



智能音乐学与 中国音乐数字媒体论

Intelligence musicology and
the digital media of Chinese music

陈根方 著

文化艺术出版社
Culture and Art Publishing House



智能音乐学与中国音乐数字媒体论

Intelligence musicology and
the digital media of Chinese music

陈根方 著

文化艺术出版社
Culture and Art Publishing House

图书在版编目 (C I P) 数据

智能音乐学与中国音乐数字媒体论 / 陈根方著.

—北京: 文化艺术出版社, 2018.3

ISBN 978-7-5039-6470-1

I. ①智… II. ①陈… III. ①数字音频技术—应用—
音乐—研究 IV. ①J6-39

中国版本图书馆CIP数据核字(2018)第045768号

智能音乐学与中国音乐数字媒体论

著 者 陈根方

责任编辑 王 红 李 冬

书籍设计 马夕雯

出版发行 文化艺术出版社

地 址 北京市东城区东四八条52号 100700

网 址 www.caaph.com

电子邮箱 s@caaph.com

电 话 (010) 84057666 (总编室) 84057667 (办公室)
84057696—84057699 (发行部)

传 真 (010) 84057660 (总编室) 84057670 (办公室)
84057690 (发行部)

经 销 新华书店

印 刷 国英印务有限公司

版 次 2018年4月第1版

印 次 2018年4月第1次印刷

开 本 700毫米×1000毫米 1/16

印 张 20

字 数 320千字

书 号 ISBN 978-7-5039-6470-1

定 价 50.00 元

国家社科基金后期资助项目 出版说明

后期资助项目是国家社科基金设立的一类重要项目，旨在鼓励广大社科研究者潜心治学，支持基础研究多出优秀成果。它是经过严格评审，从接近完成的科研成果中遴选立项的。为扩大后期资助项目的影响，更好地推动学术发展，促进成果转化，全国哲学社会科学规划办公室按照“统一设计、统一标识、统一版式、形成系列”的总体要求，组织出版国家社科基金后期资助项目成果。

全国哲学社会科学规划办公室

目 录

第一章 音乐媒体与智能音乐学	(1)
第一节 音乐媒体进化	(1)
第二节 智能音乐学概述	(10)
第三节 中国传统音乐媒体与智能处理	(16)
第二章 工尺谱音乐媒体	(25)
第一节 工尺谱概述	(25)
第二节 工尺谱的元数据方案	(30)
第三节 工尺谱可编辑语义文本表示与工尺谱语料库	(42)
第四节 类 MIDI 的工尺谱存储方案	(49)
第三章 古琴音乐媒体	(65)
第一节 古琴记谱法	(66)
第二节 古琴可编辑语义文本乐谱	(72)
第三节 基于 SVG 的古琴 Web 乐谱重建	(85)
第四章 工尺谱与古琴乐谱图像分割	(95)
第一节 图像分割概述	(95)
第二节 基于 DFS 的图像标记分割算法	(98)
第三节 基于包围盒算法的图像分割	(121)
第四节 基于自适应迭代投影平滑的图像分割	(129)
第五节 基于聚类分析的图像分割	(138)

第五章 光学乐谱识别	(142)
第一节 乐谱识别概述	(142)
第二节 模式特征提取	(145)
第三节 工尺谱乐谱识别	(152)
第四节 古琴谱字识别	(174)
第五节 乐谱图像匹配	(179)
第六章 数字音频识别	(188)
第一节 音频识别概述	(188)
第二节 民族乐器自动分类	(193)
第三节 中国戏曲唱腔自动分类	(196)
第四节 音乐智能检索	(202)
第五节 古琴的演奏音频分割	(203)
第六节 音高和节奏识别	(210)
第七节 情感识别研究概述	(217)
第七章 中国律制建模分析与古琴音色建模	(220)
第一节 中国律制建模分析	(220)
第二节 古琴音色物理建模	(239)
第八章 人工智能与数字媒体技术在音乐领域的其他应用	(255)
第一节 算法作曲	(256)
第二节 音乐的可视化	(266)
第三节 手势识别与虚拟乐器演奏模型	(269)
第四节 古琴打谱的智能分析	(275)
第五节 音乐保护	(280)
参考文献	(291)

