**学习从听去看**

**The vOICe 培训手册**

**最后更新于 2018.11.24**

The vOICe 培训手册的更新版本将提供压缩的附有 MP3 音频文件的 Microsoft Word 文件（[下载链接](https://www.seeingwithsound.com/manual_cn.zip)）。另提供[俄语](https://www.seeingwithsound.com/manual_ru/The_vOICe_Training_Manual_ru.htm)和[葡萄牙语](https://www.seeingwithsound.com/manual/The_vOICe_Training_Manual_pt_br.htm)翻译。

**目录：**

[1 介绍](#_Toc50454421)

[2 图像到声音的映射原则](#_Toc50454422)

[2.1 有声景的图像样例](#_Toc50454423)

[3 抓取物体](#_Toc50454424)

[4 理解距离和尺寸](#_Toc50454425)

[4.1 外观随距离的变化](#_Toc50454426)

[4.2 其他距离和尺寸线索：视差和遮挡](#_Toc50454427)

[5 视觉透视](#_Toc50454428)

[6 视觉地标](#_Toc50454429)

[7 地平面障碍物](#_Toc50454430)

[8 培训计划](#_Toc50454431)

[9 效果对照](#_Toc50454432)

1 介绍

看，对视力正常的人来说不费吹灰之力，他们通常认为视觉是理所当然的。然而，视觉事实上是一项高度复杂的技能，现实中，它仍然极大地挫败了计算机视觉在物体识别方面的最大努力。神经科学已经表明在视力正常的人中，一大部分的大脑直接或间接参与处理对来自眼睛的输入。另外，视觉本身往往是模棱两可的，它需要对世界的知识、对上下文的知识和广泛的观看经验来可靠地消除来自环境的典型视觉输入歧义。 The vOICe 传感替代技术将原始视觉转换为相应的声音场景（下称声景），同时保留了大量的视觉信息。从技术上讲，这可以让你用声音去看，但是你的大脑必需首先学会用视觉上有意义的术语去解码声景。这项全新的技能需要时间和练习去掌握。本培训手册旨在为你提供一套练习指南，帮助你学习理解编码为声音的视觉输入所需的基本技能。之后，通过在现实生活中广泛地、身临其境地使用 The vOICe，你便能够更深入的掌握用声音去看，最终使之成为习惯。因此，这与学习一门外语没有什么不同，你首先要学习语法和词汇以奠定坚实的基础，然后通过多年的实际运用使它变得越来越轻松自然。没有人能一夜之间学会一门新的语言，同样， The vOICe 也不是立即生效的灵丹妙药。就像学习一门新的语言一样，即使进步有时看起来很慢，最终也能熟能生巧。

本培训手册删去了某一版本的 The vOICe 在特定操作系统（如 Microsoft Windows 或 Android，或 iOS 上的 look-alikes）上如何运行以及特定硬件设置（例如摄像头眼镜、智能手机或增强现实眼镜）的技术细节，聚焦于解释如何理解声景以及适用于不同软硬件实现的 The vOICe 的普遍原理。请在 [seeingwithsound.com](https://www.seeingwithsound.com) 网站查阅实现细节和免责声明，例如针对运行在 windows 系统下的 [The vOICe for Windows](https://www.seeingwithsound.com/winvoice.htm) 软件和[摄像眼镜](https://www.seeingwithsound.com/camera_glasses.htm)，针对智能手机的 [The vOICe for Android](https://www.seeingwithsound.com/android.htm) 以及基于安卓系统的 [AR 眼镜](https://www.seeingwithsound.com/android-glasses.htm)。本培训手册同样关注与动态视觉相关的技能（在基于视觉的导航的帮助下移动、障碍物检测以及抓取物体）。本手册的重点不在任何更具体的用途，如阅读报纸头条、门牌或阅读图表以及天气图。对于视力正常的人来说，视觉的用途数不胜数。本手册关注于视觉核心技能的培养，事实上它并不仅仅为人类和人类文化所独有，而是整个生物王国生存的关键。在习得核心视觉技能后，你能够更好地独立学习与你的个人情况、个人兴趣以及工作相关的其他技能。

如果你不是完全失明，你必须蒙上自己的眼睛，或者在使用摄像眼镜时佩戴完全不透明的太阳镜夹片，这样你就无法在练习中借助残余视力。当然，一旦你完全掌握，你可以利用来自残余视力或回声定位的补充线索，但不要在练习中使用。

