

AudioMan: 电子行走辅助系统的设计与实现

徐洁¹⁾ 方志刚^{1),2)} 鲍福良¹⁾ 张丽红²⁾

¹⁾(浙江大学城市学院信电分院, 杭州 310015) ²⁾(浙江大学信息学院, 杭州 310027)

摘要 针对现有的电子行走辅助(ETA)系统在普及和应用中的缺陷,自行设计了一种易使用的ETA系统模型AudioMan。该模型结合图像处理和听觉显示技术,给盲人用户提供道路方向和障碍物位置等环境信息。系统提出通过控制乐音的参数来实现可听化,用可听化来传递图像信息。可用性测试和评估结果表明,AudioMan所建立的用户听觉模型比现有的ETA系统更有效。

关键词 电子行走辅助 人机交互 听觉显示 可听化

中图分类号: TP391.7 **文献标识码:** A **文章编号:** 1006-8961(2007)07-1249-05

AudioMan: Design and Implementation of Electronic Travel Aid

XU Jie¹⁾, FANG Zhi-gang^{1),2)}, BAO Fu-liang¹⁾, ZHANG Li-hong²⁾

¹⁾(School of Information and Electric, Zhejiang University City College, Hangzhou 310015)

²⁾(Department of Information & Electrical Engineering, Zhejiang University, Hangzhou 310027)

Abstract Aiming at the popularization and the application flaw in the existing Electronic Travel Aid (ETA) system, a model of ETA system named AudioMan is designed, which is simple in structure, easy to carry and use and of low cost. AudioMan uses computer digital image processing technology to acquire the information of the road direction and barrier positions of user's environment. Also it uses auditory display technologies such as sonification, auditory icon to show blind people the environmental information. In this system, sonification technology is realized through the control of musical parameters; and the images information transmission can also be realized by sonification and auditory icon or earicon. The result of the usability test and appraisal indicates that the user sense of hearing model AudioMan establishes is more effective than the existing ETA system; at the same time it will help the blind understand the information of the sound more quickly and effectively.

Keywords electronic travel aid, human computer interaction, auditory display, sonification

1 引言

从20世纪70年代开始,各国学者开展了电子行走辅助系统(electronic travel aids,ETA^[1])的研究工作,ETA系统利用传感设备获取环境数据信息,经过数据处理,将其转换成易于盲人理解的非视觉信号,目的是辅助盲人独立而安全地行走。从信息显示方式的角度来说,现今存在的ETA系统可分成两大类:一是听觉代替视觉系统,代表系统有

Navbelt^[2]、vOICE^[3]系统等;二是触觉代替视觉系统,代表系统有People Sensor^[4]等。第1类系统的最大缺点在于声音的不同维度的动态变化会影响到听觉感知效果,从而影响映射维度的选择;系统声音与环境声音的相互影响,降低了盲人特有的对环境声音的敏感程度。如vOICE系统采用逐列扫描方式,将在不同行列位置的像素点映射到声音某个参数,从而合成声音。该系统在理想的环境(背景噪声小,同时只存在简单物体)中比较适用;但在实际的生活环境中,背景复杂,同时障碍物也不规

基金项目:浙江省自然科学基金项目(602060);浙江省科技厅项目(2006C31006)

收稿日期:2006-10-17;改回日期:2007-01-23

第一作者简介:徐洁(1973~),女,讲师。1999年于浙江大学获工学硕士学位。研究方向为人机交互和生物认证。E-mail:xujie@zucc.edu.cn

则,合成的声音就非常混乱,且缺乏规律,不能很直观表现出环境信息,用户必须像学习语言一样经过长期训练才能理解声音所包含的内容;同时由于采用纯音合成,产生的声音效果非常令用户排斥,不可能长时间的使用。第 2 类系统通常采用对触觉具有高度敏感的部位作为传递点,当检测到障碍物时,通过振动信号来提醒用户。People Sensor 用两种不同振动信号,表示在检测范围内出现有生命和无生命的障碍物,当检测到的目标靠近盲人时,反馈振动信号变得越来越强烈。该类系统的最大缺点在于容易引起使用者局部肌肉疼痛、麻痹等副作用,并且信息显示维度比较有限,难以显示多种信息。所以,尽管各国学者和专家们已经花了 30 多年的时间研究和开发各种电子行走辅助系统,但 ETA 系统的普及和应用始终不是很理想。在国内,有关 ETA 的研究报道几乎是空白。