第一章 音乐媒体与智能音乐学

智能音乐学是研究信息时代人工智能、音乐媒体与音乐艺术融合的理论、方法和技术的新兴学科，是现代信息技术与传统艺术的交叉研究领域，它涉及机器学习、机器感知、神经网络、遗传算法、知识表示与知识系统等人工智能范畴，以及文本、图像、音频、视频、动画、虚拟现实、人机交互等数字媒体范畴。图 1-1 为与智能音乐学相关的交叉学科。

智能音乐学的研究对象包括音乐历史、音乐理论、音乐人、乐曲、音频、记谱法、乐谱、乐器、数字乐器、音乐数字格式、数字音色、音乐描述语言、音乐语料库、乐谱识别、音频识别、算法作曲、音乐软件、音乐交互、音乐挖掘、音乐机器人、演奏姿势、音乐情感分析等与音乐艺术相关的内容。

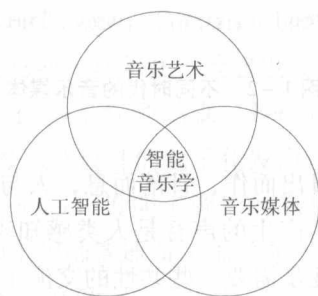


图 1-1 智能音乐学的交叉学科

第一节 音乐媒体进化

音乐与人类一样，历史久远，承载音乐的媒体进化也呈现多元化，按

音乐媒体的进化历程可以把音乐的历史划分为原始时代、符号时代、电子时代、数字时代和虚拟时代，每个时代的时间跨度不同，承载音乐的媒体也不同，见图 1-2。

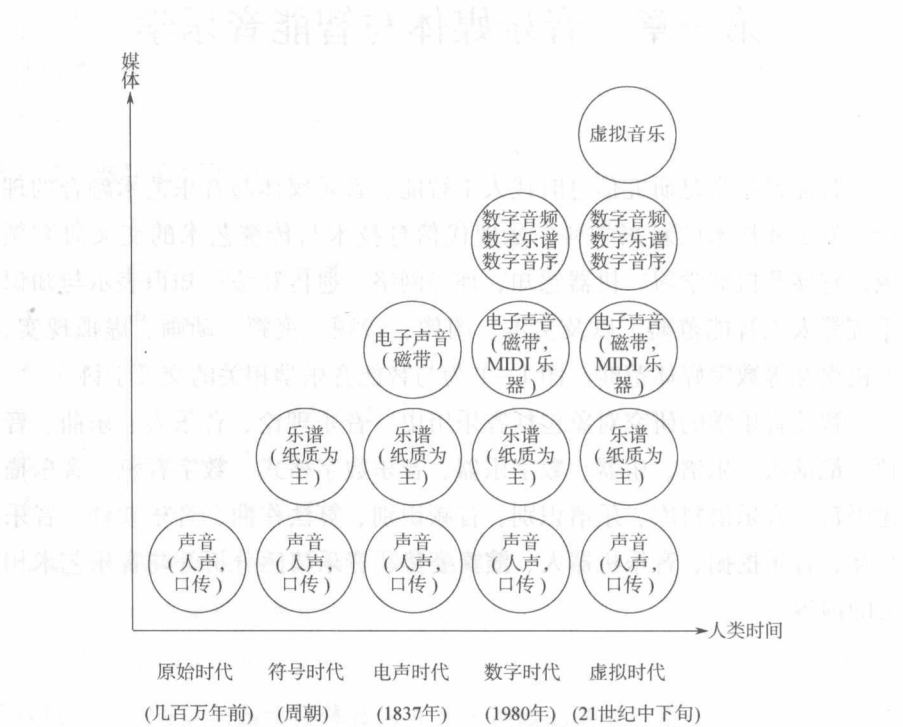


图 1-2 不同时代的音乐媒体

在远古时代，人们日出而作，日落而息，人与自然完全融合，大自然的声音和人们生产、生活产生的声音是人类感知声音的源泉。为了生存，人类以群居方式生活，逐步需要一些共性的交流工具，于是部落内的人们相互之间发出相同的声音频率，赋予不同的频率以不同的含义，语言在生活、生产中自然地产生。同时，利用工具敲击发出的声音也被赋予了不同的意义，有节律地敲击声音为音乐的节奏和节拍产生创造了条件。

在音乐符号产生前的原始时代的漫长岁月中，音乐的承载需要人们之间的口传身教，声音是音乐传承的主要载体，部落活动（祈祷、祭祀等）为主要的音乐传播和传承方式，有些部落有相对固定的巫师随着车轮等交

通工具的发明和使用,人们的活动范围得到了拓展,逐步形成了国家和城邦,人类活动范围的相对固定与交流的频繁进行,为音乐的传承与地域传播提供了可能。一个稳定的环境是音乐传承的重要条件,因此音乐在统治阶层中较为流行,如在中国周朝就设有“大司乐”^①等音乐机构,负责音乐教育与演出。

