

文章编号:1006-8309(2005)01-0059-04

可听化技术研究与发展趋势

张丽红¹,方志刚^{1,2},徐义东¹

(1. 浙江大学信息与电子工程学系,浙江 杭州 310027;

2. 浙江大学城市学院,浙江 杭州 310015)

摘要: 可听化技术是听觉显示研究的一个重要研究内容,其核心内容是对用户的听觉感知的研究。文章简要回顾近年来国内外在听觉显示和可听化的研究成果,概述了可听化技术的主要研究内容,以及可听化研究和应用的开发工具的要求;并对可听化技术的应用前景和发展趋势进行了探讨。

关键词: 可听化; 听觉显示; 感知认识; 隐喻; 可用性

中图分类号: TP391 **文献标识码:** A

1 引言

可视化技术以直观的图形方式将信息模式、数据关联或趋势呈现给决策者,在许多领域都有广泛的应用。然而,可视化的信息显示维度有限,在处理和连续多维海量数据时,单靠视觉通道已显得力不从心。人们考虑利用听觉的优势来解决视觉不能单独完成的任务,降低视觉负荷^[1]。

可听化(Sonification)将所研究领域各维度信息间关系映射成声音各维度(频率、音色、音调等)间的关系,即利用非语音信号来表征信息,解释、传达所研究领域数据的各种特征,有助于信息提取或监控^[2]。本文概述了可听化技术研究的主要内容,以及开发工具的研究情况及其要求,希望我们的工作能推动听觉显示和可听化研究的进一步发展。

2 听觉显示研究现状

在可听化技术出现前,听觉显示研究主要采用声音的突发性或声音序列来表示数据,如听标(auditory icons)和耳标(earcons)。盲人用户手机界面^[3]就是利用听标和耳标的应用实例,使手机的界面更适合盲人用户使用。听标和耳标的这种数据-声音映射关系属于离散事件的映射。

可听化技术则注重描述连续变化过程,属于连续事件映射,是听觉显示的重要研究方向。用户通过“听”声音获取所需信息。国外在这方面已有不少研究,如盖革计数器(Geiger counter)^[2],脉

冲血氧计(Pulse oximeter)^[2],“量子口哨”(Quantum whistle)^[2]等;医学数据的可听化,医生通过“听”心电图^[4]、脑电图^[5]、膝关节震动信号^[6],实现计算机辅助听诊;股市数据的可听化^[7]分析行情趋势;视觉障碍用户的听觉导盲工具研究^[8,9]。国内,浙江大学计算机辅助设计与图形学国家重点实验室研究3D虚拟现实中的声境营造^[10];利用室内声学的可听化模拟厅堂听音环境,对三维空间音质进行可听化仿真^[11],等等。

自1992年起,国际听觉显示委员会(ICAD)每两年召开一次年会,2000年后每年一次,讨论有关听觉显示的问题,并建立了一套基本的理论和应用原型。

3 可听化技术研究

可听化设计由数据预处理、数据声音映射、声音合成和可用性测试四个模块组成,如图1。

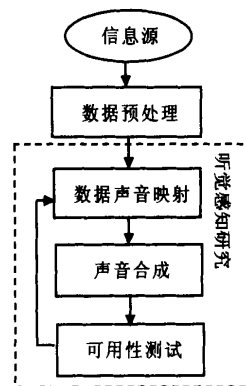


图1 可听化设计流程图

基金项目:浙江省自然科学基金资助项目(602060)

作者简介:张丽红(1981—),女,浙江奉化人,硕士研究生,研究方向为听觉显示技术和人机交互。

3.1 数据预处理

由于收集到的原始数据存在杂乱性、重复性及不完整性,数据预处理的目的是滤除不需要的信息,恢复、增强有用信号,并提取具有代表性的特征数据,降低数据声音映射的复杂性,提高用户的听觉感知能力。

