

# AudiMeters: 高维连续数据可听化研究<sup>\*</sup>

方志刚<sup>1</sup> 马卫娟<sup>2</sup>

(浙江大学城市学院 杭州 310015)<sup>1</sup> (浙江大学外语学院 杭州 310027)<sup>2</sup>

**摘要** 高维连续数据的数据-声音映射是可听化应用设计研究的重点和难点。提出了多通道-多维度数据-声音映射模型,支持多通道映射、多维度映射和混合映射,可以充分利用各种声音参数。实现了可听化汽车仪表原型系统 AudiMeters,采用混合映射对转速、时速、水温 and 油量等数据进行可听化显示,验证了多通道-多维度数据-声音映射模型的可行性。

**关键词** 可听化,听觉显示空间,数据-声音映射,映射模型

## AudiMeters: A Case Study of Sonification of Multi-dimension Continuous Data

FANG Zhi-Gang<sup>1</sup> MA Wei-Juan<sup>2</sup>

(Zhejiang University City College, Hangzhou 310015)<sup>1</sup> (Zhejiang University College of Foreign Languages, Hangzhou 310027)<sup>2</sup>

**Abstract** High-dimension data-sound mapping is emphasis and difficulty of sonification research. A kind of multi-channel and multi-dimensional data-sound mapping model, which can make use of sound parameters and supports multi-channel mapping mode, multi-dimensional mapping mode and mixed mapping mode, is presented. A prototype system, AudiMeters, which can map rev, mph, temperature and oil capacity to sound using mixed mapping, is implemented and has validated the feasibility of the data-sound mapping model.

**Keywords** Sonification, Auditory display space, Data-sound mapping, Mapping model

与可视化系统相似,可听化系统主要由数据预处理、数据-声音映射(以下简称数据映射)、声音样本合成、播放声音四部分组成。虽然数据映射也可看作一类信息编码问题,但是数据映射是特殊的信息编码。一般编码问题主要是对物理量所进行的数学变换;而数据映射还涉及到人的心理量,后者取决于人的听觉感知特点。表面上看数据映射是数据变量到声音参数的数学变换,可以精确地变换。但是人耳对声音参数的感知的过程却复杂多变(例如掩蔽效应和鸡尾酒会效应)。因此,数据映射本质上是物理量到心理量的映射,前者可以精确度量,而后者却不能。到目前为止,人们还没有找到一种指导数据映射的理论,数据映射的设计不但是一个复杂的过程,而且还无法从理论上证明设计结果的有效性,只有通过实验方法(即可用性测试)来验证和评价设计结果<sup>[1]</sup>。

在可听化系统设计中,数据映射是整个系统的关键环节,尤其是高维连续数据的可听化映射是研究的重点和难点,目前研究相对较少。国内相关研究有浙江大学 CAD&CG 国家重点实验室 3D 虚拟声音研究、清华大学、同济大学、天津大学等单位在建筑声学仿真和厅堂音质评价等方面的研究等。其中“可听化(Auralization)”<sup>[2]</sup>与本文的可听化(Sonification)意义不同,本文强调的是“数据的听觉显示”。目前,除我们正承担浙江省自然科学基金进行听觉界面研究外,利用非语音声音进行数据显示(即可听化)和听觉界面设计的研究(国内几乎为空白)。本文提出多通道-多维度数据映射模型,该模型可以支持多通道映射、多维度映射和混合映射,可以较为有效地解决高维连续数据的可听化映射。同时开发了可听化汽车仪表原型系统 AudiMeters,作为研究高维连续数据映射问题的系统平台,验证多通道-多维度数据映射模型的有效性。

## 1 数据映射的分类

听觉显示空间(Auditory Display Space)<sup>[3]</sup>是在已有的听觉显示技术基础上提出的概念,包括三个维度:

(1) 数据维度(Data dimensionality)。数据维度是一个物理量的全部属性的集合,每个物理属性对应一个数据维度。

(2) 声音维度(Sound dimensionality):声音维度是指声音的各个属性参数,包括:波形、包络、频率、幅度、空间位置、持续时间和时序等等。声音的各个维度之间不是完全独立的,地位也不等同,因此将声音维度用于数据显示时,需要慎重选取。