在音乐实践中,口口相传的传承方式受到了挑战,一些音乐人自然地利用符号来作为音乐传承的备忘录,随着时间的推移,这些音乐的记录方式,逐步以记谱法的形式固定下来,音乐得以以乐谱的方式进行传承。安妮·德拉夫孔·基尔默(Anne Draffkorn Kilmer)曾发表文章论证在美索不达米亚的楔形文字板上发现了最早的记谱法^②。中国存见最早的乐谱为中国周代的《礼记·投壶》中鼓谱的若干个记谱符号^③,而最早的完整乐谱则为中国南北朝时期丘明(494~590)的古琴琴曲《碣石调·幽兰》^④。音乐的符号时代以记谱法的方式传承音乐,中国古代流行的记谱法有工尺谱记谱法、减字谱记谱法,现代有五线谱记谱法、简谱记谱法等,直到现在乐谱仍然是音乐的主要载体之一。

到了19世纪,电子技术得到了迅猛发展,电子音乐进入萌芽时期。1837年,美国工程师查尔斯·格拉夫顿·佩奇(C. G. Page)首次让电子设备发出了声音;1857年,法国发明家斯科特(Scott)发明了声波振记器;1876年,美国工程师艾利沙·格雷(Elisha Grey)发明了音乐电报^⑤;1877年,美国发明家托马斯·阿尔瓦·爱迪生(Edison)发明了世界上第一台留声机。1885年,德国工程师珀森(Person)和埃姆斯特·路伦茨(Ernst Lorenz)发明了按乐器设计的电子化发声设备;1887年,德国人埃米乐·贝林纳(Emil Berliner)发明了圆形的唱片和水平式留声机,并获得专利。1897年,美国工程师赛地斯·加希尔(Thaddeus Cahill)设计了第一台电动机机械乐器;1898年,丹麦科学家瓦德马尔·波尔森(V. Poulsen)利用剩磁原理发明了磁性录音机。

① 杨荫浏:《中国古代音乐史稿》,北京,人民音乐出版社,2004年版,第34~35页。

② Anne Draffkorn Kilmer, 1998: "The Musical Instruments from Ur and Ancient Mesopotamian Music", *Expedition Magazine*, July.

③ 王耀华:《中国传统音乐乐谱学》,福州,福建教育出版社,2006年版,第3页。

④ 章华英:《古琴》,杭州,浙江人民出版社,2005年版,第71页。

⑤ Elisha Gray, http://en.wikipedia.org/wiki/Elisha_Gray.

