

产品声音的交互式聚类设计研究

张扬, 陈文颖, 皮珊, 丁胜年
(宁波财经学院, 浙江 宁波 315175)

摘要: **目的** 声音是产品和用户之间的一种沟通媒介, 为了增进设计师对产品声音的理解、合成与设计匹配, 提出一种交互式可视化产品声音数据聚类分析框架。**方法** 首先通过神经网络将设计师感官描述式信息与声音的特征参数进行融合嵌套; 其次基于高斯混合模型来描述非线性几何分布的产品声音数据; 最后设计师输入个人先验知识经验参与交互聚类。**结果** 基于Python的Anaconda3包开发了产品声音交互式聚类的可视化分析实验工具, 得到最优化产品声音聚类结果。**结论** 该产品声音交互聚类可视化分析工具融合了声音技术参数和人脑听觉反应机制, 在聚类过程中允许用户参与交互并融入用户的先验知识, 并行视图可以实时显示数据元素的流向和判别类别的稳定性。同时, 可视化分析可以帮助用户横向比较各聚类结果的异同, 样本的比例分布与合理性, 以期寻求最优聚类结果。

关键词: 产品声音; 融合嵌套; 交互式聚类; 可视化分析

中图分类号: TB472 **文献标识码:** A **文章编号:** 1001-3563(2023)08-0115-08

DOI: 10.19554/j.cnki.1001-3563.2023.08.011

Interactive Clustering Design of Product Sound

ZHANG Yang, CHEN Wen-ying, PI Shan, DING Sheng-nian
(Ningbo University of Finance & Economic, Zhejiang Ningbo 315175, China)

ABSTRACT: In view of that the sound is a communication medium between product and user, the work aims to propose an interactive visual analysis framework of product sound data to improve the designers' capability of understanding, synthesizing, designing and matching the product sound. Firstly, the sensory description information of designers was fused and integrated with the characteristic parameters of sound through neural network. Secondly, gaussian mixture model was used to describe the product sound data in nonlinear geometric distribution. Finally, the designers input personal prior knowledge and experience to participate in interactive clustering. Based on Anaconda3 of Python, a visual analysis experimental tool for interactive product sound clustering was developed, and the optimal product sound clustering results were obtained. The visual analysis tool for interactive clustering of product sound combines the technical parameters of sound and the auditory reaction mechanism of human brain, allowing users to participate in interaction and integrate their prior knowledge in the clustering process. The parallel view can display the flow direction of data elements and judge the stability of categories in real time. At the same time, visual analysis can help users to compare the similarities and differences of clustering results horizontally, the proportional distribution and rationality of samples, in order to seek the best clustering results.

KEY WORDS: product sound; fuse and integrate; interactive clustering; visualization analysis

人类通过耳朵接受声音而感知外界事物和意义。黑格尔在《美学》中就有“艺术作为感性事物, 只关涉视听两种认识性感觉”的论述, 可见声音与图像一

样具有重要的研究价值与意义。在产品设计领域, 通常基于信息学角度将产品声音划分为两类: 一类是信息提示和警告声, 如按键的反馈声, 洗衣机的滚筒翻

收稿日期: 2022-11-22

基金项目: 宁波市哲社规划课题 (G21-3-ZX75); 浙江省基础公益研究计划项目 (LGG19E050004)

作者简介: 张扬 (1980—), 男, 博士, 副教授, 主要研究方向为人机工学, 产品创新设计。

转声,汽车倒车的“嘟嘟”声等,从中用户可以感知到产品的信息交互反馈、运作状态及警示^[1];还有一类是引导和解释声,如汽车导航、语音机器人等智能终端就是利用了语音的亲合力给予用户温馨提示,构成了用户体验过程中可辨识的引导“线索”。产品设计师在选择上述产品声音时通常采用的是感官描述法^[2-3],即借助形容词来描述声音带给用户的直接感官刺激,以及与视觉、触觉、情绪等一起产生的联动生理与心理感知。其中代表性的声音直接描述词有“咔嚓”“滴答”“飕飕”等,联动感官描述词有“轻快”“愉悦”“紧张”“机械”等。设计师借助感官描述法和相关符号理论为交互界面的图标、产品的实体按键匹配上合适的声音,如洗衣机操控界面的产品声音设计^[4]。

除此以外,产品设计师还可凭借听觉所带来的“感官联动效应”^[5],利用声音的感觉意象作用,挖掘用户内心需求开展设计,如眼镜框架形态设计^[6],真空吸尘器的色彩和尺寸设计^[7]。从前期文献来看,由于实验所用的声音收集来源较为随机,主观描述词汇的有限性和模糊性,难以对各类声音进行更为准确地识别和归类,这在一定程度上影响了后续的相关设计的有效性。声学工程师认为声音是一种振动波,是各种频率和强度正弦波的叠加,可以利用傅里叶函数进行分解与变换,也可利用各种参数进行技术表征,常见的有基频、能量、过零率和梅尔频率倒谱系数等^[8]。然而在产品设计中用到的声音既有纯音,也有复合音,有时声音的表征参数接近,可能带给用户的听觉刺激和生理、心理反应截然不同,而且声音表征参数众多带来的高维数据,将会增加声音信息表达的复杂性和声音分类的准确性。

