

The vOICe 软件使用说明

The vOICe 软件使用说明

石蟾 译

以下是一款盲用软件测试版的简要说明。该软件可将来自PC摄像头、电脑屏幕、扫描仪或图片文件等的图像绘制成相应的（合成）声音文件。在高等教育的实际应用中，新一代读屏用户的电脑中内置声音/图形计算器，以及颜色识别器。该软件为免费版，仅供个人学习、研究之用，不得用于商业用途。

“vOICe”一词中的“OIC”意为“*Oh I See!*”（我看见了！）这是一种通过声音达到视觉效果的新型的、有效的、用途广泛的技术，它并非基于声纳技术或回声定位，而是使用来自PC摄像头或“网络摄像头”的实图输入（下文会另外讨论网络摄像头）。如今PC摄像头已经十分便宜，一般价格50美元到100美元之间。通过特殊的软件，盲人只要带着手提电脑和PC摄像头，就可以从耳机中听到周遭的景象，这样他们就能听到与明眼人所见相同的形状和事物了。该软件可将摄像头中的图像即时转换成相应声音。比如，一个光点会发出短促的“噼”声，如果光点在左面，就可以听到左面有“噼”声，如果在右面，就可以听到右面有“噼”声；如果光点上移，可以听到“噼”声音调升高，光点下移，则音调降低；有两个光点就有两个“噼”声，有三个就有三个“噼”声，依此类推。水平光线能发出长音，这是因为光线由多个光点组成，相应地就能产生多个连续组合的“噼”声，听上去就像是单个长音，同样地，如果光线上移或下移，音调也会随之改变。垂直光线是光点的叠加，虽然也是同时听到，但由于高度不同，音调也各不相同，这种光线听起来就像突然发出的短促杂音。在指导性实验中，你要将盲杖置于摄像头前，因为盲杖酷似一条光线，这样就可以通过使用盲杖的定向行走来测试声音效果了。

这种声音绘制的图形可以描绘任何景象，但是图象越复杂，声音也越复杂。描绘一幅全景大约需要一秒，摄像头捕获景象的变动情况每秒钟都会刷新一次，这点点滴滴的声音组成的图象被称作“音响风景”，通过表示位高的音高和表示亮度的响度，以及左右甄别，它就可以“听”到景象了。尽管如此，还是请注意，现实生活中的视觉景象通常相当复杂，你会被听到的声音完全搞晕。图象总是如此复杂，无论你是先天失明还是后天失明，“听”图象对你的大脑都是全新的挑战。即使明眼人也会晕头转向，因为他们只学过用眼睛看图，而非耳朵。最重要和最困难的一点是，如何通过对这一技术的全方位应用，达到精通这一技能。然而不尝试就永远不知道。希望这些视觉信息能带给你一些愉悦，即使目前似乎没有多少实际用处。也许听起来好象听中国话一样难懂，感觉毫无意义，但等你接受了全面训练并掌握这一技能后就完全不同了。

你是不是已经厌烦了上面的说明？那就快去试试这款“听”图软件——“vOICe”吧！事实胜于雄辩，“音响风景”胜于千言万语，你可以通过以下网址下载在微软Windows系统中运行的voice.exe程序：

http://www.seeingwithsound.com/voice_CN.exe

这个软件小于500K，需要PC摄像头或网络摄像头支持其识别功能和运行，你可能还没装摄像头，因此当你启动程序时，你可能会听到语音提示：“对不起，vOICe未在您的电脑上找到摄像头”，然后，一个内置的合成声音开始自动重复（测试音效描述了一张照片，上面有一辆车子，以楼房和街道为背景，还立着一根电线杆。）你应该听到立体音响分别在左右声道提示“left”和“right”，不然的话，说明你的耳机或音箱放反了。另外，别被这测试音效吓倒，因为即使你还没有安装摄像头，你仍然可以用这款软件做许多事。

比如，你只要简单地按下功能键F11，就可以打开（或关闭）内置练习模式，测试版音效消失了，你可以在随机位置听到包含着两个充满亮光的矩形图案，然后可以按空格键进入手动更新练习模式，选听另一个出现在随机位置的矩形图案。你还可以用小键盘的加减号（数字状态）选中某一图形，用光标键移动所选图形的位置，编辑菜单里还有各种类型和数量的图形可供选择，这里先跳过不说，一些细节操作须要依靠你的读屏软件和声卡支持，有时vOICe的声音会阻止读屏软件发声，可按CTRL+F2恢复读屏声音，这一情况仅出现在老版本的Windows系统中，Windows 2000，XP及以上版本的系统不存在此问题。