进入 20 世纪,电子音乐技术得到全面拓展,并直接促进了数字音乐的萌芽。1931 年,美国无线电公司(RCA)试制成功 $33\frac{1}{3}$ 转/分的密纹唱片,密纹唱片为每分钟 33.5 转,比以前转速为每分钟 78 转的唱片增加了音乐播放时间。1948 年,位于美国的哥伦比亚公司研制出了 $33\frac{1}{3}$ (转/分)密纹唱片的新技术,此技术为大批量生产提供了技术保障,是唱片发展史上的里程碑。1935 年,德国的通用电气公司研制出了磁带录音机,在第二次世界大战中被广泛应用于广播和军事。1936 年,法国工程师奥斯卡·萨拉(Oskar Sala)设计出第一个分谐波的合成器。1940 年,美国工程师荷马·杜德利(Homer Dudley)第一次合成产生人声。

1950 年,美国发明了第一台专用演奏音乐的电脑 MK1。1963 年,美国唐纳德·布奇拉(Donald Buchla)发明了第一个现代合成器 Model 100。1963 年,音频盒式磁带由荷兰飞利浦公司进行生产,渐渐地唱片开始走出人们的视线。1970 年,美国罗伯特·穆格(Robert Moog)发明第一个带键盘的单音合成器 Minimoog。1976 年,第一个可编程的合成器 PPG 1003 Sonic Carrier 被发明。1978 年,第一个波表合成器 PPG Wave-computer 360 出现。1982 年,第一个 MIDI 合成器 Sequential Circuits Prophet 600 出现。1983 年,第一个数字化 FM 合成器 DX7 由雅马哈(Yamaha)生产。1996 年 8 月,飞利浦、索尼、东芝、松下等公司联合制定了新一代高密度光盘 DVD 规格书。1996 年 11 月,松下公司在全球上市了 DVD-Video 播放机。

电子音乐时期的主要音乐媒体包括声音的存储设备、声音的合成设备、声音的传播设备、声音的播放设备、声音的放大、均衡、滤波等设备以及各种电子音乐格式,特别是乐器的电子信号系统——MIDI 协议的确立为数字音乐的产生创造了条件,其中声音的保存与重现技术是音乐事业得以迅猛发展的关键技术。电子时代的音乐技术是围绕着模拟信号展开的,以重现逼真的声音为终极目标,发展出高超的音箱制作技术和各种音乐保存媒体,如唱片、磁带、CD、VCD、DVD 等。电子乐器则包括音箱、键盘乐器、打击乐器、弦乐器、木管乐器、调音台等控制设备和系统互联设备。

1966 年,美国斯坦福大学的约翰·乔宁(John Chowning)提出了频

率调制 (Frequency Modulation, 简称 FM) 声音合成技术, 它用特定的算法来模拟真实乐器的声音, 由于设计简单、生产成本低, FM 技术得到了广泛应用^①。但是 FM 技术合成的声音和真实的乐器声音之间有一定的失真度。1979 年, Palm Products GmbH (PPG) 公司的沃尔夫冈·帕姆 (Wolfgang Palm) 提出了波表 (Wave Table) 合成技术, 首先采样真实的乐器的声音, 然后再利用相应算法对声音进行强弱、音高调整, 得到实际需要的强度和音高, 波表合成技术产生的声音相对比较逼真, 而且对处理器的要求不是很高, 因此被广泛接受与使用。1996 年, 德国 Steinberg 公司提出 VST (Virtual Studio Technology) 开发标准^②, 第一次解决了软件合成器的延迟问题, 解决了处理器的处理能力瓶颈, 推动了音色采样技术的迅猛发展。