不要用嘈杂的声景使自己感到疲劳，保持音频的音量为对你有效的最柔和的水平。因为当你在细微的声音线索中能够保持活跃时学习效果最好，并且你可能会对声音产生更高的敏感度。使用低音量可以最大限度地减少非线性失真以及降低听力损伤的风险。条件合适时可以把 The vOICe 静音。

|  |
| --- |
| ***速度要求*** |
| *在使用 The vOICe 时的一个常见错误便是行动缓慢，一直有意识地、仔细地分析视野。在某些时候，你必须跳过有意识地分析才能变得熟练，学会快速地执行动作。在保证安全的情况下迅速行动。视力正常的人在抓取东西时并不思考：他们只是去做！* |

2 图像到声音的映射原则

The vOICe 的图像到声音的映射只有三条简单的规则，每一条都涉及视觉的一个基本方面：规则1涉及左右，规则2涉及上下，规则3涉及明暗。规则如下：

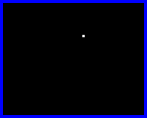
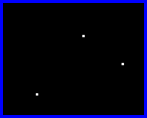
1. 左右  
   按照从左到右的顺序扫描视野，默认为每秒一个图像快照。你会听到相应的立体声从左到右平移。因此，在你的左侧或者右侧听到一些声音意味着在相应位置有一个与声景对应的视觉图案。
2. 上下  
   在每次扫描中，音高意味着高度：音高越高，视觉图案的位置越高。因此，如果音高变高或变低，你就有一个对应的上升或下降的视觉图案。
3. 明暗  
   响度意味着亮度：声音越响，亮度越高。因此，寂静意味着黑色，响亮的声音意味着白色，介于两者之间意味着灰色。

也就是说， *The vOICe 从左到右扫描视野，同时将高度与音高，亮度与响度联系起来*。换句话说，将每个视野扫描成细的垂直切片，从左侧的垂直切片开始，到右侧的垂直切片结束。生成的声音取决于当前垂直切片的视觉内容，并且在较高位置的明亮像素具有更高的音高。上述映射方法是通用的，可以表示任何灰度图像。图像，例如相机拍摄的照片，本质上是二维的，并不明确表示距离，但在稍后的“[理解距离和尺寸](#理解距离和尺寸)”一节中，我们将讨论如何使用 The vOICe 获取和解释距离线索。

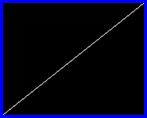
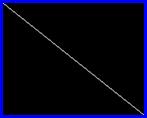
2.1 有声景的图像样例

掌握图像到声音映射原理最简单快捷的方法是根据几个简单的图像样例，从它们的视觉描述开始，分析视野听起来“应该”是什么样，下面是一些示例声景：

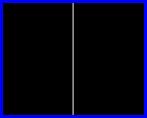
* 一个漆黑的背景意味着没有声音。黑暗是沉默的。
* 一个亮点对应一个短的哔声，音高代表高度。当视野中有多个亮点时，你会听到多个短的哔声。

[](https://www.seeingwithsound.com/manual/dot.mp3) [](https://www.seeingwithsound.com/manual/3dots.mp3)

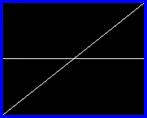
* 一条上升（下降）的亮线发出上升（下降）的音调；这条线越陡，音调的音高变化越快。

[](https://www.seeingwithsound.com/manual/lineup.mp3) [](https://www.seeingwithsound.com/manual/linestup.mp3) [](https://www.seeingwithsound.com/manual/linedown.mp3)

* 一条垂直的亮线对应咔哒声。

[](https://www.seeingwithsound.com/manual/linevert.mp3)

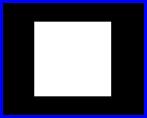
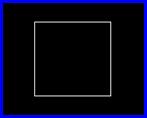
* 一条水平的亮线对应一个恒定的音调。一条水平亮线和一条上升亮线同时发出两个音调，一个是恒定的，一个是上升的。

[](https://www.seeingwithsound.com/manual/lineflat.mp3) [](https://www.seeingwithsound.com/manual/lineflup.mp3)

