

AudioMan 中方向矢量提取算法的研究和实现^①

Research and Implementation of Environment

Directional Vector in AudioMan

徐 洁 鲍福良 方志刚 (浙江大学城市学院信电分院 310015)

摘要:提出了一种电子行走辅助(ETA)系统模型——AudioMan 用于辅助盲人行走,盲人用户通过它的声音提示,来判断当前环境的情况。其重要研究内容之一是如何正确提取环境方向矢量。对提取过程中的两个主要步骤:边缘检测算法和道路偏转矢量的计算作了详细的介绍,针对现实生活中的盲道所做的实验获得了比较满意的效果。

关键词:电子行走辅助 方向矢量 边缘检测 道路偏转矢量

1 引言

从 20 世纪 70 年代开始,各国学者已开展了帮助盲人获取环境信息的电子行走辅助系(Electronic Travel Aids, ETA)的研究工作,如 Laser Cane^[1]、Fish R^[2] 的图像声音系统、vOICe^[3] 系统、Tyflos^[4] 盲人智能助手等。这些系统利用传感设备获取的环境数据信息,处理后的数据转换成易于盲人理解的非视觉信号,辅助盲人

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。因此,此类系统不能得到广泛的应用。

结合图像处理技术和在听觉显示技术研发方面的积累^[5,6,7,8],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。在该系统中,抛弃了从图像的像素点到声音的直接映射,而在环境信息向声音映射前加入了图像处理模块,提取出道路的方向和各种障碍物的位置,为声音能有效地显示环境奠定基础。本文主要介绍提取环境方向矢量过程中的两个主要步骤:边缘检测算法和道路偏转矢量的计算。

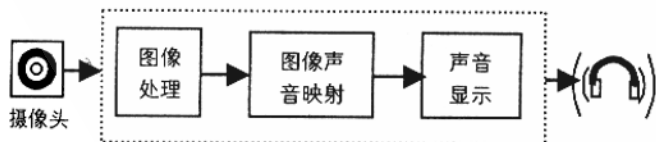


图 1 AudioMan 模型

独立而安全地行走。然而,目前存在的 ETA 系统并不能帮助盲人很好地认知环境信息,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(背景噪声很小,同时只存在简单物体的时候)中比较适用;但在实际生活环境中,背景复杂,同时障碍物也不规则,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

2 AudioMan 概念模型

AudioMan 是我们提出的 ETA 系统模型,用户通过系统提示来判断当前环境的情况。系统模型如图 1 所示:由单个摄像头、便携式计算机和耳机三部分组成。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。虚线框内所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。

图 1 中,“图像处理”是系统获取环境中信息的重要模块,通过图像处理技术,从摄像头获取的图像中的提取出道路的方向和各种障碍物的位置。“图像声音映射”是利用听标、耳标和可听化^[9],进行环境信息向

^① 基金项目:浙江省自然科学基金(602060),浙江大学城市学院教师科研基金(J52207001)

声音的映射。由模型图可见,AudioMan 运行过程中,用户与系统的交互基于听觉显示来实现,而图像处理是人与环境得以进行顺利交互的技术基础,是声音能有效地显示环境状态的保障。该模块需要滤除原始数据不需要的信息,恢复、增强有用信号,并将得到的表征环境信息的数据,送入图像声音映射模块。因此,如何正确提取环境方向矢量相当重要,它也是 AudioMan 重点研究内容之一。

3 方向矢量提取算法

在生活中,行人道一般有两种不同纹理的砖铺成,其中一种是具有方向特性的盲人砖,AudioMan 系统需要获取的方向矢量就是盲人砖在整幅图像中的位置。

3.1 边缘检测算法

图 2 - (a) 是拍摄到的某一帧图像,大小为 320 × 240 的彩色图。首先须将彩色图转换成 256 色的灰度图,如图 2 - (b)。为了计算方便,将 320 × 240 大小图像的左右两边各截去 32 列,同时向上扩展 16 行,并将补充的像素值设置为“255”,最终调整成大小为 256 × 256 (图 2 - (b))。再采用迭代阈值法^[9],对大小为 256 × 256 灰度图进行二值化处理(图 2 - (c))。二值化处理的目的是剔除图像中的噪声,有助于检测到效果好的边缘。