1947 年, 国际化标准组织 ISO (International Organization for Standardization) 成立, 致力于各类国际行业标准的制定、管理、更新和发布; 1948 年, 克劳德·艾尔伍德·香农 (Claude Elwood Shannon) 提出了信息熵的概念, 为信息论和数字通信 (包括信息编码) 奠定了理论基础; 1967 年, ASCII (美国信息交换标准代码) 成为信息交换的基本编码, 为计算机进入各个行业奠定了基础; 1980 年, 《信息交换用汉字编码字符集》GB 2312—1980 由中国国家标准总局发布, 为计算机处理和使用汉字打下了基石。1981 年 11 月, 美国工程师戴维·史密斯 (Dave Smith) 向纽约的音响工程协会 (Audio Engineering Society) 提出了数字乐器标准; 1983 年, 他和乐器厂商合作, 发布了乐器数字接口协议 MIDI 1.0 (Musical Instrument Digital Interface), 并由 MIDI Manufacturers Association (MMA) 进行维护和更新^③。MIDI 技术的出现, 使得以计算机为中枢系统的计算机音乐系统形成, 各类计算机软件如雨后春笋般出现, 大量免费资料可以通过互联网获取, 音乐创作、编曲、制作彻底进入千家万户, 数字音乐时代到来了。

在数字音频编码方面, 出现了大量的音频编码方案。1939 年, 美国

① John Chowning, http://en.wikipedia.org/wiki/John_Chowning.

② VST/VSTi, <http://www.logicmusic.com/2016/04/20/history-of-steinberg-pioneering-company-celebrates-30th-anniversary-2/>.

③ 日本合成器编程协会 (JSPA): 《MIDI 检定 3 级公式指南手册》, 北京, 中音公司, 2003 年版。

贝尔实验室的物理研究学家荷马·杜德利 (Homer Dudley) 发明了声码器; 1991 年, IBM 公司和微软公司联合发布了数字音频 WAV 格式 (Waveform Audio File Format)^①, 并成功应用于 OS2 和 Windows 系统中。同年, MP3 (MPEG Group Audio Layer III) 由德国埃尔朗根的研究组织 Fraunhofer Gesellschaft 的工程师发明, 并由国际标准组织和国际电工委员会 ISO/IEC 进行标准化处理, MP3 逐步变成了目前最流行的音频编码方案之一。其他音频编码方案还有 AVI、RIFF 等。数字音源格式大量产生, 如 Giga 格式、SoundFont 格式、AKAI 格式、ROLAND 格式、Acid 格式、Kontakt 格式、VSTi 格式、DXi 格式等, 其中使用最广泛的是 Giga 格式的音源, 许多著名的音色厂家每年都会推出大量制作精良的 Giga 音源盘, 可供选择的种类非常丰富, 如 Hypersonic, Emulator Standard, Akai Phat Phunky, East West Drum 2, NemeSys vo15 1G Grand, Sonic Implants Guitar & Bass, Yellow Tools Pure E-Basses, 维也纳失真吉他、维也纳交响乐音色盘、Dan Dean 管乐、亚洲之心等。

进入 21 世纪, 互联网得到普及, 世界进入信息时代, 信息的产生、存储、传播和获取变得非常方便。音乐创作、存储、传播和下载获取以互联网为平台展开, 基于 Web 的音乐描述语言是用于表示网络音乐的符号体系, 它通常包含基本元素的有限集合和组合规则的有限集合, 以可扩展标记语言 XML 为基础的网络音乐描述语言有 MidiXML, Apple GarageBand Format, SMDL, MPEG4-SMR, MNML-The Musical Notation Markup Language, MusicXML, MHTML, MML: Music Markup Language, ScoreML, JscoreML, eXtensible Score Language (XScore), Xmusic, Music Description Language (MDL) 等^②。

同时, 各个音乐组织和社会团体提出了表示数据的音乐元数据方案, 如美国印第安纳州立大学的数字音乐图书馆工程 (Variations 3 项目)、美国内华达大学的“国际音乐元数据”、德国教育研究部的“视音频元数据自动生成研究项目”、音频工程协会 (Audio Engineering Society, AES) 的“核心音频元数据方案”、美国国会图书馆的视音频元数据原型计划、北欧

① WAV, <https://en.wikipedia.org/wiki/WAV>.