那么这些图形是如何发声的呢？正如刚才所说，在大约一秒钟内，所有图片接受横向扫描，以音高表示位高，响度表示亮度。这一方式完全可以拓展到任何图片的听取。但是这种方式对现在众多极其复杂的图片而言实在难以奏效。然而在双矩形练习模式中，你可以分清两个响声，这两个声音的持续时间由图形宽度决定，其音高由图形高度和位高决定，其响度由图形亮度决定。一条横线上下起伏时，你可以在从左至右的扫描过程中听到上下波动的长音；而对于光点，你可以听到短促的“噼”声，而其音高则代表其位高。

注意，最初从左至右扫描出来的立体摇摄似乎是视野中的一个动作，但其实这只是扫描的效果：即使视野中没有任何变化，立体摇摄也仍然有效。就好象你反复地用手从左向右抚过一张盲文点字纸，这样就通过触觉感受到了其纹理变化。当你的手抚过纸面的时候，你所听到的“动作”并不是你视野中的一部分景象，而仅仅只是一种线性的扫描方式。

你想在尝试这款软件之前先听一下示例音效吗？那就先听听一些WAV音效吧。一个88K大小的WAV示例音效可以让你知道横线和其它基本图形是如何同时听出来的。你可以通过以下地址直接下载这个时长2秒的音效：

<http://www.seeingwithsound.com/voiscopebw2.wav>

深灰色瓷砖的背景下有一条明亮的弧线和十个小正方形，你能听出来吗？太简单了吗？好，那么就下载同一图片的1秒音效试试吧：

<http://www.seeingwithsound.com/voiscopebw.wav>

注意：为了连续多次听取同一音频以便头脑达到最为清晰的分析状态，你最好把音频播放器调成自动重播。明眼朋友可以通过以下链接中的原始GIF图片对照以上示例音效。

<http://www.seeingwithsound.com/voiscopebw.gif>

当把这张图片导入vOICe学习编辑软件后，它就能自动合成为示例音效，盲人用户也可以使用该软件的此项功能：按CTRL+O导入图形文件。听到一个图形的声音时，你可以按功能键F3将图形声音从1秒减慢到2秒。如果你想放大或者拖动图片，可以按功能键F4后使用光标键操作。

如果你正在使用平板式扫描仪，你还可以用它来做图形输入设备。按CTRL+Q进行扫描，你就能听到扫描仪上放的各种东西的图象了，甚至放上你的手也可以。按功能键F5启用反色功能，就可以听出亮色背景上的小块暗色物体，通常针对素描线条和打印材料。还有一个简单的实验：如果你把有色服饰置于扫描仪上，你只须待扫描完毕后按功能键F10就能听到它的颜色名称——如是，vOICe也可以用作一个廉价的颜色识别器，它能告诉你视野中央部位的物体颜色名称。

想知道屏幕上鼠标所在位置的图象吗？按功能键F9，然后移动鼠标或者按光标键在屏幕上的图形界面上移动，你就可以听到窗口和图标的边框以及任何图形了。

正如你所见，有那么多使用方式可供选择，而这得取决于你想做什么。软件还自带科学图形计算器供数学计算，按功能键F8，默认的是带有横纵坐标轴的一条正弦曲线，但你可以通过读屏，在一个标准对话框中设置其它函数和区间。以下网页上有自动图形计算器的详细说明：

<http://www.seeingwithsound.com/winmath.htm>

而可读图表和函数图的摄像软件在以下网页上可以找到：

<http://www.seeingwithsound.com/prgraph.htm>

以下网页上还有自我训练教程。

<http://www.seeingwithsound.com/training.htm>

尽管vOICe软件能反馈一些非常基本的信息，但是要充分、独立使用vOICe学习编辑软件的所有特性和功能，盲人用户还须使用读屏软件听取菜单和对话框中的内容。初学者可以使用一种作为第三方应用程序的免费读屏软件“Thunder”，可在<http://www.screenreader.net>下载（vOICe软件与该产品无任何关联）。

让你沉浸于真实视觉体验中的移动设备。

当你拥有一个摄像头并在手提电脑或上网本上安装了驱动程序后，你可以把摄像头缠在或缚在耳机上，这样你就得到了一台外部装有摄像头的移动装置了。你可以在以下的FlyCamOne网站上方便地找到价格适中的外置网络摄像头。

<http://www.seeingwithsound.com/flycamone2.htm>

(注：voICE软件与此产品无任何关联。)

无论你怎么使用移动装置，都请确保自身及他人的安全，因为大部分视觉信息开始时可能会非常混乱，影响注意力，而且提示声音肯定会覆盖掉一些环境声音，强烈建议你在家庭等熟悉安全的环境中开始试用，这也便于你将自己所“见”和所知事物联系起来。现在我们简单谈一下某些视觉概念，这对于那些仅了解部分视觉概念的先天全盲者有所帮助。