本文结合图像处理技术和听觉显示技术研发方面的积累^[5-9],提出一种结构简单、携带方便、易使用、低成本的 ETA 系统模型——AudioMan,该系统利用计算机图像处理技术获取环境中有关道路方向和障碍物位置的信息,结合可听化、听标和耳标的听觉显示技术向盲人用户展现环境信息。相比现存的 ETA 系统,AudioMan 系统第一次提出通过控制乐音的参数来实现可听化,用可听化来传递图像信息,数据映射采用了直观而自然的隐喻映射,使用户在不经训练或短时间的学习就能适应目标系统。另外,本文还从可用性和用户体验角度,对系统提供的 4 种方向映射方案进行了评估。

2 系统概念模型

AudioMan 是用于辅助盲人行走的系统模型,是一种支持图像处理和听觉显示的 ETA 系统。AudioMan 模型由单个摄像头、便携式计算机和耳机 3 部分组成,如图 1 所示。摄像头和耳机只是系统用于获取环境信息和表现声音信息的外部设备,系统所有的工作全部由计算机完成。图 1 虚线框内所

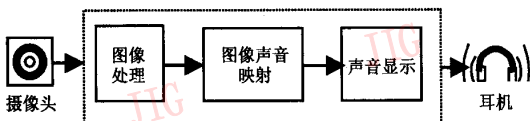


图 1 AudioMan 模型

Fig. 1 Model chart of AudioMan

列出的是计算机需要处理的工作,包括获取环境信息的数据,合成声音等处理步骤。

3 系统结构

AudioMan 的结构如图 2 所示,按照数据的流向依次为图像处理模块(image process model),数据映射/声音合成模块(mapping & sound synthesis module),用户可用性测试模块(usability testing model)。其中前两个模块受主控模块(controller model)的控制,并与主控模块进行数据交互。

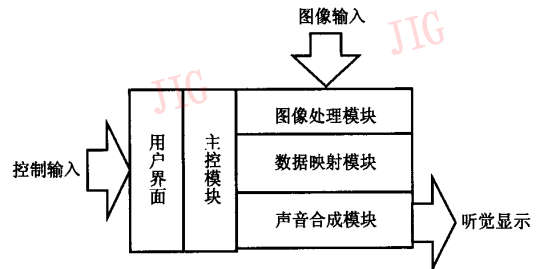


图 2 AudioMan 结构图

Fig. 2 Structure chart of AudioMan

(1)图像处理模块 其是系统获取环境中信息的重要模块,是声音能有效地显示环境状态的基础。该模块的工作是通过图像处理技术,从摄像头获取的图像信息中提取出道路的方向和各种障碍物的位置,并将这些环境信息传送给主控模块。

(2)数据映射/声音合成模块 从主控模块中得到有关环境的信息,并根据用户选择的显示方案进行环境信息向声音的映射,采用缺省的数据映射方案和多种备选的声音合成和显示方法。由于对不同的用户适合的方案可能不同,因此就必须进行可用性测试,评估各种不同的声音合成方法,从而选择最适合用户的。此外,模块能控制 MIDI 合成和 WAVE 乐音某一参数(如方位、响度、节奏等)来表现环境中的信息,这种控制方法中采用的映射规则与纯音的可听化映射方式类似。

(3)主控模块 其参数区包括图像处理状态区,缓冲区的图像数据,环境特征参数区,MIDI 设备参数控制区以及声音合成状态区。该模块的工作就是协调前两个模块数据和信息的交流。只有该模块对所有的参数区内的数据都是可见的,其他模块处理得到的结果和需要访问的数据都必须通过主控模块这个中介。