② Gerd Castan, *Datenformate of Musiknotation*, <http://www.music-notation.info/en/compmus/notationformats.html>.

视音频元数据研究小组 (Scandinavian Audiovisual Metadata Group, 简称 SAMG) 的“SAM 视音频元数据”、网络音乐数据库 MusicBrainz 的音乐元数据计划 (MusicBrainz Metadata Initiative, 简称 MMI)、国际多媒体数字图书馆工程 (Multimedia International Digital Libraries, 简称 MIND) 的“视音频元数据”、美国丹佛大学 Penrose 图书馆发布的“数字音频元数据”、美国卡内基·梅隆大学的“Informedia 项目”和传统机读目录格式 (MARC) 中的有关音乐的元数据元素集等^①。

一些基于可编辑文本的音乐描述语言也在一定范畴内得到传播, 如以 ASC II 为基础的音乐描述语言 GUIDO Music Notation Language, MusiXTeX, Musedata format, Philip's Music Writer, Note Edit, OMNL (Open Music Notation Language), Music Kit Score File, Muscript 等^②。

数字时代的音乐媒体包括数字乐谱、数字音频和数字音序, 其中数字音频技术围绕着音乐的数字信号展开, 数字音序则以 MIDI 技术为代表, 而数字乐谱主要以数字图像或字符序列的形式呈现, 如 JPG 图像格式等。各种数字音乐媒体的产生、传播、变换和存储形成了千万种丰富的音乐技术, 不同技术在历史的长河中经历了产生、使用、传播和消亡的生命周期, 有些技术在全球范围内得到接受和应用, 有些技术在一定的音乐范畴中得到应用, 有些技术则可能渐渐消失。

在不同的时代, 声音是基本的音乐载体, 人耳能感知的声音振动频率范围在 20 Hz ~ 20kHz 之间, 见图 1-3。不同频率的声音给人不同的心理感受, 在频率数轴上, 声音的频率是连续的, 即在任何两个不同的频率之间, 存在能产生无数个不同频率的声音。而在音乐中使用的乐音的频率, 选择了在声音频率数轴上有限个离散的频率点构成了基本的乐音体系, 这是在频率上的离散处理, 见图 1-3, 从声音到乐音的离散处理。

① IBM, Thinking XML: Manage metadata with MusicBrainz, <https://www.ibm.com/developerworks/library/x-think14.html>.

② Gerd Castan, *Datenformate of Musiknotation*, <http://www.music-notation.info/en/compmus/notationformats.html>.

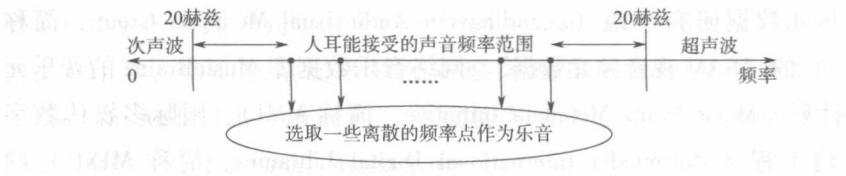


图 1-3 在频率上选择的乐音

为了进行数字化处理，对模拟的音乐乐音利用美国物理学家哈利·奈奎斯特（Harry Nyquist）的采样定理进行了时间上的离散采样，形成了数字音频，见图 1-4。数字音频的采样频率、量化精度和分辨率则决定了数字音乐与模拟声音之间的信息正相关性，如两种数字音频格式 MP3 和 WAV 对同一音乐的采样结果的存储大小不同，它们对音乐的失真率也不同。