从某种意义上说，全盲者的移动有些象在布满潜伏沟壕的事物间跳跃，除非那些事物会发出声响。这当然说得过于简单，但你应该明白了。只有中断之前靠触觉或回声捕获的事物，全盲才能再次用触觉或回声找到下一个物体或障碍物。而视觉看物体时具有较好的连续性，因为视觉可以同时捕获多个物体，下一物体出现时前一物体尚未从视野中消失，因此即使各个物体或目标之间存在很大的距离，视觉所感知的物体仍然存在较大的重叠性，这也有助于定位路线。另外，这里提到的与视觉相关的内容也适用于摄像头中的“音响风景”，因为“音响风景”中包含了相同的视觉信息。

除了近处多少总有几个物体以外，视野中还有整个视觉背景，这里的背景是指那些没有被近景物体挡住的其他事物，包括地平线上的房屋、大楼或其他远景。由于远景看起来很小，这种布满细小事物的背景总显得杂乱无章，另外要分清哪些是近景哪些是远景也需要经过一番大脑数据处理。那是很复杂的过程，尽管明眼人的这种毕生训练似乎轻松得很，但我们还是可以就大脑如何掌握这种线索说一说。其中关键点就是表面形状变化。一扇十米远的门和一座100米远的大楼也许看上去大小一样，两者都像是视野中的矩形阴影。然而，如果你朝门走几步，门看上去一下子变大了，而大楼却几乎没变化，这是因为你离较近的门的距离与离你较远的大楼的距离相比变化显著许多，因此，尽管从某些角度看，一些物体的形状相同，但是当你移动时，变化的多少就能反映出哪些物体在远处，哪些物体在近处。

当你移动时，作为一部分远景的“音响风景”正是一种持续性表现形式（当你向上或向下看，向左或向右转时，别管整体的音高变化）。因此，如果你注意“音响风景”中没有变化的部分，你就对自己要前往的地方有大致概念了，这就类似一个“视觉罗盘”。你一直朝远景走去，它也会发生变化，但速度会慢许多。

现在想象你正走向一辆静止的车辆，切记尝试前要采取适当的预防措施以保证安全，比如有明眼朋友伴随，最好有定向行走指导师。把摄像头对准汽车，很可能车子距离很远，显得很小，图象中的剩余部分将由车子周围的事物填充，包括其背景。只有你走近了才会慢慢变大，直至整个“音响风景”被车子填满，但是如果你向车子走去时有别的事物进入视野，比如另一辆汽车，那么之前那辆车就会从你的左面或右面消失。所以简言之，除了简单的水平和垂直转动以外，正是移动时感知到的变化量表明了你附近存在什么事物。当然要做到这点非常困难，但是至少体验和学习使用“音响背景”有助于你理解这种视觉原理。再次提醒，行动前请确保环境的安全，因为这种新型的信息输入是你未曾掌握的，你的头脑会被搞晕，而“音响风景”也会对你正常的听觉系统有所影响。

假如你是第一次接触视觉图象，以下的自然听觉对比将有助于你理解近景和远景的视觉概念。假设你左面远处是一条车水马龙的道路，右面远处是一所朗朗读书声的学校，尽管你向右转身时，听上去它会向左转，反过来也是如此，但是当你向前迈出几步时，这种听觉背景却会像视觉背景一样不发生什么变化。然而，根据周围声音的不断呈现，你可以听到突然出现的静止的车辆或墙壁的回声，这款软件也一样，通过它你可以更为迅速地感知周围事物的变化。

根据自然听觉背景和摄像头捕获的“音响风景”，远景可以帮助你保持前进方向。只有在这种情况下附近的事物才会混淆。再简单些说，近处的物体固然重要，因为它们可能成为行走时的障碍物，但远处可感知的景物也同样重要，特别在定位、把握总体方向、避免转向等方面尤为重要。

如果你知道那种由并列相间的栏杆组成的金属栅栏或防盗门，这种物体形象就可以帮助你理解透视的效果了。假设有一道栅栏，在使用vOICe软件时，这种栅栏就会根据其栏杆的间隔频率发出节奏性很强的声响，这种“吱吱”声一般不会漏听。栅栏离得远，节奏就快，当你走近栅栏时，节奏就会变慢，因为栏杆之间的空隙看上去变宽了，直至你走到离栅栏一臂远的地方，就只能听到很少几根栏杆的声音了。当你走到跟前顺着栅栏方向看时，你还会发现，较远的栏杆声音节奏快，较近的栏杆节奏慢。这全都是因为远处的物体看上去比较小的缘故。

当你开始全身心地体验“音响风景”时，你会发现，有些事物即使不知道它具体是什么，你都可以听懂，这种对透视法规则的理性认识以及对某物体的注意观察至少有助于初学者对摄像头捕获的繁杂声音形成初步理解。

这一用途广泛、成本低廉的技术将来的应用，很大程度上取决于你自己的决心和毅力。现在谁都不知道人们到底能够达到什么精通程度，学习曲线往往趋于陡直上升。当然，你并非必须使用它，但对那些有意愿、有需要的人来说，这也是一种选择。尝试与否由你自己决定。