3.2 听觉感知研究

根据提取的(特征)数据选择合适的映射方案,进而合成声音。系统的可用性测试则针对不同民族、地区及年龄用户的感知认识能力进行测试,并对系统整体性能进行评价。这三部分围绕用户听觉心理和生理学的研究,称为用户听觉感知研究。

3.2.1 听觉空间各维度的动态作用

从声音感知认识角度考虑,声音的维度有音调、响度和音色等。在处理海量数据时,必须利用声音多维度才能不失真的表现数据信息。但在用户听力空间(Using Listening Space, ULS)中,声音各维度之间存在着动态影响^[12, 13],如音调和频率是非线性的关系,当频率增加一倍,产生的音调并不会增加一倍;又如高音调低响度产生的效果与低音调高响度产生的效果很难被分辨出来。因此,要使声音所提供的信息与用户所理解的信息能快速地达到一致,就要充分考虑这些因素的影响,避免提供的信息模糊化。

3.2.2 隐喻映射方法

选择数据声音的映射关系时,使事件与呈现的听觉显示之间具有一定的相似性,这种映射称为隐喻映射。例如,某一数据的增大映射到声音某维度增大。在人的思维中,数据空间模型和数据显示模型具有很大的相关性,使用户能快速理解数据的变化趋势。这种具有最大相关性的映射显然是最直观和“自然”的选择。因此,数据映射的选取必须满足声音与人的听觉能力达到优化匹配,如某一剧烈变化的信号用不断变化的频率表示还是用不断变化的幅度表示更适合人的感知能力?

3.2.3 听觉流的研究

根据听觉场景分析所揭示的人自身的听觉感知机理,在实现多维数据的可听化时,根据数据的特征,可以考虑多元声音显示,即结合连续和离散的映射方式,以听觉流的形式展示数据,丰富声音

的内涵。但用户对听觉场景的分析是一个非常复杂的过程,仅仅考虑声音的可分辨性并不能满足听觉显示的要求。而且,人们所熟知的听觉鸡尾酒会效应绝大多数是关于语音,而不是非语音信号。因此,听觉流的研究有多少潜力,这一过程中的记忆负荷、注意力以及其它影响信息传递的因素还需进一步的研究。

3.2.4 “训练”的重要性

用户经过“训练”,能提高对声音的感知能力,从而提高系统的有效性和效率。如电子行走辅助(Electronic Travel Aid, ETA)系统研究中,盲人利用系统提供的声音获取路面信息。实验表明,未接受“训练”的用户对环境的认知能力只有4.3%,而接受“训练”的用户的认知能力高达80.7%^[14]。可见,通过“训练”,用户积累了一定的知觉经验,对可听化设计的成功与否起到很大的作用。

3.3 可用性测试

人们对听觉感知这一心理过程进行了大量的研究,但数据映射的实现比较盲目,缺乏统一的标准和坚实的理论基础,系统可用性测试就显得非常重要。以心理学理论和工效学评价为基础对映射结果进行测试,需要从3个方面考虑:①有效性:用户能否准确地理解系统声音所提供的信息;②效率:用户能准确理解声音的信息与所需时间之间的比率;③满意度:用户在使用过程中的主观感受和接受程度。

4 可听化研究和应用工具

国内外在可听化工具及相关方面做了许多工作。Stanford 大学开发的 Synthesis Tool Kit (STK)^[15]可听化应用研究软件包;德国 Bielefeld 大学的 AVDisplay^[16]工具实现听觉显示与视觉显示结合;浙江大学的听觉显示开发平台(Auditory Display toolKit, ADK)^[17]能实现实时和非实时的声音合成,支持离散和连续事件的可听化,既能进行可听化的应用研究,也可以对听觉感知、可用性测试等方面进行研究。对可听化开发工具或软件包而言,须具备以下特点:

(1)灵活性(Flexibility):除合成技术的多样性外,数据映射方案的改变等简单控制在设计中也是必要的。从而能从多种途径控制数据映射,实现数据可听化,并适用不同层次和喜好的用户,提高工具的人性化。