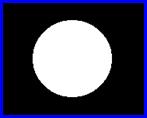
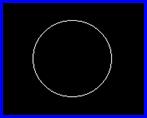
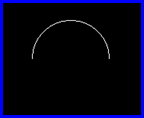
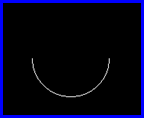
* 两条平行的亮线对应两个同步的音调。三条平行的亮线对应三个同步的音调。

[](https://www.seeingwithsound.com/manual/lineflat2.mp3) [](https://www.seeingwithsound.com/manual/lineflat3.mp3)

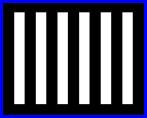
* 一个垂直的光亮填充的正方形或长方形对应一个骤然发生和戛然而止的突发声响。总体的音高和音域决定了高度和范围。对于一个垂直的亮线描边，内部为空的正方形或长方形来说，左边缘发出一条垂直亮线对应的声音，然后顶部边缘和底部边缘发出两条平行的水平亮线对应的声音，最后右边缘发出一条垂直亮线对应的声音。一个有五个声景的[视频](https://www.seeingwithsound.com/manual/squareani.wmv)演示了如何构建一个亮填充的正方形。首先描绘底部边缘，然后描绘右边缘、顶部边缘和左边缘，最后进行填充。

[](https://www.seeingwithsound.com/manual/square1.mp3) [](https://www.seeingwithsound.com/manual/square0.mp3)

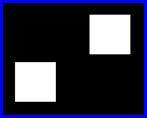
* 一个光亮填充的圆听起来像一个骤然发生，逐渐停止的突发声响。一个亮线描边，内部为空的圆听起来像两个同步的音调，一个音调表示上半圆先上升随后下降，另一个表示下半圆先下降随后上升。

[](https://www.seeingwithsound.com/manual/circle1.mp3) [](https://www.seeingwithsound.com/manual/circle0.mp3) [](https://www.seeingwithsound.com/manual/circle0tophalf.mp3) [](https://www.seeingwithsound.com/manual/circle0bottomhalf.mp3)

* 一组等间距明亮竖条对应了一段由突发声响构成的持续的节奏。例如，华盛顿特区的一个栅栏或者白宫的柱子都有这样的图案。

[](https://www.seeingwithsound.com/manual/pillars.mp3) [](https://www.seeingwithsound.com/manual/fence.mp3)  [](https://www.seeingwithsound.com/manual/whitehouse.mp3)

* 两个垂直的光亮填充的正方形或长方形听起来像两个突发声响。

[](https://www.seeingwithsound.com/manual/squares.mp3)

上述示例的图片链接了含有相应声景 MP3 声音文件，单击图片将会播放它对应的声景。下载文件后，你可以用 Windows Media Player 打开，并循环播放，来捕捉声景的细节。

掌握上述方法之后，你就可以开始探索现实生活中的物体看起来（听起来）是什么样子。比如，把你的白色手杖放在一个黑色的表面上，注意它是如何根据它的方向来发出上升或下降的音调的。在黑色表面的几个闪亮的硬币将会发出相应的哔声。你也可以试试书架上的书脊所形成的节奏，或是窗帘合拢时发出的节奏。请注意，一般情况下从室内看到的窗户听起来像一个明亮的矩形，因为室外环境光通常更亮。事实上，城市环境中许多人造建筑从正面看都是矩形的：不仅是窗户，还有门以及大楼。

当你听到一个声景，并通过触摸或其他方式验证了对应的视觉内容是什么，请试着理解它为什么听起来是这样。听到一个音调，可能对应一个边缘。听到一个节奏，可能对应一段垂直的格栅。听到一个平滑的突发声响，可能对应一个平滑的表面。有意识的、合理的分析将会帮助你理解那些不易理解的景象。同时，注意细节会提升你的倾听技巧。永远保持好奇心和求知欲！

3 抓取物体

需要掌握的最重要的实践技能之一就是抓取物体的能力。通过使用 The vOICe，你现在无需搜掠或摸索就能够学着去看：你可以通过声景看到并直接抓取物体。为此，你必须首先掌握如何协调相机和手，就像视力正常的人眼手协调一样。这意味着你要理解声景中的点对应的视觉方向，并且需要机动性以准确地到达特定的方向，最终达到不假思索的程度。本节中的描述中假设你使用的是头戴式摄像头，摄像头指向你鼻子的方向。但适当更改叙述后，同样的步骤也可以适用于手持摄像头，例如智能手机。但是，强烈建议使用能够提供最直观一致的观看体验的头戴式摄像头，尤其是摄像眼镜。