声音是产品和用户之间的一种沟通媒介,聚类分析是研究声音的重要方法之一。为了增进产品设计师对声音的理解、合成及与设计符号的有效匹配性,提出了一种交互式可视化产品声音数据聚类分析框架。该方法融合了声音技术参数和人脑听觉反应机制,并且交互式聚类过程能让设计师参与其中,输入个人先验知识经验,得到更符合预期需求的最优结果;过程与结果的可视化表达,有助于设计师对产品声音及其媒介沟通机制的理解,提升设计的有效性。

1 声音数据的分析

1.1 感官描述式

早期产品设计师在分析声音时主要聚焦于声音与感知意象关系的映射上^[9]。通过招募专家与用户来对声音进行感官与情感描述,或者通过文献搜寻声音的相关描述词,经统计筛选后得到若干组简练且有差异性的描述词汇,以此作为衡量维度对声音进行分类。

参照上述思路,通过两种途径收集声音资源:从 Soundsnap、Sound Ideas、声音网等国内外网站下载音效资源;作者团队前期数字合成。本次实验共收集150种音效资源,时长在2~10 s,内容与产品声音相关度较高,涉及的产品功能操作行为有开、关、旋转、按压等,发声场所有户外、办公室、厨房等,时间上有长短快慢、节奏重复等变化,结合前期文献,整理出可供选择的属性描述,见表1。

表1 产品声音相关的属性描述列表
Tab.1 Attribute description list of product sound

功能行为	发声场所	时间属性	发声源	感情属性		
打开	户外	短	电子	紧张	愉悦	含蓄
关闭	办公室	长	机械	轻松	压抑	突兀
旋转	厨房	快	家居	细腻	急剧	熟悉
按压	浴室	慢	环境	粗犷	迟钝	陌生
倾倒	卧室	不变	其他	机械	尖利	宏大
吹风	公共场所	重复		人工	低沉	微软
		规则		刚硬	柔弱	

在此基础上,邀请32名自述听力正常的用户(2名声音专家与30名工业设计专业学生)参与本次实验。其中,参与实验的1名专家用户来自某汽车厂商从业4年的音效设计师,1名专家用户来自作者校内人机交互实验室工业设计专业的教师,年龄: $M=32.100$, $SD=2.012$;30名工业设计专业学生为作者校内大二、大三在校生,年龄: $M=20.865$, $SD=3.652$,其中女性18名。在音效类别上,32名用户经投票认为,按发声源属性类别来划分更符合现代产品主题。在具体类别的细分上,按得票数量的多少划分为电子警报类、机械音效类、空气音效类、液体音效类4大类(簇),记为 $X_1=\{R'_s\}$ 其中, $s=1,2,\dots,150$ 为音效资源序号, $t=1,2,3,4$ 为音效所属类别。

1.2 特征参数式

近年来,各种传感器技术和算法被用于检测声音事件与发生场景的分类。如在Dcase2020挑战赛中,Task3就是音频事件检测和定位,要求在10 h的音频数据集中检测出包括警报、婴儿哭啼等14个声音类别。研究人员相继开发出的各种算法最优模型中,包括随机森林(RF)模型、图神经网络(GNN)模型,这几种模型都结合了音频的特征信息。

为了克服音频特征提取方法上的鲁棒性等问题,论文采用了基于听觉谱的矩特征向量法^[10]。该方法不仅抗干扰能力强,其原理更符合人耳对音效的处理过程,相关的音频特征参数分布如下:

$$t_1 = \arg \max(P_{f_i}) \quad (1)$$

$$t_2 = \bar{f} = \sum_i f_i \times P_{f_i} / P \quad (2)$$

$$t_3 = \sqrt{\sum_i f_i^2 \times P_{f_i} / P} \quad (3)$$

$$t_4 = \sqrt{\sum_i (f_i - \bar{f})^3 \times P_{f_i} / B^3 P} \quad (4)$$

$$t_5 = \sqrt{\sum_i (f_i - \bar{f})^4 \times P_{f_i} / B^4 P} \quad (5)$$

$$P = \sum_i P_{f_i} \quad (6)$$

其中, $t_1 \sim t_5$ 分别表示听觉谱的最大幅值频率、中心位置、中心均方根的带宽、中心的谱三阶中心距及谱四阶中心距, f_i 和 P_{f_i} 分别为第 i 个等效矩形带宽率及对应的幅值。由于各音频特征参数之间存在一定的差异性, 需要对其进行数据的归一化。

至此, 实验中的 150 种音效资源均可用一组数值在 0~1 的 5 维的向量来表示:

$$\mathbf{T}_s = [t_1, t_2, t_3, t_4, t_5] \quad (7)$$

$$\mathbf{X}_2 = \{\mathbf{T}_s\} \quad s = 1, 2, \dots, 150 \quad (8)$$

1.3 数据降维与可视化展示

高维的特征参数 $\mathbf{X}_2$ 一方面可以构建出更为准确的信息表达式, 但同时也会增加计算负载和可视化难度。因此, 采用神经网络算法对其进行降维处理, 并联合感官描述式聚类数据 $\mathbf{X}_1$ 进行可视化展示^[11]。相关算法思路如下。

