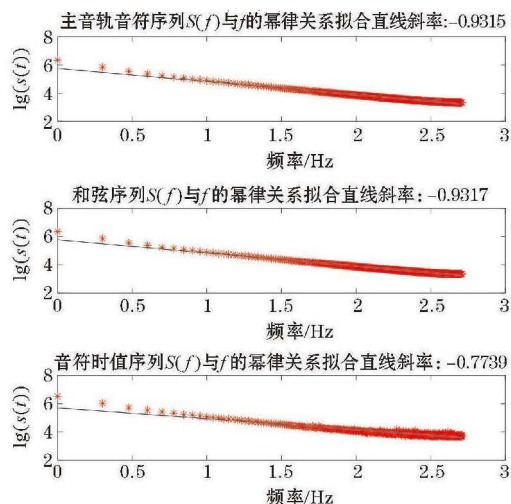


图5 各音轨的功率谱密度拟合曲线

该实验结果说明,本文的处理过程以及合成方法得到的音乐符合 $1/f$ 波动特性,满足相似不重复特性,并会给人舒适的感觉。

4.3 耳鸣模型分析

为验证合成的音乐是否能对特定频段的耳鸣有抑制作用,参考文献[18]提出的模型。该耳鸣模型将人耳听觉范围均分成 20 个频率范围不同的分路。当耳蜗中该通路功能受损后,为了将进入听觉中枢的信号强度维持在某一稳定值,该通路会自适应的放大信号,而通路的自发信号也会被放大从而形成耳鸣。当外界输入有效的治疗音后,耳鸣的声音将被抑制。可以通过模型输出的时频图观察到自发信号被抑制。

本实验假设 11 通路(中心频率为 2609 Hz)受损。图中前 10 s 为模型有耳鸣时的原始输出情况,在 10 s 时输入 60 s 的频率集中在 2000 ~ 3000 Hz 的治疗音信号作为输入刺激信号,在第 70 s 时撤掉治疗音。由图 6 可以观察到在治疗音加入的 60 s 中耳鸣被抑制,表现为模型输出功率减小,而治疗音撤掉后,模型输出功率增大,表示耳鸣复发情况。

频段匹配的音乐刺激效果分析

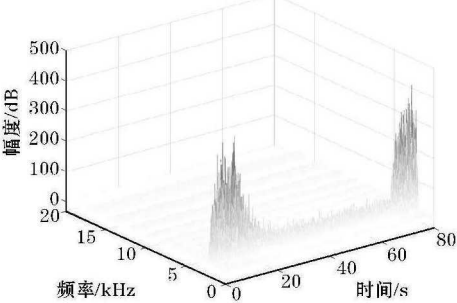


图 6 匹配频段音乐刺激时模型输出

当将该频段音乐加入第 7 通路(中心频率为 5568 Hz)的受损通道时,模型输出如图 7 所示,对耳鸣情况没有抑制效果。实验说明本文合成的康复音可以有效地抑制对应频段的耳鸣。

频段匹配的音乐刺激效果分析

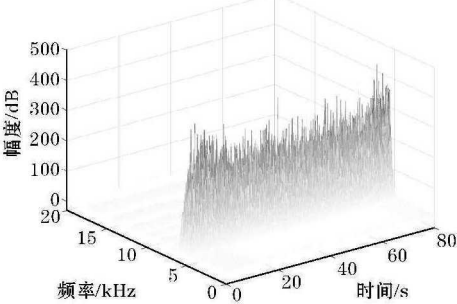


图 7 频段不匹配的音乐刺激时模型输出

4.4 试听分析

以上 3 种验证方式都是对音乐的客观性能进行验证,音乐治疗耳鸣不但要符合耳鸣治疗的特点,更多的

是提高耳鸣患者对音乐的接受度。所以本文采用试听实验对生成的音乐进行主观评价。将文献[6]和本文所产生的音乐进行对比实验。文献[6]也采用 IFS 分形序列生成音乐,将生成的音乐标记为 Frag_1,本文的分形音标记为 Frag_2,本文方法产生的分形音乐与自然音乐合成的音乐标记为 Frag_3,每种音乐截取 3 分钟供试听。该实验在普通人群中招募 32 名志愿者进行试听。结果如表 1 所示,听完音乐后试听人员先从音频的重复度来选择,从如下 3 个标准中选择一个所听音频最符合的特征,把选择的人数百分比统计得到 1 ~ 3 行的结果,然后分别从流畅度、舒缓度、自然度 3 项标准在 0 ~ 10 的范围内进行打分,分数越高表示该特征越强。将所有评分取平均值得到如表 1 所示的 4 ~ 6 行。

表 1 试听结果

音乐片段	Frag_1	Frag_2	Frag_3
完全不重复	15.6%	6.25%	18.8%
相似不重复	68.8%	87.5%	65.6%
重复	15.6%	6.25%	15.6%
舒缓度	5.2	6.9	7.0
流畅度	5.6	7.3	7.3
自然度	5.4	6.7	7.0