(3) 映射维度(Mapping dimensionality):数据映射是指数据属性到声音参数的映射,即特定数据到特定声音的对应关系。主要利用声音的多维度属性,为多维数据域映射到声音多个参数找到合适的函数关系。声音的多维度特性对表达多维数据很有利,但有时也使研究者在选择维度时无所适从。在不同应用中,表征同一个量可能使用了不同的维度。目前还没有标准对维度选取作一些规定,数据映射的设计结果的好坏需要通过可用性测试来评估。

在听觉显示空间的三个维度中,最重要的维度是数据映射。数据映射目前没有统一的分类方法,可以根据所要表达的数据类型、声音参数、声音的方式以及映射函数关系进行分类。根据数据属性和声音属性的函数关系可以分为:线性函数映射、幂函数映射、指数函数映射、阶梯函数映射、分段函数映射等。根据数据属性可以将数据映射分为:

(1) 无结构离散信息映射。这种映射方式是比较简单的点对点的映射,一般不需要研究声音的具体参数值,只要将信

<sup>\*</sup>浙江省自然科学基金资助项目(602060)。

息映射为与之意义关联的声音。声音所代表的意义一般与人们熟悉的或自然的联系相一致。无结构离散信息映射中最简单的是听标<sup>[3]</sup>，一个听标对应一种声音，例如用某种动物的叫声来代表此动物。

(2)结构化离散信息映射。将短的音调序列与某种事件或物体相联系，可以成为耳标<sup>[3]</sup>。耳标可以是任何声音消息，比如自然声音、音符、乐器音、语音以及自己合成的声音数据等等。可以通过改变耳标的声音参数来构造层次化的耳标，层次化耳标适合表现结构化离散信息。结构化离散信息映射的典型应用是菜单的可听化。

(3)高维点数据映射。早期的一个可听化试验<sup>[4]</sup>将7种化学物的变量分别对应于频率、响度、衰减、方位、持续时间和间隙等各声音维度。测试对象为化学专家，他们能理解声音表达的不同模式。训练前区分化学物的准确率为90%；训练后的准确率为98%。实验表明，专家很容易适应复杂的听觉显示。

(4)一维连续数据映射。数学和物理中有很多二维函数图(线图)，线图通俗易懂，给人们的学习带来很多方便。为了盲人教学也能使用函数图，科学家研究用可听化来表现函数图<sup>[5]</sup>，将线图映射为声音；选取声音的某一维度，将函数值随横轴的变化映射为声音参数随时间的变化。例如：直线影射为频率不变的一段声音；正弦曲线映射为频率随时间起伏变化的声音。线图可听化属于一维连续数据映射的范畴。

(5)高维连续数据映射<sup>[6]</sup>。在数据可听化中，目前还没有统一的规则来指导映射设计。声音参数的属性、优劣性有待进一步研究，映射关系是否合适，必须通过系统设计者不断地测试。我们开发的原型系统 AudiMeters 就是针对高维连续数据映射问题的研究。

在听觉显示技术中，听标和耳标分别是无结构离散信息映射和结构化离散信息映射；而可听化主要研究信息和声音之间的连续映射。可听化技术以新颖有用的方法在听觉显示信息方面具有最大的潜力，它能显示多种可能的数据，能处理连续的、多维的数据。

## 2 多通道-多维度数据映射模型

在高维连续数据的数据映射设计中，数据具有多个属性且为连续的；相应地，要求声音具有多维度的参数且能够连续调制。但是人们不能任意选择声音的多维度参数来表达多维数据，原因在于人耳对声音的许多参数是不敏感的。人类工效学告诉我们，在声音的众多可调制参数中，人耳对频率(音调)的变化最为敏感，对幅度(响度)的变化则相差甚远。也就是说，在可听化设计中，人们往往倾向于在几个人耳敏感的声音参数中进行选择。这样带来的问题是可供选择的参数变得非常有限。传统的数据映射方法要求不同的数据属性映射到不同的声音参数，在应用设计中造成了矛盾。为克服上述矛盾，我们提出一种多通道-多维度数据映射模型(如图1)。