Step1: 构建 4 层结构神经网络 (见图 1), 其中输入层为某音频高维特征参数 $\mathbf{X}_2$, 神经元节点数为 5; 输出层为该音频所属类别 $\mathbf{X}_1$, 共有 4 个神经元节点, 即该音频的感官描述式分类结果与感情属性。

隐藏层 H_1 神经元节点数为 3, 隐藏层 H_2 为待学习的数据嵌套, 即为该音频寻求合适的平面坐标点 (x, y) 。因此, 利用该神经网络框架进行数据训练, 提取得到的隐藏层 H_2 , 其数值能有效融合该音频的感官描述信息和特征参数信息。

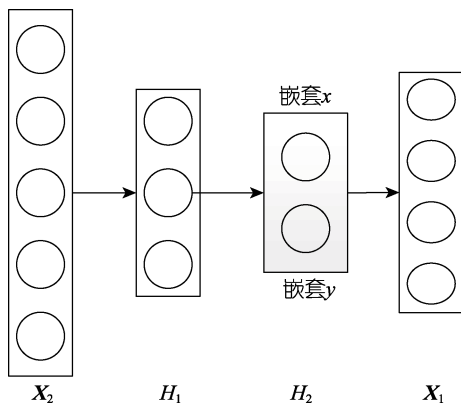


图 1 基于神经网络的产品声音降维嵌套框架

Fig.1 Nested framework of product sound dimensionality reduction based on Neural Network

Step2: 隐藏层 H_1 和输出层 $\mathbf{X}_1$ 分别采用 ReLu 型和 Sigmoid 型激活函数, 隐藏层 H_2 使用线性函数, 并采用交叉熵损失函数为其目标函数进行参数训练

$$H(f, \mathbf{X}) = \arg \min_f - \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^S X_{t,s} \lg[f_s(X_t)] \quad (9)$$

式 (9) 中, f 为待训神经网络模型, $f_s(X_t)$ 为输出层音频 $\mathbf{S}$ 的神经网络模型预测值, $X_{t,s}$ 则为输出层音频 $\mathbf{S}$ 的真实值。

Step3: 将前期已经感官描述式聚类数据 150 个声音样本的特征参数 (数量共 750 个) 作为输入, 以此来确定神经网络模型的阈值和权值。训练采用 Adam 优化器, 在训练了 85 轮迭代后, 神经网络模型的误差率基本接近于 0。一共训练了 120 轮后, 提取了神经网络模型的隐藏层 H_2 数值 (嵌套 x , 嵌套 y)。

感官描述式聚类数据 $\mathbf{X}_1$ 结合隐藏层 H_2 数值 (嵌套 x , 嵌套 y) 即可得到产品声音初始分布图 (见图 2)。从图 2 的 150 种产品声音初始聚类结果可知, 虽然在数据上已经融合了感官描述信息和音频特征信息, 但仍存在可进一步优化的空间, 比如空气类音效、液体类音效和机械类音效很多元素相互交织在一起。这种聚类边缘的不清晰性, 促使人们去探索更优的聚类方法。

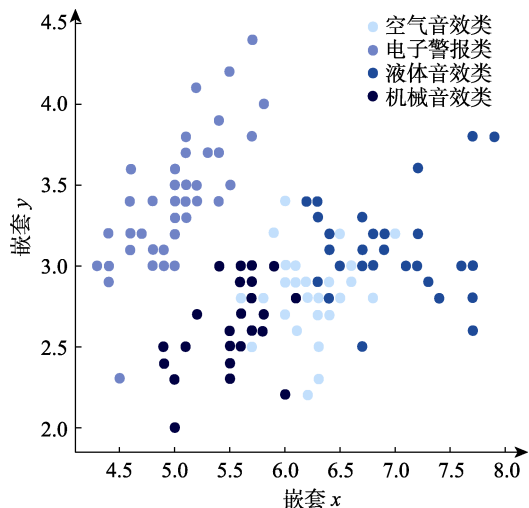


图 2 产品声音嵌套的初始分布

Fig.2 Initial distribution of product sound nesting

2 声音数据的交互式聚类算法

考虑到产品声音初始聚类结果现状及设计师用户对产品声音的特殊需求, 在接下来的交互式优化聚类过程中赋予用户 (设计师) 更高的自由度, 即用户可以根据自身需求, 自由调节声音类别数量, 以及各类别之间的关联性等因素, 并在多次迭代后得到优化结果。同时, 用户可以对多次聚类结果进行比较, 这不仅可以加深对声音及其类别的理解, 也可以找到更符合预期需求的最佳结果。

2.1 高斯混合模型

与 K-means 聚类算法基于元素之间的欧式距离

判断^[11],且只对圆形或椭圆形的元素分布形状敏感相比较,高斯混合模型(GMM)假设元素是一种混合的高斯分布,并将元素软分配到不同聚类类别,适用于任何形状的元素分布^[12],包括实验中嵌套呈现的非线性几何分布的产品声音。因此,论文将基于高斯混合模型来设计交互式聚类算法。