练习确定方向与抓取物体的先决条件：

* 一个深色桌面。比如深色木制桌面，或者盖上深色布的桌面。非反光黑色，以及黑毡是有最佳视觉对比度的首选。
* 一些明亮的东西，供你扔在桌子上并伸手去拿。亮黄色或白色的 DUPLO 砖是很好的选择，但你也可以使用白色聚苯乙烯泡沫塑料包装材料切割成的手掌大小的物体，只要这些物体与深色桌面形成良好的视觉对比，并且可以像骰子一样投掷而不从桌上滚下来。下面的一些描述中，我们假设您使用的是 DUPLO 砖块，它是一个有八个顶点的立方体。

坐在深色的、没有任何分散注意力的东西的桌面旁，确保当所选的亮色物体放在桌面上时光线充足，能够被“看”到。并且检查当没有物体放置在桌面上时，桌子表面是否显得空荡荡（声景中无声或有柔和的突发声响）。深色表面作为一个不分散注意力的背景，让你聚焦于物体以及它们在相机视野中的位置。

现在把一个明亮的物体放在桌子上，使它在桌子上反弹一些落在桌子上一个随机的位置而不掉下来。你的任务是在不用手搜掠桌子的前提下抓住这个物体。因此你需要在视觉上定位物体，并伸出手臂，用手抓住物体。最可靠的方法是首先在你的（相机）视野中使物体居中，然后伸手去抓。

为了做到这一点，你必须首先通过环顾四周，让尚未出现的物体在相机视野中被捕获。一个明亮的小物体听起来很像哔声。一旦你注意哔声，你必须置其于你的相机视野的垂直及水平中心，这样的条件下，物体就在正前方，也就是你鼻子指向的方向上。当一个物体在从左到右的扫描过程中发出一半声音并且音高中等时，说明它处于声景的中心。接下来将更详细地描述这种定心对准方法。

[](https://www.seeingwithsound.com/manual/duplo5.mp3)

对于垂直对准，你需要上下倾斜你的头部直到物体的哔声为中等音高。接下来，在保持这个音高的同时，将你的头部左右转向直到每次扫描开始半秒后发出哔声。也就是说，在每次默认一秒的声景扫描时间的一半位置。接下来保持物体发出声音的方向一直向前，而不是向左或向右。然后你可以伸手抓住物体，想象它在你鼻子指向的方向。（同样的，如果您使用的是手持式摄像头而不是摄像眼镜，请根据您的手持方式想象摄像头的观看方向。）

在初期的练习中，有些偏差是很正常的，但是在几个小时的练习之后，你应该在大部分情况下可以抓住物体。持续练习，直到它变成一个自然而然的动作。你可以同时试着把你正在寻找的物体形象化，以强调你正在完成的事情的视觉本质。尽量避免陷入旧习惯，即用手在桌子上搜掠来定位物体（尽管快速修正几厘米是可以的）。我们的目标是快速而正确地抓住物体，这完全是可以通过练习做到的。学习相机和手的协调是很快的，因为投下并抓住 DUPLO 砖块只需要五秒钟左右，仅仅十五分钟就可尝试两百次。

此外，尽管只有一瞬间，你仍可以通过判断音高上升或下降的速度很容易地分辨出一块 DUPLO 砖块的朝向。判断砖块是否倒立着（或是侧着）很困难，但砖块的八个顶点以及底部的斜脊图案都有一种有助于增加“看”的真实感的特殊的声音质感。

一旦你掌握了上述方法，你就可以放宽将物体居于视野中心的条件，直接在音高和声源的指示下抓取任何偏离中心的物体。这有些困难但更有效率，因为你可以在大约一秒钟的时间内从第一个出现物体的声景中抓取它。在掌握如何确定物体在视野中心之后，更为推荐这种无需首先确定中心的抓取物体方式。