(2)便携性(Portability): 开发工具能在不同的计算机平台之间、实时和非实时声音合成产品之间转移,可移植性强。

(3)可集成性(Integrability): 能与可视化程序、电子表格、实验设备等软硬件进行方便的连接,提供标准化的软件接口,整合数据输入、声音映射和合成。

5 可听化应用前景和发展趋势

可听化系统能帮助用户快速理解声音所表达的信息,并缓解视觉过载或代替视觉完成其不能完成的任务。因此,可听化技术具有广阔的应用前景:

(1)数据挖掘: 随着听觉感知研究的不断深入,可听化技术与可视化技术的结合,将成为数据挖掘技术更强有力的工具。

(2)导盲设备: 帮助视觉受障或失明用户获取环境信息。如解决消防队员因现场的恶劣环境使其视觉受障问题; 电子行走辅助研究解决盲人的出行问题等。

(3)数据监控: 当用户需要同时从事多项活动时,可以考虑用听觉通道来缓解视觉负担。如实现报表数据的可听化,使操作师能一边控制机器,又能“听”取不断变化的数据,提高了操作的安全性。

(4)其他还有数据表示/理解、数据探查/分析、分类任务等。

可听化研究与听觉心理和生理研究的进展是密切相关的。根据不同的应用领域与知识背景,可听化研究和发展方向将遵循以下几个方面:

(1)随着听觉感知和认识研究的进一步深入,需要加强听觉心理学和生理学的研究。可听化是通过数据到声音的映射向用户传达信息,因此,对声音的感知理解的准确性是任何可听化应用成功的关键。

(2)可听化研发工具的研究,包括硬件和软件工具。处理大容量数据时,希望用常用的 DSP 和 ASIC 提高算法的速度和减小系统体积,使可听化技术走向实际应用系统。

(3)可听化应用研究。虽然可听化已成功地应用于许多领域,但在实际应用中仍无法设计一个可听化应用并能预先知道这一设计能很好地实现指定的任务。可听化的继续发展需要有前瞻性

的设计原则。只有这样,可听化才能够从一些特殊的实验发展为相连贯的设计、研究和应用的领域。

参考文献:

- [1] 方志刚, 吴晓波, 徐新民, 等. 非语音声音在人机交互技术中的应用[J]. 计算机科学, 1997, 6(24): 62-65.
- [2] Kramer G, Walke B, Cook P, et al. Sonification report: status of the field and research agenda [EB/OL]. <http://www.icad.org>. 2004-06-12.
- [3] 胡国兴, 方志刚, 李清水. 听觉界面和盲人用户手机界面[J]. 人类工效学, 2003, 9(2): 7-10.
- [4] Ballora M, Pennycook B, Ivanov P, et al. Detection of obstructive sleep apnea through auditory display of heart rate variability[A]. Proc of CiC 2000 [C]. Boston: Computers in Cardiology, 2000. 739-740.
- [5] Jovanov E, Wagner K, Radivojevic V, et al. Tactical audio and acoustic rendering in biomedical applications [J]. IEEE Transactions on Information Technology in Biomedicine, 1999, 3(2): 109-118.
- [6] Krishnan S, Rangayyan RM, Bell GD, et al. Sonification of knee-joint vibration signals[A]. Proc of the 22nd Annual Int Conf of the IEEE(vol 3)[C]. Chicago IL: EMBS, 2000. 1995-1998.
- [7] Keith VN, Stephen B. Evaluation of a multimodal sonification and visualization of depth of market stock data[A]. Proc of the 2002 Int Conf on Auditory Display[C]. Kyoto: ICAD, 2002. 1-6.
- [8] Milos E, Kapralos B, Kopinska A, et al. Sonification of range information for 3-D space perception[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2003, 11(4): 416-421.
- [9] Nagarajan R, Yacob S, Sainayanan G. Role of object identification in sonification system for visually impaired[A]. Conf on Convergent Technologies for Asia-Pacific Region (vol 2)[C]. Bangalore, India: TEN-CON, 2003. 735-739.
- [10] 张琼, 石教英. 分布式虚拟环境中的三维音频支持[J]. 计算机辅助设计与图形学学报, 2002, 14(12): 1152-1155.
- [11] 曾向阳, 郭晓元, 陈克安, 等. 三维空间宽带音质可听化仿真[J]. 计算机仿真, 2003, 20(1): 62-64.
- [12] Kaper HG, Wiebel E, Tinei S. Data sonification and sound visualization[J]. Computing in Science & Engineering, 1999, 1(4): 48-58.