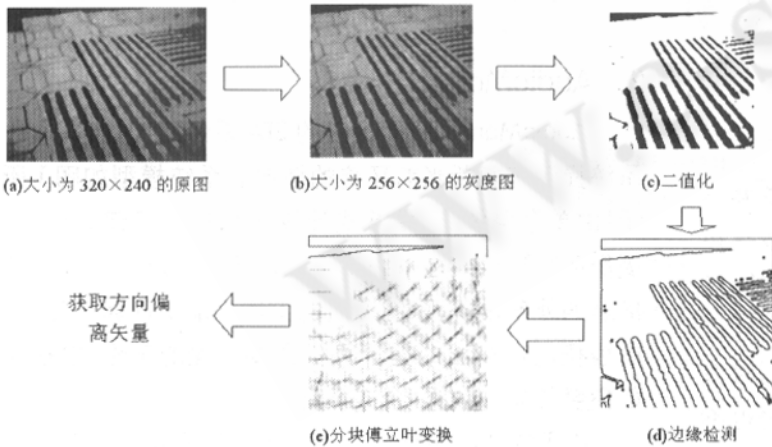


图 2 某一帧图像方向矢量提取示意图

经典的边缘检测方法,利用边缘邻近一阶或二阶方向导数变化规律,来检测边缘。其中 Kirsch 算子^[10]是经典的边缘检测方法。

Kirsch 算子采用 8 个模板,如图 3 所示。

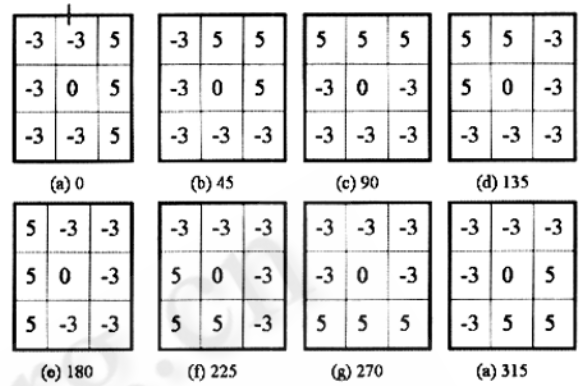


图 3 Kirsch 3 × 3 模板

3.2 计算道路偏转矢量

功率谱含有图像纹理的主要特征,其波峰给出了纹理的空间周期信息。对如图 2 - (d) 所示的方向性强的纹理图像,其功率谱沿着纹理方向相垂直的直线汇聚。设图象 $f(n, m)$ 的离散傅立叶变换为 $F(\omega_1, \omega_2)$, 其功率谱 $P_{ff}(\omega_1, \omega_2)$ 可按式计算:

$$P_{ff}(\omega_1, \omega_2) = |F(\omega_1, \omega_2)|^2 = F(\omega_1, \omega_2) F^*(\omega_1, \omega_2) \quad (2)$$

如果在功率谱 $P_{ff}(r, \phi)$ 描述中使用极坐标:

$$r = \sqrt{\omega_1^2 + \omega_2^2} \quad (3)$$

$$\phi = \arctan \frac{\omega_2}{\omega_1} \quad (4)$$

均值为:

$$P_{\phi}(\phi) = \int_0^{\max} P_{ff}(r, \phi) dr \quad (5)$$

则 $P_{\phi}(\phi)$ 可以较好地反映纹理的方向性。若图像纹理沿 θ 角方向表现处较强的方向性,那么 $P_{\phi}(\phi)$ 在角度 $\phi = \theta + \frac{\pi}{2}$ 处有最大值。

根据这个特性,将大小为 256 × 256 的二值图像分块处理,在对每块子像进行傅立叶变换。子块的大小适中,并能使每个子块所包含一定的纹理特性。系统分割得到 8 × 8 块子像,但只对下面 7 × 8 个子像进行傅立叶变换(在调整图像大小过程中,其中在图像最上层补充了 16 行,因此分割后最上层的 8 个子块图像不需要进行变换)。变换后各子块得到的功率谱如图