多通道-多维度数据映射模型的基本思想是：(1)如果数据维度较低，则优先选择人耳敏感的几个声音参数进行映射，每个声音参数对应一个数据属性，进行声音合成并输出。(2)如果需要也可以将不同的数据属性映射到相同的声音参数，但必须保证在不同的声音参数值区间进行映射，这类似于复调音乐的几个声部。(3)如果数据维度很高，超出了可选择声音参数的上限，则将数据属性加以分组，然后以组为单位分别选择声音参数进行映射并进行声音合成，最后将合成后的声

音叠加起来进行输出。基于这个思想，可以将多通道-多维度数据映射模型分解为三个层次：

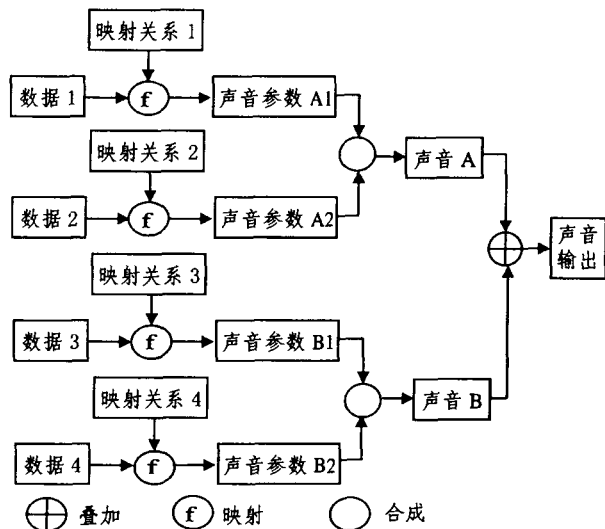


图1 多通道-多维度数据-声音映射模型

(1)多通道映射。多通道声音就是将来自几个独立通道的声音叠加后播放出来。根据声音的“鸡尾酒会效应”，在一个嘈杂的环境中，我们照样能将注意力集中于某个声音上，获得需要的信息。利用人类听觉的这一特性，我们可以提出数据映射的多通道映射模型。简单来说，多通道映射就是将多个数据属性映射到不同通道声音的同一参量，一般选择人耳最敏感的声音参数(如频率或音调)。多通道映射设计的一般步骤为：(a)对应 $n$ 个数据属性，选择人耳容易辨认的 $n$ 种不同声音，一般选择音色不同的声音。(b)选择声音参数，将不同数据属性映射为相应通道的某个声音参数(如频率、幅度或方位等)，各个通道应选择相同的声音参数。因为只利用声音的一个参数，所以一般选取信息表达能力最强的频率或音调。(c)为每一个声音通道选择合适的映射函数关系，如线性、二次、指数等。(d)合成、叠加并输出声音。对于多通道映射，虽然理论上对数据维度的数目没有限制，但随着数据属性的增多，在具体实现中通道之间的干扰会明显增强。

(2)多维度映射。多通道映射只利用了声音的一个维度，可能存在几个声音通道的相互干扰。利用声音的多维度特性，我们可以提出多维度映射模型。在多维度映射中，每个数据属性映射为同一个声音的不同参数。多维映射设计的一般步骤为：(a)对应 $n$ 个数据属性，选择人耳容易辨认的 $n$ 种不同声音参数。(b)将不同数据属性映射为不同的声音参数，一般选取信息表达能力最强的声音参数表达最重要的数据。(c)为每一个映射维度选择合适的映射函数关系。(d)合成并输出声音。类似地，对于多维度映射，虽然理论上对数据维度的数目没有限制，但随着数据属性以及对应声音参数的增多，声音参数之间的干扰(称为交互作用)也会增强，声音参数的选取会变得很难。

(3)混合映射。混合映射就是图1所示的完整模型，它实际上兼有多通道映射和多维度映射两种方式的特征。上面讨论了多通道映射和多维度映射的方法、实现步骤以及各自的不足之处。我们不妨将两者综合起来得到一个混合映射模型。混合映射设计的一般步骤为：(a)对数据属性进行分组，比如属性1到属性 $m$ 为第一组，属性 $m+1$ 到属性 $2m$ 为一组，等等。(b)分组设计，为每组选择一种声音(不同组之间的