正在使用或有兴趣使用vOICe软件的盲人朋友也可以加入“The vOICe user group”（vOICe用户小组邮件列表），你可以发一封主题为“subscribe”（不包括引号）的邮件到：

seeingwithsound-request@freelists.org

你也可以听听一篇有关用户的音频报道，这篇报道名为“Seeing with sound: A journey into sight”（《听图—畅游五彩世界》），发表于2002年4月8日的图森2002“知觉”大会。音频地址是：

<http://www.seeingwithsound.com/tucson2002f.ram>

你也可以听听CBC Radio One的科技节目《Quirks and Quarks》中的一篇特辑，MP3音频地址是：

<http://www.seeingwithsound.com/media/qq-2005-04-02a.m3u>

你可以考虑使用租来的或二手的配件，通常建议投入较小的成本，这样，根据你对这种声音定位方式的适应程度，你可以自己决定是否使用，因为这种技能不是一夜就能掌握的，而且要达到预期效果也并不容易，无法保证其可行性。

Giuseppe Masciopinto编纂的意文版vOICe学习编辑使用手册（PDF格式）可以在以下网页找到：

http://www.seeingwithsound.com/extra/guida_vOICe_1.0.pdf

Radion Mynayev编纂的俄文版vOICe使用手册（PDF版）可以在以下网页找到：

http://www.seeingwithsound.com/extra/The_vOICe-russian.pdf

vOICe补充说明

1 热键功能表

1.1 常规

空格 声音暂停/继续；练习模式中手动模式切换；井字游戏中划圈

小键盘加减号(数字状态) 练习模式中更换选中图形

光标键 放大状态下移动图形显示焦点；练习模式中调整选中图形位置；井字游戏中上下左右移动

1.2 文件

CTRL+D 设置运行路径

CTRL+O 导入图片文件

CTRL+U 从网上下载并导入图片文件

CTRL+Q 运行扫描仪

CTRL+M 将声音及图片以邮件附件形式发送

CTRL+SHIFT+C 复制摄像头当前图象到剪贴板

1.3 各项功能

注意，以下热键均为键，即按第一次时生效，按第二次时取消

F1 3D模式

F2 停止声音播放

CTRL+F2 恢复声音播放

F3 二分之一扫描速度

CTRL+ALT+F3 四分之一扫描速度

SHIFT+F3 二倍扫描速度

ALT+F3 四倍扫描速度

CTRL+F3 八倍扫描速度

F4 图象放大二倍

SHIFT+F4 图像放大四倍

CTRL+F4 图像放大八倍

CTRL+SHIFT+F4 图像放大十六倍

F5 反色

F6 明暗对比加强

SHIFT+F6 明暗对比加倍加强

F7 边线加粗

F8 函数设置

F9 鼠标区域显示

ALT+F9 激活窗口显示

CTRL+F9 激活窗口客户端显示

SHIFT+F9 全屏显示

F10 提示颜色

ALT+F10 无提示

F11 随机图形练习

SHIFT+F11 井字游戏

2 颜色中英对照表

light 淡色的

dark 深色的

red 红色

orange 橙色

yellow 黄色

green 绿色

cyan 青色

blue 蓝色

magenta 紫色

white 白色

black 黑色

grey 灰色

3 练习模式详解

3.1 首先可以设置练习模式中的图形：

先按ALT键弹出菜单，找到设置对话框：

Edit - Exercise Preferences - Randomly Placed Shapes

打开对话框后，焦点依次停留在以下编辑框中：

矩形数量，默认为2；

圆形数量，默认为0；

三角形数量，默认为0；

线条数量，默认为0；

声音自动重复次数，默认为3；

图形像素，默认为10；

之后还有两个副选框，分别是“黑色边框”和“彩色模式”。

3.2 练习举例：

比如设置两个矩形，两个圆，一个三角形和一根线条，则在开始练习模式后，每次在一张图片中出现所设置的所有图形，当然大小、角度、位置都是随机的，且有可能重叠在一起，此时按空格键切换成手动模式，按小键盘加减号（数字状态下），就可在图片中的所有图形中来回切换，并可听到语音提示所选中的图形，选中某图形后，可按上下左右光标键移动图形位置，如听到“嘀”的一声提示，表明图形已不能再移动了。

以下是提示音中各种图形的中英文对照：

rectangle 矩形；

circle 圆形；

triangle 三角形；

line 线条。

因此在本例中，按加减号可以依次听到语音提示：

“rectangle,rectangle,circle,circle,triangle,line”，如听到第二声“circle”时停止按键，即选中第二个圆形，此时按光标键调整该圆形位置，听到的声音发生了变化，声音变化的部分就是相应的圆形位置变化。