[13] Neuhoff JG, Kramer G, Wayand J. Pitch and loudness interact in auditory displays: can the data get lost in the map? [J] . Journal of Experimental Psychology, 2002, 8(1): 17-25.

[14] Capelle C, Trullemans C, Arno P, et al. A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution [J] . IEEE Transactions on Biomedical Engineering, 1998, 45 (10): 1279-1293.

[15] Cook PR, Scavone GP. The synthesis toolkit in C++ (STK) [EB/ OL] . <http://www-ccrma.stanford.edu/software/stk>, 2004-06-12.

[16] Thomas H, Christian N, Helge R. Interactive visualization and sonification for monitoring complex processes [A] . Proc of the 2003 Int Conf on Auditory Display [C] . Boston: ICAD, 2003. 247-250.

[17] 徐义东. 听觉显示开发平台的设计及应用 [D] . 杭州: 浙江大学, 2004.

[收稿日期] 2004-02-26

[修回日期] 2004-06-21

(上接第 58 页)

[10] 王伟, 葛列众, 李宏汀. 微光环境下显示器(CRT)颜色视标相对辨别能力的研究 [J] . 人类工效学, 2000, 6(4): 14-17.

[11] 葛列众, 王伟, 李宏汀. 微光环境下显示器(CRT)颜色视标呈现时间和视标面积交互作用的研究 [J] . 人类工效学, 2002, 8(2): 8-10.

[12] 钟浩, 姜国华, 王颖, 等. 液晶汉字尺寸对视觉识别工效的影响 [A] . 龙升照. 人-机-环境系统工程研究进展(第三卷) [C] . 北京: 北京科学技术出版社, 1997. 160-165.

[13] 牟晓非, 田志强, 王颖, 等. 液晶显示器在不同的环境照明和亮度对比度条件下可读性的研究 [A] . 龙升照. 人-机-环境系统工程研究进展(第五卷) [C] . 北京: 海洋出版社, 2001. 224-229.

[14] 许百华, 傅小贞. 液晶显示的红、绿色光对视觉暗适应影响的比较研究 [J] . 人类工效学, 2002, 8 (4): 17-19.

[15] 郭小朝, 刘宝善, 武国城, 等. 战斗机话音告警汉语用语的选定 [J] . 中华航空医学杂志, 1994, 5 (4): 230-233.

[16] 刘宝善, 武国城, 郭小朝, 等. 战斗机汉语合成话音适宜音量测定和频谱分析 [J] . 中华航空医学杂志, 1995, 6 (2): 91-96.

[17] 刘宝善, 武国城, 郭小朝, 等. 战斗机汉语合成话音告警用语设计参数的测定 [J] . 中华航空医学杂志, 1995, 6 (3): 176-180.

[18] GJB 185-1986, 飞机飞行品质 [S] .

[19] 刘宝善. 握杆操纵技术中手的数据和手指功能 [J] . 中华航空医学杂志, 1992, 3(4): 206.

[20] 牟伟民. 座舱显示器周边各键尺寸、方位和间隔的工效学实验研究 [J] . 人类工效学, 1997, 3(4): 10-12.

[收稿日期] 2004-09-16

[修回日期] 2004-12-01