声音要容易辨认)并且位于不同通道,然后进行多维度映射子系统的设计。(c)按不同通道合成声音,并将来自不同通道的声音叠加后输出。

很明显,在数据属性比较多的情况下,纯粹的多通道映射存在通道间干扰,纯粹的多维度映射存在声音参数间的交互作用(干扰)。混合映射模型采用折衷的办法可以减少声音之间的干扰,较好地解决高维连续数据的数据映射设计的听觉感知问题。

### 3 AudiMeters:可听化汽车仪表原型系统

#### 3.1 研究意义

在可听化技术研究中,数据映射设计不仅要基于理论模型,还要根据具体应用的实际情况,结合可用性测试来进行研究。我们选择“可听化汽车仪表”AudiMeters作为原型系统加以实现,演示和验证所提出的数据映射模型。汽车驾驶员在驾车过程中需要不断变换视线注视前方、后方及仪表盘,特别是扫视仪表盘会引起重新聚焦,这不仅容易引起视觉疲劳,还存在安全隐患。模拟驾驶舱利用计算机屏幕以图形方式显示汽车各主要仪表,同时利用实时参数化合成声音连续显示汽车工作时的瞬态信息,并在需要时显示视觉告警和听觉告警信息。

#### 3.2 汽车运动模型

简化的汽车模拟运动模型的参数为:

(1)中心位置:以驾驶员眼睛的位置为参考,作为三维场景显示的视点。

(2)方向:控制三维场景显示,由方向盘控制。

(3)加速度:汽车启动、制动、踩油门时设定不同加速度,作为计算速度的依据。加速度变化按照分段函数取值,速度越大,加速度越小。

(4)速度:速度数据被直接显示到速度表上,同时用作温度等仪表取值的参考依据。速度按照牛顿第二定理变化。

(5)时间:时间是温度、油量等仪表取值的依据,同时作为控制方向变化快慢的依据。

#### 3.3 AudiMeters的图形界面

利用OpenGL绘制普通驾驶室和汽车挡风玻璃前方的三维场景。利用ActiveX控件制作绘制包括速度表、转速表、温度表、油量表在内的汽车仪表盘。提供三维图形界面的目的是提高系统演示的真实感。

#### 3.4 AudiMeters的交互方式

系统的交互方式选用市面常见的游戏方向盘,用户可以控制方向、油门、制动、功能键。其中,部分功能键用于改变数据映射方式,其它功能键用于控制转向灯等附属功能。

#### 3.5 AudiMeters主要结构

采用面向对象技术,利用VC++和OpenGL,在听觉显示开发平台ADK(Auditory Display Kit)<sup>[7]</sup>的基础上具体实现了可听化汽车仪表原型系统AudiMeters。系统主要结构如图2。

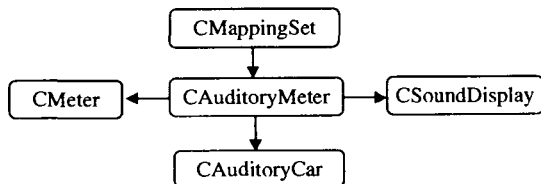


图2 AudiMeter的主要C++类

(1)CAuditoryMeter:实现主程序的界面,负责接受并处理用户(方向盘)输入,控制模拟汽车仪表外观显示和指针转动,并根据仪表数据调用声音显示等功能,可以说是数据中转站。汽车仪表界面中的主要控件有各模拟汽车仪表、各数据显示框及一些功能按钮,如映射关系预置等。

(2)CAuditoryCar:实现汽车驾驶模拟环境,包括三维场景以及控制转向灯等附属功能。

(3)CMeter:实现模拟的汽车仪表,用于存储仪表相关数据,如数据动态范围、当前示数等,同时负责仪表的视觉显示。

(4)CMappingSet:实现数据映射关系设置,是进行数据映射的设计和试验的平台,同时也为系统提供了交互式定制映射规则的能力。CMappingSet是系统的核心类,负责设置和维护汽车仪表系统的数据映射规则,它关系到对可听化实际效果的控制。

(5)CSoundDisplay:实现实时的参数化声音合成,可以提供的参数主要是频率、方位和音量等,可以同时播放多个声音,既可以满足多通道映射,也可以满足多维度映射的要求。