在单物体抓取练习之后，下一阶段是把两个或三个 DUPLO 砖块扔到桌子上，并且无需逐个将其确定于视野中心来进行抓取。因此，你可以从单个声景中非常高效地抓取多个物体中的每一个。这也很像掌握一门外语，你首先要掌握对严格的语法规则的有意识的应用。一旦这些规则烂熟于心，你就可以“忘记”有意识地应用它们。因为从那时起，有意识的应用只会减慢你的速度。事实上，某些时候你只有跳过有意识的分析才能变得熟练。以此类推，没有人能通过思考何时左转右转来骑自行车：这样只会导致摔倒。感官技能必须变成潜意识的、自然而然的才能为需要注意力的更高级别的目标服务。

[](https://www.seeingwithsound.com/manual/duplo_flash.mp3)

一旦你能轻松完成桌面抓取练习，就可以开始针对移动环境进行训练——但是一定要在安全的环境中（家里）。从距离桌子几米远的位置开始，将物体置于视野中央，然后朝着它走，同时保持物体在视野中居中。注意，你需要弯腰以防止物体从视野中消失。最后，当你离的足够近（物体占据了你视野的一个明显的部分）的时候抓住它。把这个练习想象成一个捕食者在追踪猎物的过程。开始时，相距遥远的物体会发出微弱的哔声。当你逐渐靠近，它的外观会变得更加明显，而且物体看起来“更大”。下面这段包含五个声景的[视频](https://www.seeingwithsound.com/manual/duploani.wmv)演示了这个过程。

[](https://www.seeingwithsound.com/manual/duplo1.mp3)[](https://www.seeingwithsound.com/manual/duplo2.mp3)[](https://www.seeingwithsound.com/manual/duplo3.mp3) [](https://www.seeingwithsound.com/manual/duplo4.mp3) [](https://www.seeingwithsound.com/manual/duplo5.mp3)

为了使这项练习更贴近日常生活，你可以用白色的咖啡杯或餐盘代替 DUPLO 砖块。单物体抓取练习听起来很简单，虽然确实如此，但首要任务是正确掌握基本知识。单物体抓取能够帮你建立自信，去很好的利用通常极其复杂的声景的某些方面。它也是移动环境中许多更复杂的行为和场景的一部分。伸手去抓门把手是这项技能的另一个直接实际应用。

*开始时，需要进行每天半小时的单物体抓取练习，至少持续两周。*这会帮助你建立能够应用于日常生活的相机和手的协调。通过更为长期的练习，你能够做得更好。但是你必须达到一定的技能水平才能从 The vOICe 中受益。

一旦你完全掌握这项练习，你就可以在有足够视觉对比度的前提下，伸手去抓取你能够到的物体，比如咖啡杯、门把手、硬币，当然还有你的白手杖。

[](https://www.youtube.com/watch?v=AZ1xigijrxQ)

视频1 使用 The vOICe 掌握相机和手的协调

4 理解距离和尺寸

图像并不显式地表示视野中物体的距离和大小，但是存在许多可供学习的关于距离和大小的间接视觉线索。

4.1 外观随距离的变化

最有力、最普遍的距离和尺寸线索是，当你来回移动时，物体的表观尺寸是如何变化的。这里也有一个非常简单的视觉透视法则：*当距离减半，物体会看起来是原来的两倍大。*这意味着，如果一个物体已经很近而且已经占据了视野中的很大一部分，那么仅仅接近一点就会导致它占据更大部分。因为当已经距离物体很近时，不需要花费太多就会使距离减半。此外，当物体占据了视野的大部分同时距离很小时，物体的物理尺寸将与其物理距离相当（对于 50 到 120 度的典型相机视角，给定或取两倍系数）。一个物理宽度或高度为半米的物体通常在你的视野中有一个胳膊的长度。但是，一个大得多的物体会超过你视野的宽度，导致你的视野中的表观尺寸不足以辨别距离。