由统计结果得知,在重复度这个标准中,3 个片段中选择相似不重复特征的人数百分比为 68.8%, 87.5%, 65.5%, 表明试听人员大部分都认为这 3 个片段满足相似不重复的特征,这是由于这 3 种音乐都是基于 IFS 分形算法产生的,所以结果都满足分形音乐的相似不重复特征。在舒缓度、流畅度、自然度中, Frag_2 和 Frag_3 都有明显的提高。而 Frag_3 因为加入了自然音片段所以音频在自然度中有了一定的提高。实验结果说明,本文合成的治疗音符合分形音乐的特点,且试听效果舒缓,更有益于患者在音乐治疗过程中放松心情,得到舒缓。

5 结束语

提出一种具有个性化频段的舒缓耳鸣康复音数字合成新方法。从曲调、节奏、和弦方面合成分形音乐并加入了与耳鸣同频段的自然音,使得合成的音乐既满足相似不重复特性,符合 $1/f$ 波动,也能与耳鸣频段相适配,在耳鸣模型中抑制特定频段的耳鸣,同时在试听实验中舒缓度明显增加。因此,本文的合成方法对耳鸣治疗有一定的借鉴价值。

参考文献:

[1] 卢兢哲,钟萍,郑芸. 欧洲多学科耳鸣指南:诊断、评估和治疗[J]. 听力学及言语疾病杂志, 2020,28(1):110-114.

- [2] 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.
- [3] 郭振平,杨东,李欣.耳鸣患者心理健康状况及影响因素[J].心理月刊,2020,15(13):104.
- [4] Sweetow R W, Sabes J H. Effects of acoustical stimuli delivered through hearing aids on tinnitus [J]. J Am Acad Audiol, 2010, 21(7):461.
- [5] Mohanty P R, Walikar B N, Rashinkar S M, et al. Outcome of audiometric masking therapy on subjective tinnitus patients in different age groups[J]. International Journal of Otorhinolaryngology and Head and Neck Surgery, 2017, 3(3):699.
- [6] 徐媛媛,何培宇,陈杰梅.一种基于IFS分形算法和分解和弦的耳鸣康复音合成新方法[J].四川大学学报(自然科学版),2017,54(3):517-522.
- [7] 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.
- [8] 程冉,何培宇,陈杰梅.一种基于LS文法的耳鸣康复音合成新方法[C].中国高科技产业化研究会智能信息处理产业化分会.第九届全国信号和智能信息处理与应用学术会议专刊.中国高科技产业化研究会智能信息处理产业化分会:中国高科技产业化研究会,2015:39-44.
- [9] Heike Argstatter, Miriam Grapp, Elisabeth Hut-ter, et al. The effectiveness of neuro-music therapy according to the Heidelberg model compared to a single session of educational counseling as treatment for tinnitus: A controlled trial[J]. Journal of Psychosomatic Research, 2015, 78(3).
- [10] Argstatter H, Grapp M, Plinkert P K, et al. "Heidelberg Neuro-Music Therapy" for chronic-tonal tinnitus-treatment outline and psychometric evaluation [J]. International Tinnitus Journal, 2012, 17(1).
- [11] Henry J A, Schechter M A, Zaugg T L, et al. Clinical trial to compare tinnitus masking and tinnitus retraining therapy [J]. Acta Oto-Laryngologica, 2006, 126(S556):64-69.
- [12] 李重光.音乐理论基础[M].北京:音乐出版社,1962:39.
- [13] 邓阳,周莉,许多,等.基于动态规划的自适应和弦编配算法研究[J].复旦学报(自然科学版),2019,58(3):393-400.
- [14] 宋知用. MATLAB在语音信号分析与合成中的应用[M].北京:北京航空航天大学出版社,2013:293-296.
- [15] 蔡丽,何培宇,陈杰梅.一种基于迭代函数系统的个性化耳鸣康复自然音合成方法[J].生物医学工程学杂志,2018,35(4):631-636.
- [16] Beauvois M W. Quantifying aesthetic preference and perceived complexity for fractal melodies [J]. Music Perception, 2007, 24(3):247-264.
- [17] Colley I D, Dean R T. Origins of 1/f noise in human music performance from short-range autocorrelations related to rhythmic structures [J]. PLOS ONE, 2019, 14(5).
- [18] 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 Soothing Tinnitus Rehabilitation Sound with Personalized Frequency Band

QI Leiyu, HE Peiyu, CHEN Jiemei, CAI Li

(College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China)

Abstract: Tinnitus is a subjective auditory perception without external source of sound in clinical. Researches show that fractal sounds which are similar but unrepeatable or sounds which are similar to the frequency of tinnitus can therapy or relieve tinnitus. The current synthetic tinnitus music has the matter of single frequency and poor audition effect, so a new digital synthesis method of personalized tinnitus rehabilitation sound which has both soothing fractal music and specific frequency natural sound is proposed. This method is based on iterative function system algorithm (iterative function system, IFS), key notes, chord accompaniment to generate the first and second tracks of music, and then it is based on iterative function system algorithm, linear proportional overlap and addition, the natural sound segment containing specific frequency bands generates a third soundtrack of music. The results show that the sound synthesized by this method is consistent with the basic features of fractal sound and the audition effect is soothing. In the simulation tinnitus model, it can also inhibit the tinnitus of corresponding frequency, which may have reference value for tinnitus treatment.

Keywords: linear proportion overlap-adding algorithm; audio concatenation; IFS; rehabilitation sound