(6)CUabilityTest:为可用性测试实时采集数据预留的接口,目前尚未实现。

### 4 数据映射设计

#### 4.1 数据映射设计原则

数据映射设计的第一步是维度选取。不同的维度适合表现不同的数据类型,设计者首先要研究声音各维度的特性,应该根据可听化系统的数据类型来选择声音维度。比如音色适合表现离散数据,频率可以表现连续数据并且范围较大。

第二步是确定数据与声音的映射关系。设计者要选择用户易于理解的映射,可以考虑用户头脑中传统的、自然的数据映射,还要注意不能令用户厌烦。

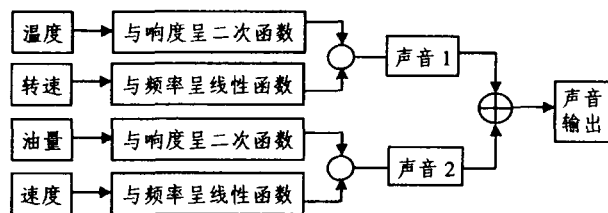
第三步是进行可用性测试。理想情况下,用户会主动适应听觉显示,会学习这一映射。但是,设计者并不一定能把适应性设计得很好,可能导致用户会拒绝适应或学习映射。因此,可听化系统设计中,选取维度和确定映射关系除了应当尽可能多地表现信息,还必须考虑系统使用者的实际情况、喜好程度等主观因素。在条件允许的情况下,多做一些试验和测试,反复比较,使其最好地适应用户和系统使用环境。

#### 4.2 AudiMeters数据映射设计

AudiMeters主要实现速度、发动机转速、水温 and 油量等4个仪表,每个仪表包括5个属性:(1)参数都有起始值和终止值;(2)对应的声音通道;(3)对应声音参数;(4)对应声音参数的起始值和终止值,由于仪表值大小不一定与声音参数值大小成正比,终止值可以比起始值小。比如某些人可能习惯于速度越大、频率越小;(5)映射关系,可以设定为各种映射函数(默认为线性函数)。

数据映射设计基于多通道-多维度数据映射模型,映射方式为混合映射。图3是方案之一,具体考虑是:(1)转速和速度变化较快而且范围大,听觉显示的实时性强,所以让这两个数据与频率作线性映射以充分利用声音频率的表现力。(2)温度和油量变化缓慢,在相当长的时间内往往近似保持不变,所以让这两个数据与响度作二次函数映射。(3)温度和转速在汽车面板上位于同一组,所以安排在同一个通道进行合成与显示,合成声音1;而油量和速度也是一组,合成声音2;最后两个声音叠加在一起输出。(4)用户可以打开或关闭对于某一仪表参数的映射,以突出对重要参数的听觉显示。(5)用

户可以选择不同的映射函数、参数范围等设定,以适应用户的听觉感知特点。



通道 1 映射设置如下:

①温度映射到声音的幅度:输入参数最小值 0,最大值 105,输出参数类型选择幅度,映射函数表达式为  $0.01 * \exp(0.5 * x)$ ;  $A = 0.01 * \exp(0.5 * T)$ ;  $A$  为声音 1 的幅度,  $T$  为温度。

②转速映射到频率:输入参数最小值 0,最大值 7000 (rpm),输出参数类型频率。映射函数表达式  $x+50$ ;  $F = R + 50$ ;  $F$  为声音 1 的频率,  $R$  为转速。

通道 2 映射设置如下:

①油量映射到声音的幅度:输入参数最小值 0,最大值 4,输出参数类型选择幅度,映射函数表达式为  $0.01 * \exp(1.2 * x)$ ;  $A = 0.01 * \exp(1.2 * G)$ ;  $A$  为声音 2 的幅度,  $G$  为油量。

②速度映射到频率:输入参数序号 1,最小值 0,最大值 200,输出参数类型频率,映射函数表达式  $20 * x + 50$ 。  $F = 20 * V + 50$ ;  $F$  为声音 2 的频率,  $V$  为速度。

### 4.3 AudiMeters 实验评价

可听化应用设计能否成功,最大的不确定因素就是数据映射。因为映射不单纯是一个信息编码问题,还涉及到用户的听觉感知的心理学问题,即复杂的人因(Human Factors)问题。成功的可听化设计不能仅靠设计者主观决定,需要进行严格的可用性测试。图 3 所示为通过反复试验后确定的初步