假设用户需要将所用产品声音分成 K 类,那么每类都将是一个高斯分布,且可以由三个参数来描述,分别是 u_k 、 Σ_k 和 $N(x|u_k, \Sigma_k)$ 。其中 u_k 、 Σ_k 分别为第 k 类分布的2维均值向量和 2×2 协方差矩阵, $N(x|u_k, \Sigma_k)$ 为该分布的概率密度函数,可以表示如下:

$$N(x|u_k, \Sigma_k) = \frac{1}{2\pi \times \det(\sqrt{\Sigma_k})} e^{-\frac{1}{2}(x-u_k)^T \Sigma_k^{-1} (x-u_k)} \quad (10)$$

因此,将某声音元素 x_n 通过GMM软聚类到第 k 类分布的概率可以表示如下:

$$p(k|x_n) = \frac{N(x_n|u_k, \Sigma_k)p(k)}{p(x_n)} \quad (11)$$

式(11)中, $N(x|u_k, \Sigma_k)$ 为条件概率函数, $p(k)$ 为先验概率函数, $p(x_n)$ 为观测密度函数。在三个参数 u_k 、 Σ_k 和 $N(x|u_k, \Sigma_k)$ 无法直接找到的情况下,可以通过需求模型的最大似然估计来求解^[13],即构建目标函数和迭代优化(EM算法)。

2.2 列联表函数的构建

用户在对产品声音进行 K 类分类的基础上,若需根据聚类结果进行 K' 再分类,可以提供一个 $K \times K'$ 的目标列联表^[14]。该目标列联表融合了用户先验知识经验,并可将其转化成新分类数据来源概率。

以 $K=4$ 分类结果为例,第一次聚类GMM₁结果中机械类、空气类、液体类三类音效元素之间存在一定的交织现象,用户想构建一组 $K'=5$ 的新聚类GMM₂,因此需构建一组 4×5 的目标列联表,见图3。由图3可知,新的第5类其元素主要来自原来第1(机械类)、3(空气类)和4(液体类)类。在此基础上,需要构建新第5类数据来源的比例概率,论文采用的是平均等比分配的方式,即在新的分类中,第1~4类均来自原来的第1~4类,第5类来自第1、3和4类,比例各占1/3,由此构成列视图分布矩阵 J_X ,见式(12)。

图3 目标列联表

Fig.3 Target contingency table

将矩阵 J_X 规范化后,得到行视图矩阵 $J_{X'}$,即式(13)。在式(13)中以第一行为例,原第1类减少的1/3元素转至新第5类。论文采用的数据来源等比概率,经过多次迭代后会随产品音效特征、目前聚类及目标函数逐渐优化,这在以往文献和前期实验中都得到了验证^[14]。

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0.33 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0.33 \\ 0 & 0 & 0 & 1 & 0.33 \end{bmatrix} \quad (12)$$

$$\begin{bmatrix} 0.67 & 0 & 0 & 0 & 0.33 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.67 & 0 & 0.33 \\ 0 & 0 & 0 & 0.67 & 0.33 \end{bmatrix} \quad (13)$$

同样,条件概率的列联表也是由行视图分布矩阵和列视图分布矩阵来表示:

$$\phi_i(j) = \frac{W_{ij}}{\sum W_{ij}} \quad (14)$$

$$\phi_j(i) = \frac{W_{ij}}{\sum W_j} \quad (15)$$

$$W_{ij} = \sum_{n=1}^N P(i|x_n) \cdot P(j|x_n) \quad (16)$$

式(13)~(14)分别为条件概率的行视图分布和列视图分布, $i \in [1, 2 \dots K]$, $j \in [1, 2 \dots K']$,式(15) W_{ij} 为前原聚类属于 i 类与新聚类属于 j 类的乘积。

2.3 期望函数与优化

论文的期望函数是基于目标列联表和条件概率列联表之间的相对熵来(D_{KL})构建的,即期望步骤(E-step)具体公式如下:

$$E = \frac{1}{K} \sum_{i=1}^K D_{KL}(\phi_i \| J_{X'}(i,:)) + \frac{1}{K'} \sum_{j=1}^{K'} D_{KL}(\phi_j \| J_X(:,j)) - \frac{1}{N} \sum_{n=1}^N D_{KL}(w_1(x_n) \| v(1/K)) - \frac{1}{N} \sum_{n=1}^N D_{KL}(w_2(x_n) \| v(1/K')) \quad (17)$$

式(17)中,为 $w_1(x_n)$ 产品音效元素 x_n 第一次聚类中归属于各类别的概率向量,同理 $w_2(x_n)$ 为新聚类中的归属概率向量; $v(1/K)$ 和 $v(1/K')$ 分别表示前后聚类中各类别的分布概率。其中, $J_{X'}$ 、 J_X 和 $\sum_{n=1}^N D_{KL}(w_1(x_n) \| v(1/K))$ 为常量。