2-(e)所示。与图 2-(d)对照可见,功率谱中具有较大的能量所对应的角向与图像纹理的方向角基本垂直;而不规则的纹理特征对应的频率谱的能量却是杂乱而分散的。从而,根据各个子块的频率谱分布,计算各个子块中具有有一定能量的角度,分析得到方向矢量。

由 2、3、4 式,可将功率谱转换成极坐标形式,并计算得到各个角向 ϕ 所具的能量。系统将 $[0, \pi]$ 等分成 16 个方向,计算这 16 个方向各自的能量 $\rho(\phi)$ 。图 4 是图 2-(a)的分块频率谱,第 8 层包含了非环境信息,故不需要进行变换。盲人砖处的能量谱集聚在某一方面上,而在其他地方能量谱比较分散。根据这个特点,系统对每层的 8 个子块的能量谱进行比较,记录具有相同(或相似)能量谱的区域及其位置,再得到相似区域的中心位置,即这一层的盲人行走道路的中心位置。最后得到 7 层道路中心位置,这七个数值组成道路方向矢量,用 $O = (O_1, O_2, \dots, O_7) = (x_1, x_2, \dots, x_7)$ 表示(其中 $0 < x_i < 8$)。

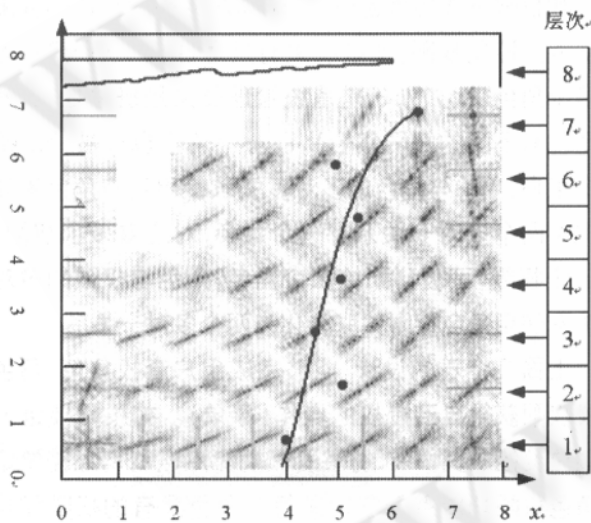


图 4 某一帧图像的分块频率谱

图 4 中黑点是系统按照这种方法处理后得到的各层道路的中心位置,其基本上能表示出当前取得的环境图像中道路的偏离情况。虽然与理想情况还有一点差距,但结果足以为盲人提供当前的道路偏离情况,而且这样的差距并不会给盲人带来严重的方向判断错误。

4 结束语

AudioMan 系统的研究是一项交叉性研究,它将有助于进一步加强工效学领域与各领域专家的联系与合作。AudioMan 改变了现存 ETA 中普遍使用的从图像的象素点到声音直接映射的算法,利用边缘检测算法和道路偏转矢量的计算,提取出环境的方向矢量参数。经过图像处理后的声音映射,是清晰的,有规律的,使用户不经训练或仅通过短时间的学习就能适应。希望本文的工作能为以后设计不同功能和复杂度要求的 ETA 系统提供一定的参考。

参考文献

- 1 D Bolgiano, EJ Meeks. a laser cane for the blind [J], IEEE Journal of Quantum Electronic, 1967, 3(6): 268-268.
- 2 R Fish. auditory display for the blind [P]. USA: 3800082, March 1974.
- 3 Meijer P. Image - Audio transformation system [P]. US A: 5097326, March 1992.
- 4 Bourbakis NG, Kavraki D. An intelligent assistant for navigation of visually impaired people [A]. Proc. of the IEEE 2nd Inter. Symposium on Bioinformatics and Bioengineering Conference [C]. Bethesda: IEEE Press, 2001: 230-235.
- 5 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 徐义东、方志刚、张丽红, 可听化技术综述 [J], 计算机辅助设计和图形学学报, 2004, 16(1): 14-18.
- 7 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 徐义东, 听觉显示开发平台的设计及应用 [硕士学位论文], 杭州: 浙江大学, 2004.
- 9 杨枝灵等, 数字图像获取、处理及实践应用 [M], 北京: 人民邮电出版社, 2003. 4.
- 10 崔岐, 数字图像处理技术与应用 [M], 北京: 电子工业出版社, 1997. 3.