AudioMan 与用户的交互主要是通过声音进行的,因此在系统设计过程中尽量减少要求用户输入的信息量。当然用户也有一定的权力选择自己喜欢的显示方案和适合的乐音。另外,模块化的设计使得 AudioMan 的功能扩展比较容易。

4 系统数据流图

4.1 图像处理模块

图像处理模块的数据流图如图 3 所示。

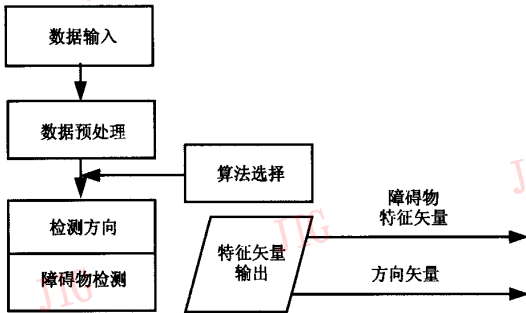


图 3 图像处理模块的数据流图

Fig. 3 Data flow chart of video processing model

该模块主要完成两件事情:检测道路方向和检测障碍物位置。实验初期,AudioMan 系统工作在实验室模拟的环境中。在该环境中,设置了几种基本道路信息,如道路转弯的情况,简单的障碍物等,因此,在这种环境中外界干扰少,加上盲人行走速度比较缓慢,AudioMan 原型系统对视频的单帧图像进行处理。

图像处理模块根据从主控模块读入的图像处理启动控制参数。该模块启动后,读取并处理缓冲区的单帧图像,提取有关道路方向和障碍物的数据特征,输出的数据特征参数通过读写接口输出到主控模块的环境特征参数区。每当缓冲区内装满一帧数据,系统就自动调用帧回调函数 (fpFrameCallback)。按照摄像头的规格,设置帧显示率为 66ms/frame。每次调用回调函数时,都会检测“图像开始处理”标志。该标志在用户启动方向映射被置位。当检测到该标志被置位,图像处理模块被启动,即启动帧图像处理线程。当帧图像处理线程 (FrameThreadProc) 被启动后,在该线程内完成方向与障碍物的检测,以及启动相应的听觉显示线程。

图像处理模块中定义了 6 个图像处理和数据读

取接口:

(1) 图像预处理接口 ADIPreprocess:完成单帧图像的灰度化、大小调整和二值化的处理。

(2) 方向检测接口 ADIGetOrientationVector:进行分块图像傅里叶变换,进而取得方向的特征矢量 O 。

(3) 障碍物检测接口 ADIGetObstaclePos:对二值化后的图像进行腐蚀处理,从而得到障碍物在整个环境中的位置。

(4) 方向特征矢量提取接口 ADIReadVector:读取 O 。

(5) 最大方向偏离值提取接口 ADIReadOrientation:读取方向矢量 O 中的偏离中心最大的参量 O_{max} ,将这个值进行方向的映射。

(6) 障碍物特征提取接口 ADIReadObstacle:从图像处理模块中取得障碍物特征矢量。

4.2 数据映射/声音合成模块

图像处理模块的数据流图如图 4 所示。

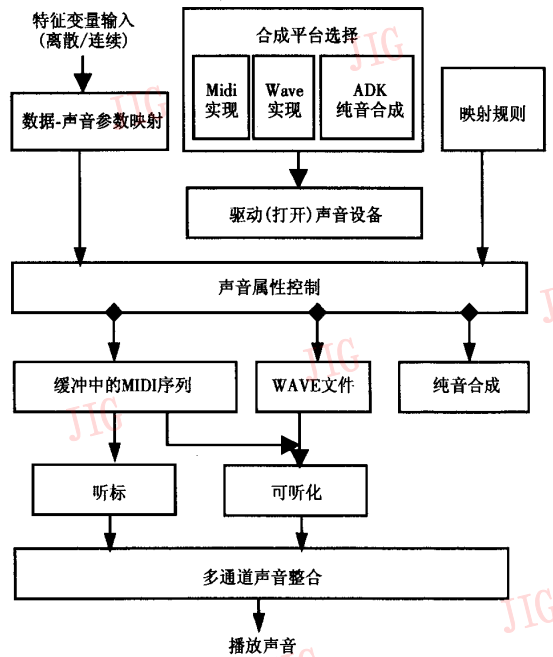


图 4 图像处理模块的数据流图

Fig. 4 Data flow chart of video processing model

该模块包含了从图像处理模块中得到的环境特征参数映射到声音某一维度的过程以及声音合成与实现。采用缺省的数据映射方案和多种备选的声音合成和显示方法,通过读写接口取得有关道路方向

和障碍物的数据特征参数,作为映射的输入,同时根据映射规则,用于确定映射输入到声音参数间的具体映射函数关系,最后映射模块根据计算出的声音参数调用纯音合成模块中的发声函数。该模块设计应用了心理听觉感知因素、MIDI 规范、计算机底层的硬件知识、操作系统等知识,采用多线程的工作方式,处理和完成图像处理线程和各种声音映射/合成线程。为了使各种线程之间能很好地协调和分配系统的资源,引入了一些线程的同步信号。系统提供了 3 种方向合成的方案,每种方案都定义了各自的线程函数。包括:

(1) 纯音合成线程: UINT AMPureThreadProc (LPVOID pParam);

(2) WAVE 乐音方位控制可听化线程: UINT AMMusicThreadProc (LPVOID pParam);

(3) MIDI 节奏控制的 可听化线程: UINT AMRhythmThreadProc (LPVOID pParam);

(4) 语音显示线程: UINT AMVoiceThreadProc (LPVOID pParam);

(5) 垃圾桶听觉显示启动线程: UINT DemoGarbageFun (LPVOID pParam);

(6) 消防栓听觉显示启动线程: UINT DemoPoleFun (LPVOID pParam);

5 系统的设计实现和评估

AudioMan 采用了 VC4U 型号的摄像机,成像速度为 15fps,分辨率为 320 × 240;成像距离选择在 2m 到 3m 之间。

AudioMan 能运行在两种工作方式下:实时处理的运行方式和脱机运行方式。实时处理的工作方式是指系统实时地从摄像头中获取数据并对得到的数据进行处理,实时地控制数据声音映射和合成。相比用摄像头实时地获取图像数据并进行处理和映射,系统提供了另一种运行方式,原始数据的提取并不是用摄像头实时获取,而是事先已经存放在指定的目录中;但对界面当前显示的数据的处理和数据映射是实时的过程,这样的工作方式称为准实时的过程,又称脱机运行。对复杂的感知系统,通过一定的人机交互的学习和“训练”,使用户积累一定的听觉经验,能提高用户的听觉感知能力,提高系统的有效性和效率,这对成功地听觉显示设计起到很大的作用。因此,脱机运行方式的设计能帮助用户学习

各种映射方案的工作情况,并对各种候选方案进行初步的测试和评估。

可用性测试和评估是可听化设计的重要部分。在设计 的 6 种映射方案中,选择其中 4 种区别比较大的方向映射方案进行测试:①纯音的频率显示;②纯音的方位显示;③乐音节奏的正极性变化显示;④乐音方位显示。采用“用户测试”的评估技术,得到各种映射方案用户满意度平均值和错误率数据如表 1 所示(测试用户样本为 10 个,每种方向映射方案的询问被测试次数为 20 次)。由表 1 得到各种方案的显示效率、有效性和用户听觉满意度比较结果为:方案④ ≥ 方案② ≥ 方案③ > 方案①。评估结果表明:采用乐音的方位显示是比较合乎用户的听觉心理。该方案不仅采用乐音作为映射的声音源,解决了用纯音显示容易引起听觉疲劳的问题,同时采用了比较直观而自然的隐喻映射方式,即事件与呈现的听觉显示之间具有一定的相似性,满足声音与人的听觉能力的最优化匹配,并帮助用户更容易理解声音所表现的信息。