如果你坐在深色的桌面旁，把一个明亮的咖啡杯放在手臂够的着的地方。当你前倾或将杯子移向自己，杯子会立刻显得更大。另一方面，一个远处的物体，比如一个远处的大楼，无论在你的视野中看起来多大，当你朝着它走几步时，它的表观尺寸几乎不会改变。因为距离太远，所以将会需要走许多步来使距离减半并在视野中尺寸显示为原来的两倍。因此，通过判断前后移动时表观尺寸的变化量，可以判断和物体的物理距离。请记住，而且你必须充分意识到这一点——表观尺寸本身与距离无关：距离 50 米的建筑可能具有与距离 5 米的门有相同的矩形外观和表观尺寸，尽管建筑实际高度和宽度是门的 10 倍。因此，你永远无法从一个单一的静态视野中做出可靠的距离判断（除非你识别出物体，并知道其物理大小，可以将它的表观尺寸与物理尺寸联系起来）。你必须要移动才能可靠地判断距离。这是一种叫做动态视觉的概念，因为在四周移动时比静止时能够获取更多的信息，尤其适用于获取距离和大小的信息。

这为障碍物检测和导航提供了一种强大的方法。如果一个静止的物体在走近几步之后明显变大，那么它必然距离较近，可能是一个你希望绕过或至少在用手杖接触它之前预料到的障碍物。另一方面，如果一个静止的物体走近一步之后表观尺寸几乎没有变化，那么它就距离遥远，是一个可以用来导航的潜在的地标。例如保持希望的行进方向，这一部分在“[视觉地标](#视觉地标)”中将进行叙述。

上述的方法也可以应用于等距垂直格栅。如果当向前迈一步时，相关的节奏变慢了两倍，那么格栅距离很近。比如有一个在手臂够的着的书架。如果几步之后节奏的速度变化不大，那么可能在十米的距离处有一道栅栏。

接下来，用不同大小、不同距离的物体广泛练习如何捕获这种距离线索。这种线索通常是相当微弱的，至少在刚开始练习的时候，它需要有意识地努力去不“忽略”一些嵌入在典型环境的声景中的变化。微弱的线索很容易被淹没，你需要学会“倾听”那些对你重要的细节。能够更好地感知明显的变化对你的移动非常有利，可以帮助你绕过障碍物，追踪各种远处的地标。理解原理远远不够，你必须每天进行至少十分钟的练习。

*刚开始练习在前后移动中感知尺寸变化时，每天练习半个小时，至少持续两周。*加上抓取物体的练习，意味着每天要花一个小时进行训练，至少持续两周。之后，练习时间可以减半，使用不同的物体在不同的距离下前后移动，进行十五分钟抓取物体的练习和十五分钟理解距离和尺寸的练习。每天练习半小时是能够取得良好进步的最低标准，就像你需要每天至少练习半小时才能在学习一种乐器或一门外语方面取得好的进步一样。你应该维持这种强度的日常训练至少一年。当然，这是一个权衡的问题：成为一个钢琴演奏家需要每天练习几个小时的乐器，但由于其他社会或工作责任，期望每个人这样做是不现实的。如果在提升视觉技能方面集中注意力并保持持续的兴趣，那么每天半小时的练习一定能使你达到一个不错的水平。

4.2 其他距离和尺寸线索：视差和遮挡

除了在前后移动中得到的距离提示，你还可以通过视差现象，从侧向移动中获得距离提示。另一种视觉效果是，从你的视角看，物体后面的任何东西都会因遮挡而隐藏起来。此外，阴影和投影也可以提供更为微妙和隐含的距离线索，尽管后者在 The vOICe 的应用中通常不是非常有效。视差和遮挡是“[外观随距离的变化](#外观随距离的变化)”一节曾讨论的距离线索的有力补充。

*视差*

当向侧面移动时，视野中的物体似乎在朝相反的方向移动，但位移量取决于它们的距离。远处的背景看起来根本不会移动，而近处物体相对远处背景的位移表明这些物体确实距离更近。你可以根据按位移量的排序判断物体的距离顺序。

在开始的练习中，站在一根柱子前，并且更远处的背景（比如建筑物）有一些视觉图案。注意当你向旁边走一步和向后走一步时声景是如何变化的。如果操作正确，也就是说在没有同时转身的情况下，你会注意到背景保持不动而柱子移向了与你移动方向相反的方向。另外，当柱子距离较远时，它看起来会随着你的步子移动得更少。如果背景似乎也在移动，说明你在不经意间稍微转身了。反之亦然，你可以通过保持视野中的背景不变来确保你的朝向没有改变。

*视觉遮挡*

除非一个物体碰巧是透明的，比如玻璃，你看不到它后面的东西：物体后面的东西被遮挡住了。这使得更远处的物体和部分背景会全部或部分隐藏起来。与视差结合可以使隐藏的东西随着侧向移动而改变。