期望函数的优化是将其值极值化,主要包括将目标列联表(J_X 和 $J_{X'}$)和条件概率列联表(ϕ_i 和 ϕ_j)之间的差距最小化, $w_1(x_n)$ 与 $v(1/K)$ 、 $w_2(x_n)$ 与 $v(1/K')$ 之间的差距最大化,以此多次迭代后来吻合用户预期,即M-step。

3 实验及其结果分析

产品声音交互式聚类的可视化分析实验基于 Python 的 Anaconda3 包开发,其中交互式可视化界面基于 Tkinter 库,声音的特征参数基于 Librosa 库,神经网络的数据嵌套基于 Tensorflow 和 Keras 库来实现。在此基础上,实验部分侧重于感官描述与特征参数嵌合后的产品声音的聚类实验与评估验证。其中,在测试人员安排上邀请了参与感官描述式测试的 1 名声音专家与 5 名学生,年龄: $M=23.165$, $SD=3.812$,其中女性 3 名。测试人员接受了 5 min 的使用界面交互教学和 10 min 的自由练习。

3.1 最优分组聚类实验

本阶段实验小组的目标是寻求产品声音的最佳分组。小组成员经协商认为最佳分组的标准应该包括:类别内具有相似性,类别间有差异性且边缘清晰,整体分类效果能吻合感官判别。

实验将图 2 产品声音初始分布数据导入后,在初始聚类 GMM_1 中设定类别数 $K=4$,聚类结果见图 4。从图 4 可知,电子警报类声音(类别 2)与其他 3 类差异性大,仍单独归类;同时,与初始分布(见图 2)相比较,类别 1 和类别 3 在视觉上能较好地区分,但是图 4 方框标红处类别 3 和 4 之间边缘的清晰性较差,通过聆听比较发现个别元素归类不够准确。

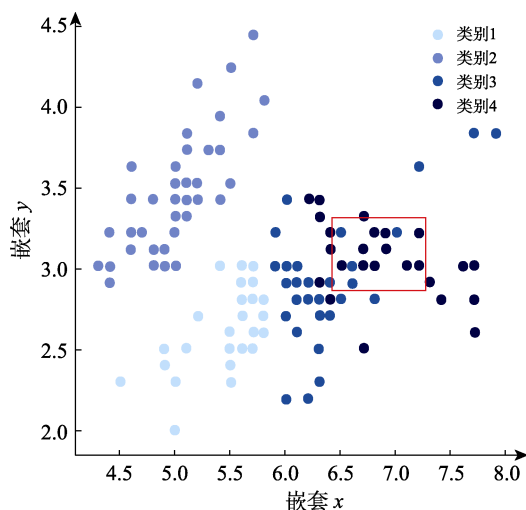


图 4 四类初始聚类 (GMM_1)
Fig.4 Initial clustering of four kinds(GMM_1)

在此基础上,为了进一步探寻更优聚类,实验在第二次聚类 GMM_2 中减少了类别数,设定 $K'=3$ 。由其聚类结果(见图 5)可知,机械类音效和空气类音效在频率上较为接近,算法将两者大部分元素合并成一类,即图 5 中的类别 1。然而该分类仍然未能将类别 1 和类别 3(原液体类音效)恰当归类,在视觉上存在一定的交织现象,在感官通过产品元素的比较聆听,也没能取得较好的聚类结果,有待于进一步探索

产品声音的归类。

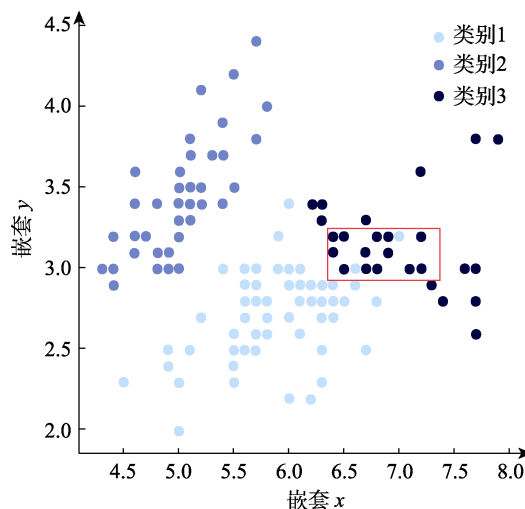


图 5 三类二次聚类 (GMM_2)
Fig.5 Quadratic clustering of three kinds(GMM_2)

进而实验增加二次聚类 GMM_2 的类别数,即设定 $K'=5$ 。从其聚类结果(见图 6)可发现该分类有几点较为明显的特征:从初始分布的空气音效和液体音效中分离产生一个新分类,即图 6 中的类别 5;5 个类别在视觉上边缘较为清晰。