表 1 用户满意度平均值和错误率

Tab. 1 User's average satisfaction and mistake rate

	方案①: 纯音频率	方案②: 纯音方位	方案③: 乐音节奏	方案④: 乐音方位
满意度平均值	2.5	4.035	3.107	4.821
即时辨别困难概率(%)	27.08	10.19	21.87	5.21
分辨错误率(%)	22.91	18.75	13.54	8.333

6 结论

根据对现有 ETA 系统的优缺点分析,认为采用图像输入和听觉显示技术的 ETA 系统具有很大的发展潜力。其核心是利用计算机视觉技术检测环境图像中方向和障碍物信息,并根据合理的映射方案将环境信息用声音展现给用户。AudioMan 原型系统首次提出采用可听化和听标(或耳标)两种听觉显示来传递图像信息,并尝试通过控制乐音的参数来实现可听化。评估结果表明,AudioMan 采用使用的隐喻映射方式,使用户不需要长时间学习,就能快速理解系统提供的用来表现环境信息的声音。因此,AudioMan 所建立的用户听觉模型比现有的 ETA 系统更有效。当然,如果要想 AudioMan 得到广泛的应用,系统的整体有效性和用户使用的满意度仍需要进行进一步的测试和研究。

参考文献 (References)

- 1 Pissaloux E. A characterization of vision systems for blind people mobility[A]. In: IEEE Interventional Conference on SMC (Vol 4) [C], Hammamet, Tunisia, 2002; 6~11.
- 2 Shoval S, Uirich I, Borenstein J. NavBelt and the GuideCane-robotics-based obstacle-avoidance systems for the blind and visually impaired[J]. IEEE Robotics & Automation Magazine, 2003, 10(1): 9~20.
- 3 Meijer P. Image-Audio transformation system[P]. USA: 5097326, March, 1992.
- 4 Ram S, Sharf J. The people sensor: A mobility aid for the visually impaired[A]. In: IEEE Second International Symposium Wearable Computers[C], Pittsburgh, USA, 1998: 166~167.
- 5 Hu Guo-xing, Li Qing-shui, Fang Zhi-gang. Auditory interface& mobile phone interface for blind users[J]. Chinese Journal of Ergonomics, 2003, 9(2): 7~10. [胡国兴,李清水,方志刚. 听觉界面与盲人用户手机界面[J]. 人类工效学, 2003, 9(2): 7~10.]
- 6 Xu Yi-dong, Fang Zhi-gang, Zhang Li-hong. Survey of sonification research [J]. Journal of Computer Aided Design & Computer Graphics, 2004, 16(1): 14~18. [徐义东,方志刚,张丽红. 可听化研究综述[J]. 计算机辅助设计和图形学学报, 2004, 16(1): 14~18.]
- 7 Zhang Li-hong, Fang Zhi-gang, Xu Yi-dong. Survey of sonification research and developing trend[J]. Chinese Journal of Ergonomics, 2005, 11(1): 59~62. [张丽红,方志刚,徐义东. 可听化技术研究与发展趋势[J]. 人类工效学, 2005, 11(1): 59~62.]
- 8 Xu Yi-dong. Design & Application of Auditory Display Development Platform[D]. Hangzhou: Zhejiang University, 2002. [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.]
- 9 Fang Zhi-gang, Hu Guo-xing, Wu Xiao-bo. Study of aural user interface based on non-speech Sound [J]. Journal of Zhengjiang University(Engineering Science), 2003, 37(6): 684~688. [方志刚,胡国兴,吴晓波. 基于非语音声音的听觉用户界面研究[J]. 浙江大学学报(工学版), 2003, 37(6): 684~688.]