*阴影*

表面与光源方向的夹角会影响反射光的视觉亮度。因此，一个物体，比如一个球体，在它的表面上会显示出一个取决于它的三维形状的特定的亮度变化，从而提供了一个距离（和形状）线索。这很难给出明确的规则，需要你从各种类型的物体及其在各种照明条件下的视觉外观中学习。

*投影*

来自平行光源的光可能被物体遮挡，导致物体后面的表面区域显得更暗。这些较暗的区域称为投影。与阴影一样，物体的投影提供有关物体在环境中的三维位置的隐含信息，从而提供距离（和形状）的补充线索。

5 视觉透视

你已经在“[理解距离和尺寸](#理解距离和尺寸)”的叙述中了解到了视觉透视的一种效果，就是距离减半时物体看起来变大两倍。但是，这种效果会产生一系列相关的影响。例如，当你沿着一条道路向远处看时，道路在两倍距离处显得窄了一倍，而在非常远的地方，道路看起来变得非常窄，这就是为什么人们有时会说所谓的消失点。它让长而直的道路看起来像一个三角形，其中上顶点位于消失点。如果道路两旁都是一排排的建筑，那么这些建筑会在从左到右的声景扫描中呈现出节奏。当朝向消失点方向扫描一排建筑时节奏速度加快，然后当随着道路右侧的一排建筑被追踪直到声景在附近的建筑处停下时，节奏会逐渐变慢。因此节奏在远处时加快而在近处时变慢，而且由于建筑物通常在物理尺寸和形状上大致相当，如果你已经知道自己在路上，这可以给你关于距离的线索。这样的练习并不简单，因此我们用有类似的变化的电脑键盘的声景来代替。找到一个普通的电脑键盘在下个练习中使用，并从上面俯视键盘。

[](https://www.seeingwithsound.com/manual/keyboard_topview.mp3) [](https://www.seeingwithsound.com/manual/keyboard_topcloseup.mp3) [Computer keyboard seen from above and on the left side
](https://www.seeingwithsound.com/manual/keyboard_leftslanted.mp3) [](https://www.seeingwithsound.com/manual/keyboard_rightslanted.mp3)

键盘上的一排排键给出了一种特有的节奏。如果你把相机移近键盘，这种节奏就会变慢，而随着距离的增大，节奏会加快，就像先前的书架一样。这是因为视觉物体在近距离看起来更大，因此视野中出现的按键数目更少。现在从某个角度看键盘。你会注意到，在声景中，节奏会加快或减慢，因为键盘较远的部分的节奏比较近的部分快。准确地说，节奏快慢的变化取决于相机的方向。试试不同的角度和距离，直到你完全熟悉声景如何随距离和方向而变化。

通过听键盘节奏有多快以及它是加速或减速，你可以分辨出距离和方向。同样的效果也适用于道路沿线的一排排建筑物和每栋建筑中的一排排窗户，只是这种情况下，由于涉及的距离较大，可能要走上几步才能注意到声景的变化。

6 视觉地标

远处背景中的独特的视觉图案可以作为视觉地标。由于距离很远，当你向前走时，地标在视觉上的变化很慢。它们有助于保持恒定的朝向。例如，如果你在一个开放的停车场，保持一个远处的建筑在声景中的固定位置可以像一个视觉“指南针”帮助你。如果视觉图案足够独特，它们也可以帮助你知道你在一条路线上的位置，而不需要计算步数。有些建筑的正面可能有一些柱子，这些柱子具有独特的节奏；有些建筑的窗户有特殊的排列方式，使它们在其他建筑中脱颖而出，而有些商店可能在入口上方用大号字母标出店名。毕竟，大多数公司都想脱颖而出，受到关注。试着记住在你的环境中有什么东西有特殊的声景，这样随着时间的推移，在熟悉的路线上的所有位置上，通过环顾四周并注意路线上该位置或多或少的独特图案，你就可以知道你在哪里。当然，这不适用于看起来都一样的城市街区。