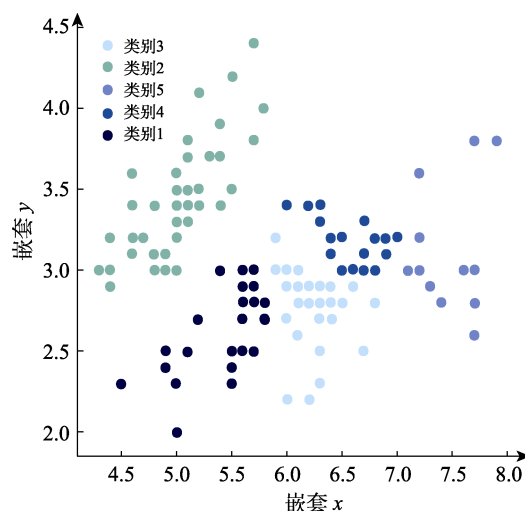


图 6 五类二次聚类 (GMM_2)
Fig.6 Quadratic clustering of five kinds(GMM_2)

为了进一步验证 $K'=5$ 分类结果中新增类别的正确性,首先在目标列联表中设置新增类别的源数据来源,论文分别选择来自图 2 原 1~4 类, 2~4 类和 3~4 类 3 类数据,结果显示数据来源为原 3~4 类时聚类效果最佳。再者,通过聆听发现,原始分布中的空气类音效和液体类音效确实在感官上有些相似之处。

综上所述,测试人员认为 $K'=5$, $K=4$,且新增类别来源于空气类音效和液体类音效时分类效果最佳。同时,实验小组认为,该工具在界面操作上比较简单,可以在线聆听和比较音效;算法上能嵌合感官评价和

特征参数,能通过目标列联表灵活选择新聚类数据来源,并能多次比较聚类结果,从而选择最优结果。

3.2 聚类的可视化分析

聚类的可视化分析主要分成两部分,一是聚类的过程和结果的可视化,包括数据元素流向、类别比例的可视化,以及聚类结果的可视化;二是交互操作的可视化。

采用数据流向并行视图来直观地显示初始聚类 GMM_1 和二次聚类 GMM_2 过程中产品声音元素的流动趋向,见图7。由图7可知,在最优分组聚类过程中,电子警报类音效(初始类别2)由于声音的特殊性,不因聚类数量的变化而变化,始终保持“类别2→类别2”的流动趋向;机械类音效(初始类别1)在3大类聚类过程中先和空气类(初始类别3)合并,再在5大类聚类过程中又独立开来,即“类别1→类别1(新合并)→类别1”;空气类音效(初始类别3)先和机械类音效(初始类别1)合并,再在5大类聚类过程中部分独立,部分和液体类音效形成新的分类,即“类别3→类别1(新合并)→类别3(或新类别5)”;同样,液体类音效(初始类别4)元素的流动趋向是“类别4→类别1(新)→类别4(或新类别5)”。

同时,采用旭日图的形式来表示聚类的元素类别比例,见图8。在图8中由内向外的3个环形分布

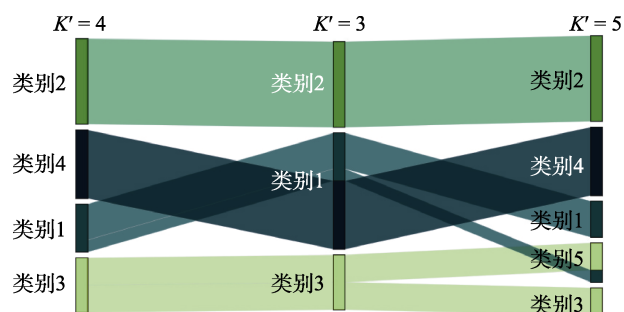


图7 数据流向并行视图
Fig.7 Parallel view of data flow

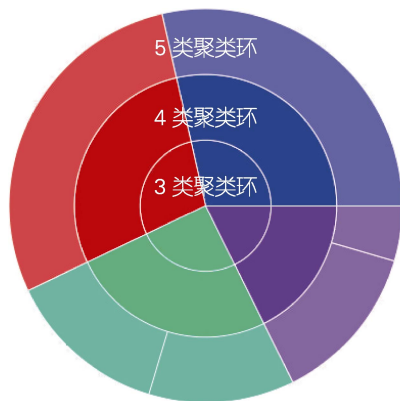


图8 聚类比例旭日图
Fig.8 Sunrise chart of clustering scale

分别是 GMM_1 四大类、 GMM_2 三大类和 GMM_2 五大类的产品元素比例,还可通过鼠标交互操作显示某一类音效在不同聚类过程中的流动趋向及其相关比例。

在交互界面的设计上,如图9所示有6个直接显示视图和1个弹出式显示视图(目标列表视图)。整个界面的操作逻辑可以分成以下步骤:

Step1: 从左上角的文件导入产品声音的感官描述和特征参数数据,经后台神经网络算法嵌套生成数据的初始分布图。

Step2: 根据系统后台得出的产品声音初始分布图,结合控制台上声音元素的聆听与比较,用户可以根据需要通过鼠标拉伸重新选择聚类数(初始聚类 GMM_1)。

Step3: 根据初始聚类的结果,用户可以进一步进行最优化聚类 GMM_2 ,这其中需要构建目标列联表和选择合适的聚类数。最优化聚类 GMM_2 可以进行多次,直至符合用户预期。

Step4: 在多次最优化聚类 GMM_2 过程中,可以通过并行视图来判断聚类过程中的数据流向,旭日图可以用来进一步判断聚类过程中类别数据的整体比例和局部比例(可通过鼠标交互操作得到)。

