构建用于在人类身上安装蝙蝠回声定位技能的训练界面

在人类身上

Miyoko TSUMAKI[1,a], Yu TESHIMA[1], Takao TSUCHIYA[2], Kaoru ASHIHARA[3],Kohta I. KOBAYASI[1], Shizuko HIRYU[1,b]

[1] 生命与医学科学研究生院、同志社大学、日本

[2] 理工学院、同志社大学、日本[3] 人类信息学与交互研究所、日本产业技术综合研究所（AIST）、日本

摘要

蝙蝠利用一种称为回声定位的复杂超声感知方法来识别周围环境。近期有研究报告称，没有回声定位经验的视力正常参与者可以通过训练聆听回声，提高感知各种环境空间布局的能力（Norman 等，2021）。本研究利用眼动追踪仪开发了一种新的人类回声定位训练系统。我们采用波动方程有限差分时域方法，模拟了受回声定位蝙蝠觅食策略启发的连续向下线性调频脉冲所产生的双耳回声。在实验中，视力正常的参与者需要识别隐藏目标的形状，而虚拟回声则根据他们在显示器上的眼球运动进行呈现。我们的分析表明，训练缩短了形状识别所需的感知时间；通过将注意力集中在目标边缘的策略，形状识别的准确率也有所提高。我们认为，该系统有助于对人类回声定位进行定量验证。关键词：眼

-追踪 ，人类回声定位 ，虚拟回声

1. 引言

盲人回声定位专家（EEs）通过自行发出的口腔咔嗒声，利用反射声波（回声）感知周围环境[1]。与此同时，据报道，即使是没有回声定位经验的视力正常的参与者，也能通过训练[2]提升空间识别技能。然而，实际的回声测量需要投入时间和精力进行实验[3]，，而回声定位技能的有效训练方法和学习过程仍有许多尚不明确之处。

本研究提出了一套训练系统，利用三维声学仿真，根据不同形状的虚拟目标生成回声，训练参与者学习回声定位技能。参与者根据显示器上的眼球运动，通过聆听生成的虚拟回声对目标进行回声定位。我们使用该系统，根据眼动追踪器获取的注视位置，分析了参与者的感知时机、方向和形状识别过程，并考察了缺乏回声定位经验的视力正常参与者所采用的感知策略。

2. 方法

2.1 基于脉冲响应估算回声

采用基于波动方程有限差分时域法（WE-FDTD）[4]，的三维声学仿真，将仿真人头（HATS；头部与躯干模拟器 Brüel

[a]mtsumaki2021@gmail.com[b] shiryu@mail.doshisha.ac.jp

& Kjær TYPE 4128-C) 的形状数字化，并保存为立体光刻（STL）文件。将 HATS 和一个虚拟目标的 STL 模型置于仿真空间（1。5 × 1。5 × 1.5 m）（图 1）。 仿真的空间分辨率为 1.5 mm，Courant-Friedrichs-Lewy（CFL）=0.5，声速设为 344 m/s，满足三维仿真稳定性条件。由 HATS 嘴部发射一个信号（脉冲），并通过数值分析计算虚拟目标在左右耳道入口处（接收点）的回声。将计算得到的回声定义为回波脉冲响应。此时，将目标划分为 5 x 5 cm 的网格，并将声源（HATS 嘴部）位置的 x 和 z 坐标设为每个网格中心位置的 x 和 z 坐标，以计算虚拟回声。训练任务包含 8 种虚拟目标，另加 5 种未训练版本，测试任务共使用了 13 种不同目标（图 2）。