7 地平面障碍物

Title: Handling ground level hazards - Description: Schematic illustration (for sighted trainers) showing how to adjust your viewing angle as you move closer to a ground level obstacle.
虽然在可靠地探测到附近的地面危险（如障碍物、台阶或坑洞）方面，没有什么比白手杖更好的了，但适当的视觉策略也能起到帮助作用，尤其是在用手杖击中潜在危险之前，很好地预见到潜在的危险，甚至让你在通过左右转向绕开这些危险而无需用手杖接触到它们。此处我们将更详细地考虑这样的情况：除非障碍物被手杖碰到，你一直朝着它走，直到有被绊倒的风险。

一般来说，走路时最好稍微向下看。最佳视角通常不是正前方（除非你面对的是远处的地标），因为头顶以上的物体很少会出现碰撞威胁。相反，与头部高度相比，地平线通常应该显示在相机视野的顶部，而不是中间。这与直视前方相比，能让你在察觉头部高度危险的同时看到更多的地面。最佳视角取决于相机的视野，对于具有 45 度垂直视野的相机来说，向下倾斜 20 度较为合适。然后，一旦你发现一个潜在的地面危险，逐渐将你的头向下倾斜，以避免障碍物离开你的视线。当你靠近它时，让它靠近你的视野底部来跟踪它。视觉尺寸的变化，再加上你头部倾斜程度的增加，能够告诉你离得有多近，如果是地面障碍物，你可以避免被绊倒。建议偶尔短暂抬头，以确保在同一时间内没有出现头部高度危险，因为在跟踪地面危险时，难以注意到由于头部向下倾斜增加而造成的危险。

8 培训计划

如前几节所述，所建议的最低程度训练计划如下：

第一周和第二周：

* 三十分钟[抓取物体](#抓取物体)的日常练习
* 三十分钟[理解距离和尺寸](#理解距离和尺寸)日常练习

第三周及以后，至少一年：

* 十五分钟抓取物体的日常练习
* 十五分钟理解距离和尺寸日常练习
* 在日常生活中运用 The vOICe

更多的训练时间是有帮助的，但是不要过度，不要超过建议时间的两倍。然而，你在日常生活中使用 The vOICe 的时间并没有真正的上限，就像视力正常的人整天用眼睛一样。

9 效果对照

在自我训练中，至关重要的是你要对自己的表现非常挑剔，确保达到预期的表现水平，为进一步使用 The vOICe 打下坚实的基础。如何知道你已经达到了本培训手册的目标呢？为此，你可以对照以下几点来判断：

* 在一个深色的桌面上同时随机地投下两个明亮的物体，比如 DUPLO 砖块，之后，我可以直接在三个声景内抓住它们，而不必用手搜掠去纠正偏差。我至少有一半时间是对的。这证明我已经掌握了很好的相机和手的协调。
* 在一个熟悉的环境（家）中，我曾经练习过 The vOICe 的移动使用，现在我可以自由走动，而不必接触任何墙壁、门柱或家具。我可以向前或向后走一步，根据物体表观尺寸的变化来判断我与附近物体的距离。接下来，我可以在无需尝试且没有失误情况下伸手去触碰物体的特定的点，比如物体的边缘。
* 从一个熟悉的房间的一端开始，我可以走到房间另一端的有显著物品的地方，一旦声景中表观尺寸快速变化，表明物品在手臂可及的范围内，我就会停下来。然后我可以伸手去触摸物品的某个特定点。显著物品可以是远处墙上的一幅画，或者是房间另一端门口的门柱。
* 从一个熟悉的房间开始，我可以转身几次，然后很容易地从声景中调整自己的方向，朝着特定的方向走，比如朝门口走去。
* 当我把钥匙扣掉在一块深色地毯上时，我可以很容易地用 The vOICe 找到它，弯下身子，用一个平稳的动作直接捡起，而不需要任何摸索。

永远记住，如果你在训练后仍在执行一项基本的视觉任务时动作缓慢，那说明你做得还不够好：你应该能像有视力的人一样很快地完成。不断地催促自己提高速度，否则你可能会陷入一个缓慢的有意识分析的水平，无法充分发挥你的潜力（也不能发挥 The vOICe 的潜力）。

Acknowledgement: Thanks for the translation work by [Yake Wei](https://echo0409.github.io/) and [Di Hu](https://dtaoo.github.io/).

致谢： The vOICe 中文训练教程由[卫雅珂](https://echo0409.github.io/)和[胡迪](https://dtaoo.github.io/)翻译完成，特此感谢。