测试人员通过操作认为,该产品声音交互聚类可视化分析工具在聚类过程中允许用户参与交互并融入用户的先验知识,并行视图可以实时显示数据元素的流向和判别类别的稳定性。同时,可视化分析可以帮助用户横向比较各聚类结果的异同、样本的比例分布与合理性,寻求最优结果。

3.3 设计应用案例

为了进一步验证论文所提出的产品声音聚类方法的有效性,课题组将为一款智能中药煎煮机操作界面的图标匹配合适的音效,见图10。该款产品主要面向注重生活品质和美学感官的中老年用户群体。

首先,在图6五类产品声音二次聚类的结果中进行选择。这与以往直接在150种产品音效中进行海选相比,极大地节省了时间。在大类别的选择上,课题组成员主要倾向于类别1和类别2的音效。其中类别1的音效主要是机械类音效,类别2为电子类音效。考虑到老年用户视力衰退和思维固化等特性,最后采用的是机械类音效,即机械运作音效。

再者,结合表1为“启动/取消”“功能”“热烘”和“清洗”4个产品图标进行音效的选择匹配设计,见表2。以“启动”图标为例,功能行为为“打开”或“启动”某一功能,所选择的是一段3s“起音向上”的产品音效,标志着机器开始进入某一工作状态。当然匹配的过程也可以基于产品声音交互聚类可视化分析平台,但在小样本声音聚类过程中所花的时间会比较大。

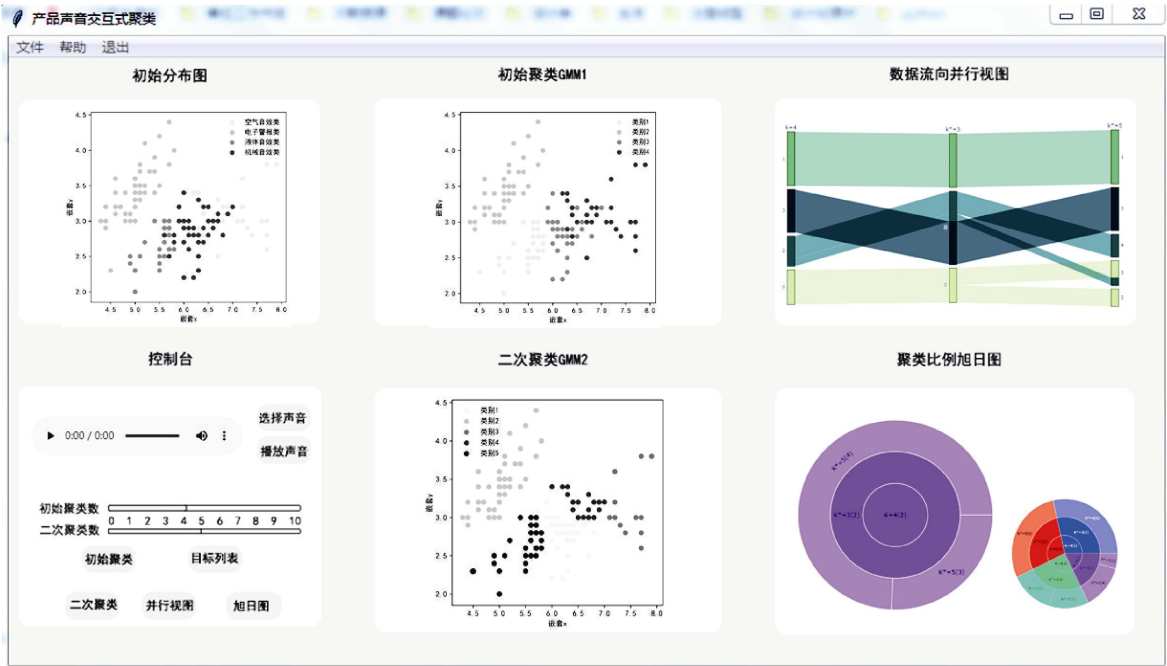


图 9 产品声音交互聚类可视化分析界面
Fig.9 Visual analysis interface of product voice interactive clustering



图 10 智能中药煎煮机及其操作界面
Fig.10 Intelligent traditional Chinese medicine extracting machine and its user interface

表 2 界面图标与音效的匹配

Tab.2 Matching of product interface icon and sound effect

图标名称	功能行为	操作方式	音效特性
启动/取消	打开/关闭	长按	起音向上/尾音向下
功能	功能选择	短按	平声短音
热烘	吹风	短按	风机吹动
清洗	旋转	短按	高速旋转

最后, 课题组招募了 10 名老年用户为音效匹配结果进行聆听与答题, 用户年龄: $M=63.865$, $SD=5.600$, 其中女性用户 3 名。论文采用的是 SUS 可用量表来对音效的可用性、匹配性和满意度进行打分, 共计 10 个问题, 测试用户需要在“非常同意”到“非常不同意”