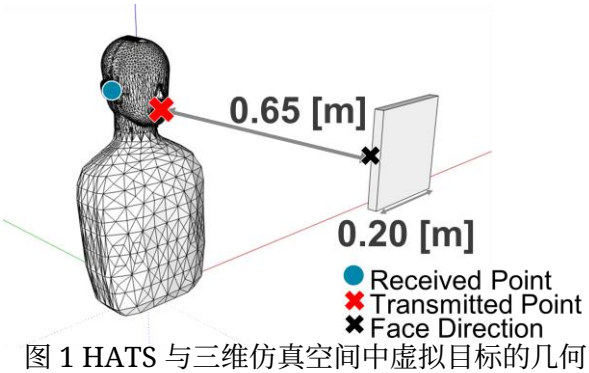

图 1 HATS 与三维仿真空间中虚拟目标的几何配置示例。通过倾斜 HATS 改变嘴部方向，使发射信号入射到目标上的任意位置。

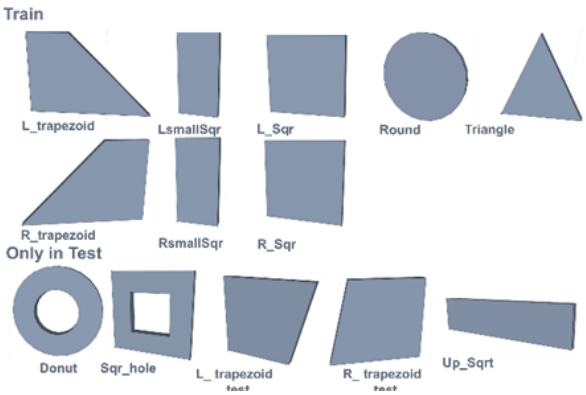

图 2 训练环节和测试环节中使用的虚拟目标。

2.2 心理声学实验

心理声学实验在隔音室（2.3×1.6×1.4 m）内进行。参与者为1名男性和4名女性（22 - 25岁），视力正常。根据使用听力计（AA-77A，RION，东京，日本）进行的标准纯音测听，确认所有参与者在0.25至3 kHz范围内的听阈均在30 dB以内。涉及人类的心理声学实验已获同志社大学伦理委员会批准。在取得参与者的知情同意并进行听力测试后，我们开始了心理声学实验。每个声音刺激由一个10 ms的向下线性调频声音（7 kHz - 1 kHz）构成，该声音与2.1节所述的回波脉冲响应进行卷积；随后以30 ms的间隔重复播放30次，以模拟受回声定位蝙蝠觅食策略启发的嗡鸣声。通过个人电脑上的PsychoPy3（适用于64位Windows的独立版2021.2.3（使用Python3.6））[5], 控制向参与者播放声音刺激；声音刺激由个人电脑经音频接口（OCTA-CAPTURE UA–1010；Roland）传输，并通过耳机（MDR-CD900ST；SONY）播放。

参与者的注视位置（注视点，POG）通过一台 150Hz 桌面式眼动追踪仪（GP3HD眼动追踪仪，Gazepoint，加拿大温哥华）进行记录。网格状辅助线显示在显示器上，耳机则播放与 POG 对应的网格虚拟回声。参与者可自由地通过自主注视，感知隐藏在显示器上的虚拟目标的形状。当参与者认为自己已经能够画出该形状时，便自行结束该环节，随后在带有网格的纸上画出他们识别出的目标形状。训练试次共 8 次（提供正确反馈），测试试次共 7 次（不提供正确反馈）。完成试次后，参与者填写了一份书面问卷。

3. 结果与讨论

3.1 虚拟回声定位训练中参与者的感知策略

图 3 展示了一名参与者在感知后绘制的目标形状（红色虚线所示）。相似度按以下方法计算。为了定量评估参与者绘制的形状与所呈现的虚拟目标轮廓之间的差异（下文简称“差异”），我们使用基于 Hu 矩的 OpenCV 库 matchShapes 函数进行计算 [6]。如果阈值小于 0.02，则判定为形状相同（例如，结果越低，匹配度越高；返回值越接近零，相似度越高）。所有参与者对训练过的目标所产生的差异平均为 0.11，而对未训练目标（即首次尝试）的平均差异为 0.17，这表明即使面对未训练过的目标，参与者也在一定程度上具备形状辨别能力。 此外，训练环节中达到形状识别所需的平均感知时间为 179.6 秒。测试环节中，针对训练过的目标，平均感知时间为 155.0 秒；针对未训练目标，平均感知时间为 167.0 秒，均短于训练环节。由此可认为，训练对回声定位具有学习效果。