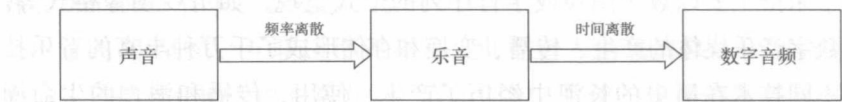


图 1-4 音乐媒体在频率与时间上的两次离散选择

虚拟音乐时代必将在数字音乐时代的迅猛发展下出现，在虚拟音乐时代，音乐的展现方式、人与音乐之间的交互等都将发生翻天覆地的变化。如音乐的同步问题，在传统交响乐团的演出中，需要不同的音乐家处在一个相同的时空中，彼此在指挥的引领下，演奏出完美的乐曲；在数字时代，作曲家可以利用软件对不同的乐器轨道分别进行录音和编辑，最后合成为一首完整的乐曲；在虚拟时代，不同的音乐家、作曲家在不同的地点、不同的时间完成不同的乐器的音频，而对听众而言，这些乐器的声音通过互联网同时在私人空间响起，形成一个完整的乐曲，这里需要对不同的音频进行同步处理。

当人们弹奏钢琴、拨弄吉他等乐器的时候，需要与乐器进行接触，传递力量给相应的接触面，不同的动作和力度能产生不同的音像效果。在虚拟时代，人们只要戴上数据手套或者面对摄像机，就可以演奏不同的乐器，不需要真实的乐器存在，触觉等不同的感觉器官被解放，人与音乐的

交互将发生根本的变化。作曲将更加容易，只要在大脑中形成相应的旋律，乐曲就被自动记录，乐谱就自动产生。对于实际演出而言，《第九交响曲》的乐队指挥可以是机器人，也可以是虚拟贝多芬。

1988年，马克·维瑟（Mark Weiser）提出了普适计算的概念^①，普适计算能随时随地向人提供信息和服务的计算资源，大大改进了人与环境的关系。在耳机市场，可穿戴式无线耳机逐步被人们所接受。2008年以来，通过蓝牙技术的分米波无线耳机渐渐成为中国市场的主流，广场舞由于声音扰民被人们诟病的现象也将消失，人手一个无线耳机的广场舞将逐步普及。

2013年8月，华盛顿大学公布了一项研究成果^②，人类脑对脑接口实验获得成功，脑信号通过互联网远程传输给同伴，同伴接受后，完成同样的动作，实现了动作同步的效果，这项研究的影响深远。在音乐教学中，面对面个性化教学被学者们认为是最有效的教学方式，这项技术使得师生之间的动作模仿成为可能，极大地提高了动作模仿的有效性。再者，一位没有任何音乐基础的普通人通过互联网利用动作同步效果，能瞬间成为著名演奏家的替身。

虚拟现实技术在音乐演奏与体验中得到充分应用，乐器仿真变得很方便，每个人都能产生不同的虚拟乐器，虚拟乐器具有声像逼真、人与乐器交互（演奏）逼真、乐器形象逼真等特征，沉浸式体验逐步取代真实乐器演奏。

虚拟音乐时期需要解决的两个基本问题：其一为任意时间的音乐组合，其二为跨空间的音乐组合。打破时空限制，使得任意时间中产生的音乐能任意组合，无论何地的音乐能无缝组合。同时还应该解决声响空间的分割问题，即把一个有限空间分割为若干个子空间，随意地向指定的子空间传送不同的音乐信号，使得在不同子空间中产生不同的声响效果。

① 郑增威、吴朝晖：《普适计算综述》，《计算机科学》2003年第3期。

② Doree Armstrong, Michelle Ma, Researcher controls colleague's motions in 1st human brain-to-brain interface, <http://www.washington.edunews2013/08/27/researcher-controls-colleagues-motions-in-1st-human-brain-to-brain-interface/>.