5 个等级中进行选择答题。测试用户答题分值总和乘以增量常数 2.5 即为最终的可用性分值。本次测试结果的 SUS 分值为 83.2, 对应的可用性评级为 A。这一方面说明课题组所匹配的音效具有较好的满意度, 满足产品使用标准, 另一方面也验证了声音在用户操作交互中所起的辅助性表达, 以及与图标的语义匹配性。

4 结语

听觉是人类感知外界过程中的重要信息渠道, 因此声音是产品和用户之间的一种沟通媒介, 聚类分析是研究声音的重要方法之一。论文针对设计师对产品声音识别与归类的需求, 提出一种交互式可视化的声

音数据分析算法框架。该算法框架融合设计师感官描述信息和声音技术特征参数信息,使声音数据信息来源更为全面系统;基于高斯混合模型的聚类算法不仅能适应非线性几何分布的产品声音数据,还能在交互过程中进一步参考设计师用户先验知识和需求,得到吻合设计师预期的最优聚类结果。同时,聚类的可视化分析可以清晰地展示聚类过程的数据元素流向、聚类结果的类别比例,增进了设计师对产品声音的理解。实验结果表明,论文所提出的声音数据分析算法框架能有效地帮助设计师用户进行产品声音的聚类,方便后续设计实践中的产品声音的合成、设计与匹配应用。

参考文献:

- [1] ÖZCAN E, VAN EGMOND R. Memory for Product Sounds: The Effect of Sound and Label Type[J]. *Acta Psychologica*, 2007, 126(3): 196-215.
- [2] 李檬, 王安霞, 邓丽. 感官设计理念在包装设计中的应用[J]. *包装工程*, 2008, 29(3): 128-130.
LI Meng, WANG An-xia, DENG Li. Application of Sense Organ Design Idea in Packaging Design[J]. *Packaging Engineering*, 2008, 29(3): 128-130.
- [3] 王年文, 徐丽. 基于感官评价方法的多线切割机造型设计研究[J]. *机械设计*, 2018, 35(5): 116-120.
WANG Nian-wen, XU Li. Research on Modeling Design of Multi-Wire Cutting Machine Based on Sensory Evaluation Method[J]. *Journal of Machine Design*, 2018, 35(5): 116-120.
- [4] CHANDRAN K M, PRABHUNE P, SEN D. Sound Association for Product-Sound Design Using Semiotics[J]. *ICoRD'15-Research into Design Across Boundaries*, 2015, 1(2): 491-503.
- [5] CUMMINGS A, ČEPONIENĖ R, KOYAMA A, et al. Auditory Semantic Networks for Words and Natural Sounds[J]. *Brain Research*, 2006, 1115(1): 92-107.
- [6] LU Wei-hua, PETIOT J F. Affective Design of Products Using an Audio-Based Protocol: Application to Eyeglass Frame[J]. *International Journal of Industrial Ergonomics*, 2014, 44(3): 383-394.
- [7] 林霜, 黄若涵. 基于多模态感官用户体验的产品声音感知研究[J]. *装饰*, 2019(8): 102-105.
LIN Shuang, HUANG Ruo-han. Research on Product Sound Perception Based on Multimodal Sensory User Experience[J]. *Art & Design*, 2019(8): 102-105.
- [8] 李伟, 李硕. 理解数字声音——基于一般音频/环境声的计算机听觉综述[J]. *复旦学报(自然科学版)*, 2019, 58(3): 269-313.
LI Wei, LI Shuo. Understanding Digital Audio—A Review of General Audio/Ambient Sound Based Computer Audition[J]. *Journal of Fudan University (Natural Science)*, 2019, 58(3): 269-313.
- [9] ÖZCAN E, EGMOND V R. Basic Semantics of Product Sounds[J]. *International Journal of Design*, 2012, 6(2): 41-54.
- [10] ZHANG Wen-juan, DAI Ming, ZHANG Hong-wei. A Comparison of Hardware Implementation of BP Neural Network Based on NIOS II Single-Core and Multi-Core Processor[C]//2011 International Conference on Electric Information and Control Engineering. Wuhan: IEEE, 2011: 2853-2856.
- [11] 魏娜, 辛向阳. 模糊 C 均值聚类算法在自动供送装置改进设计中的应用[J]. *机械设计*, 2016, 33(12): 105-108.
WEI Na, XIN Xiang-yang. Application of Fuzzy C-Mean Cluster Algorithm on the Improved Design of the Automatic Feeding Device[J]. *Journal of Machine Design*, 2016, 33(12): 105-108.
- [12] SANCHEZ-LENGELING B, WEI J N, LEE B K, et al. Machine Learning for Scent: Learning Generalizable Perceptual Representations of Small Molecules[EB/OL]. (2019-10-25) [2022-10-22] <https://arxiv.org/abs/1910.10685v1>.
- [13] IMAM N, CLELAND T A. Rapid Online Learning and Robust Recall in a Neuromorphic Olfactory Circuit[J]. *Nature Machine Intelligence*, 2020, 2(3): 181-191.
- [14] HOSSAIN M S, OJILI P K R, GRIMM C, et al. Scatter/Gather Clustering: Flexibly Incorporating User Feedback to Steer Clustering Results[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(12): 2829-2838.

责任编辑: 陈作