我们分析了 POG 数据，以比较受训目标与未受训目标的分布差异。比较训练环节与测试环节（受训目标）发现，在测试环节（受训目标）中，参与者的注视更倾向于落在目标边缘及目标外部，尤其是那些对形状差异表现出较低敏感度的参与者。此外，在测试环节中，与未受训目标相比，参与者对受训目标边缘的关注更多，感知时间也更短。这些结果表明，参与者在训练过程中调整了感知策略，将注视固定在目标边缘或目标外部。各参与者的 POG 分布以及所有环节结束后的问卷回答均表明，来自目标边缘的虚拟回声具有显著特征。

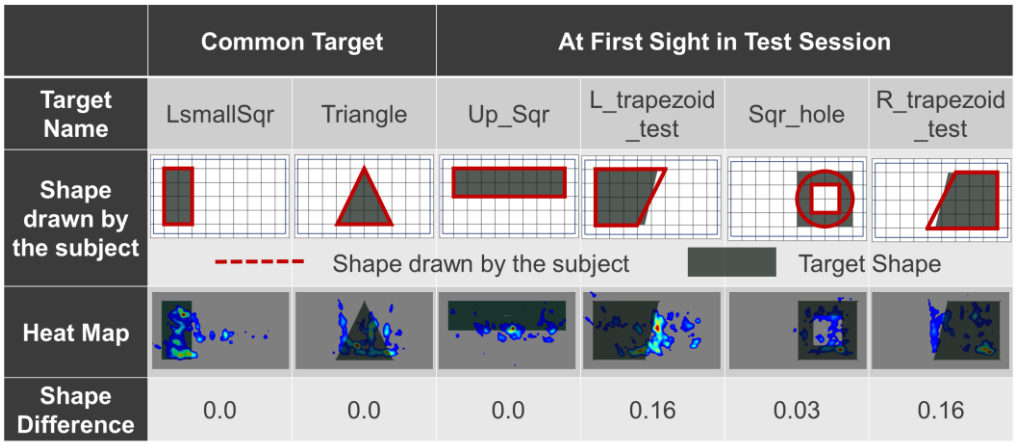

图 3 问卷中绘制的形状（上）、测试环节中的热图（中），以及目标与受试者绘制形状之间的差异值（下）。红色虚线表示受试者绘制的形状。

4. 结论

本研究提出了一种系统，参与者可以在自由通过视线感知的同时聆听虚拟回声，并借此探究人类回声定位的学习过程。结果表明，参与者转而采用聚焦于虚拟目标边缘的感知策略后，利用回声定位进行形状识别的准确率有所提高。该回声定位训练系统为量化回声定位的学习过程提供了一种有效方法。

参考文献

1. Thaler L，Goodale MA。人类回声定位：概述。《WIREs Cogn Sci》。2016；7：382–393。pmid:27538733。

2. Norman L、Dodsworth C、Foresteire D、Thaler L（2021）人类点击式回声定位：失明与年龄的影响，以及为期十周的训练计划中的现实生活意义。《PLoS ONE》16（6）：e0252330 3. Sumiya M、Ashihara K、Yoshino K、Gogami M、Nagatani Y、Kobayasi KI 等。受蝙蝠启发的信号设计在人类回声定位中的目标辨别作用。《J Acoust Soc Am》。2019；145（4）：2221–2236。pmid:31046316 4. Y. Miyazaki 和 T. Tsuchiya，“波动方程有限差分时域方法的完全匹配层”，《Jpn. J. Appl. Phys.》51，07GB02-1–5（2012）。5. Peirce, J. W.、Gray, J. R.、Simpson, S.、MacAskill, M. R.、Höchenberger, R.、Sogo, H.、Kastman, E.、Lindeløv, J.（2019）。PsychoPy2：轻松开展行为实验。《Behavior Research Methods》。10.3758/s13428-018-01193-y 6. OpenCV、OpenCV模块，https://docs,opencv.org/4.6.0/（2022/7/15）