第二节 智能音乐学概述

笔者通过对 20 世纪下半叶至 21 世纪初国内外关于数字音乐的发展与研究成果的梳理和分析,认为人工智能已经在音乐领域取得了丰硕的成果,随后笔者于 2013 年首次提出“智能音乐学”^①这一学科概念,并认为:数字音乐=音乐媒体+智能音乐学。其中音乐媒体是音乐信息的载体,是音乐数字化表示技术的总和,而智能音乐学则主要利用人工智能的方法来研究音乐媒体之间的相互转化及变换过程,它体现了人工智能在音乐艺术领域的拟人化、个性化、最优化的愿景目标。智能音乐学可以说是人工智能一个新的研究领域,下文首先对人工智能进行概述,然后将重点论述智能音乐学的主要内容。

一、人工智能概述

人工智能(Artificial Intelligence,简称 AI)是关于人类知识的获取、知识的表示和知识使用的科学,以人的智能的理论、方法、技术及应用系统为研究对象,模拟、延伸和扩展人的智能的一门技术科学。它的研究领域涉及数学、哲学、心理学、信息论、认知科学、神经生理学、计算机科学、控制论和仿生学等跨领域学科。

人工智能的研究目标是研究如何利用信息技术特别是计算机技术去完成人才能完成的工作,因此,它的研究主要依赖于计算机去模拟人类的一些智力活动的基本理论、技术和方法。它的研究领域包括问题求解、模式识别、专家系统、机器视觉、自然语言理解、自动程序设计、自动定理证明、机器学习和机器人系统等。

模式识别是利用信息技术对信号、图像、音频、字符等多种不同的信息模式进行自动分类识别的科学,是对指定对象的信息描述方式的转换研究。如对图像中的人脸进行识别,识别后用文本方式来表示人名,这是一种从图像媒体到文本媒体的信息描述方式的转换。一般的模式识别过程包

^① 中国科技论文网(www.payoer.edu.cn.)。

括待处理对象的样本采集、对象的多维度信息获取和表示、信息特征选择和分类器设计等过程。常见的分类方法有贝叶斯决策理论、神经网络、隐马尔可夫链、决策树、支持向量机、感知机、最近邻法、最大似然估计、核函数、蚁群算法、模糊方法和聚类分析等。模式识别技术有广泛的应用,在农业、工业、生活、航空、军事、医学等领域有深入的应用,如语言识别、说话人识别、字符识别、车牌识别、数字识别、文字识别、图像识别、指纹识别、地图识别、照片识别、脑电图识别、心电图识别、震动波识别、植被遥感识别等。

问题求解是指通过给定的搜索方法对问题的状态空间进行匹配,寻找一个状态空间的子空间,满足问题的要求。问题求解面对的状态空间往往是高维或者大数据的集合,对搜索方法的要求较高。一般问题求解步骤有状态空间表示、搜索算法、解空间表示三个阶段,常见的搜索算法有二分法、深度优先法、广度优先法、动态规划法、回溯法、贪心法、概率法、近似法、分支限界法、启发式搜索策略等。问题求解的搜索算法有广泛的应用,在生活生产的各个领域都有应用。如经典的排序问题、拓扑排序问题、凸包问题、字符匹配问题、最短路径问题、最大流问题、旅行商问题、装箱问题、背包问题、八皇后问题、整数规划问题、哈密顿回路问题等。

自然语言理解是指机器能理解与正确回应人类发出的自然语言所携带的信息,包括人机问答系统、自然语言理解系统、联机或脱机手写文字识别系统、机器翻译系统和多语种互译系统等。这是一个有挑战性的研究领域,有着广泛的应用领域。如呼叫中心的自动信息问答服务,屏幕(手机等终端)手写文字识别,多语种交流中的语言互相翻译等应用。

机器视觉的对象是图像,机器视觉涉及图像处理、图像分割、图像描述、特征提取、图像理解和分类器设计等方面的内容。通常,图像处理的对象是图像的像素,处理的结果是图像或者图像的各种变化,如几何变换、傅立叶变换、小波变换等;图像分割的目的是从图像中提取出感兴趣的焦点区域,图像描述主要考虑选择合理的数据结构来描述相关的图像信息,特征提取是从图像信息中提取出用于分类的数据集,图像理解以对图像的对象进行识别为目的。机器视觉广泛应用于机器人、自动驾驶、3D视觉分析、步态分析、运动分析等应用领域。