设计方案之一,这一方案已经通过了非正式的实验评估,在实验室环境下证明是可接受的<sup>[7]</sup>。系统运行表明,在两通道四维线索的声场中,用户能优先辨别速度和转速的变化。同时,不同声音线索之间的干扰是存在的,但是用户可以通过关闭对于某些仪表参数的映射来消除或减轻干扰。

结束语 AudiMeters 作为一个实验平台进行数据映射探索和试验,目前其技术层面的工作基本完成,可以进行实时的演示,并可以设定不同的数据映射方案和参数设定。下一步的工作是实现可用性测试的数据采集和处理接口,并进行严格的可用性测试和工效学研究。

### 参考文献

- 1 Kramer G, et al. Sonification Report: Status of the Field and Research Agenda. See <http://www.icad.org/websiteV2.0/References/nsf.html>.
- 2 张琼,宋哲,石教英.可听化技术综述[J].计算机辅助设计与图形学学报,1999,11(2):189~192
- 3 Brewster S A. Providing a Structured Method for Integrating Non-speech Audio into Human-computer Interfaces[D]; [PhD Thesis]. University of York, UK, 1994
- 4 Bly S. Multivariate Data Mappings. In: Kramer G ed. Auditory Display: Sonification, Audification and Auditory interfaces Reading, MA. USA: Addison-Wesley Publishing Company, 1994. 405~416
- 5 Mansur D L, Blattner M M, Joy K I. Sound-Graphs: A Numerical Data Analysis Method for the Blind. J. Med. Sys., 1985, 9: 163~174
- 6 Yeung E S. Pattern Recognition by Audio Representation of Multivariate Analytical Data. Anal. Chem. [J], 1980, 52: 1120~1123
- 7 徐义东. 听觉显示开发平台的设计及应用[D]; [浙江大学硕士论文]. 2004. 3

(上接第 110 页)

收敛的性质,当  $K$ 、 $n$  和  $D$  增长时,  $getNextCP(Q, R_i)$  调用次数增长并不明显,因此具有较好的伸缩性能。

由于最小距离聚集查询可以看作一种特殊的多路空间连接<sup>[11]</sup>,下一步的研究工作将把多路连接方面的研究成果扩展到最小距离聚集查询处理以进一步提高查询处理效率。

### 参考文献

- 1 Roussopoulos N, Kelley S, Vincent F. Nearest Neighbor Queries. In: Proc. ACM SIGMOD, 1995. 71~79
- 2 Papadopoulos A N, Manolopoulos Y. Performance of Nearest Neighbor Queries in R Trees. In: Proc. 6th Int. Conf. on Database Theory (ICDT'97), 1997. 394~408
- 3 Ferhatosmanoglu H, Stanoi I, Agrawal D, Abbadi A. Constrained Nearest Neighbor Queries. SSTD, 2001
- 4 Mamoulis N, Papadias D, Tao Y. All-Nearest-Neighbors Queries in Spatial Databases. In: IEEE Conf. on Scientific and Statistical

Database Management (SSDBM), 2004

- 5 Hjalason G, Samet H. Incremental Distance Join Algorithms for Spatial Databases. SIGMOD, 1998
- 6 Corral A, Manolopoulos Y, Theodoridis Y, Vassilakopoulos M. Closest Pair Queries in Spatial Databases. SIGMOD, 2000
- 7 Guttman. R-trees: A Dynamic Index Structure for Spatial Searching. SIGMOD, 1984
- 8 Beckmann N, Kriegel H P, Schneider R, Seeger B. The R\*-tree: an Efficient and Robust Access Method for Points and Rectangles. SIGMOD, 1990
- 9 Carey M J, Kossmann D. On saying "enough already!" in SQL. In: Proc. of ACM SIGMOD, 1997. 219~230
- 10 Fagin, Lotem R, Naor A. Optimal Aggregation Algorithms for Middleware. PODS, 2001
- 11 Mamoulis P N, Theodoridis Y. Processing and Optimization of Multi-way Spatial Joins Using R-trees. In: Proc. 18th ACM PODS Conf. 1999. 44~55