

## 鉴赏音乐如同品酒：音乐厅的感知式评价

(北京大学 朱星 编译自 Tapio Lokki. *Physics Today*, 2014, (1): 27, 原文详见 <http://ptonline.aip.org>)

一个音乐厅的音响效果是如何影响听众的感觉？如果运用合适的方法，通过询问听众的感受，我们会获益良多。

当幕间休息开始时，我的夫人迫不及待地走向门厅。在饮完一杯葡萄酒后，她不停地讲述我们刚刚听过的音乐是多么棒，多么不可思议地听出小提琴独奏中非常轻柔的音色。我感到愕然。刚才的演出并没有使我感动，尽管那是我所喜欢的让·西贝柳斯的小提琴协奏曲。虽然独奏者表演得很好，但是不知为什么，我没有感到音乐的魅力，也没有使我想去欣赏。难道是音乐厅的声音效果破坏了美妙的演出吗？我的夫人一直在赞扬演出，音乐厅并不高明的音响效果并不能转移她的享受感。显然，我们对葡萄酒和音乐的感受如此之不同。

在六年前那个时刻，我开始意识到葡萄酒和音乐厅具有如此之多的共性。它们都具有多重的认知特性。如何对它们进行评价涉及到个人的喜好。当对不同的音响效果和葡萄酒进行比较时，不同人对于声音和品味的关注点不同，他们表达的感受也不相同。酿酒人发明了不同的方法，用来确定使得一种酒与众不同的因素。酒香轮盘就是用于评价葡萄酒多种口味和香气的方法之一。那么葡萄酒业是如何通过多位品酒者的不同认知，对酒的质量进行总体评价呢？能否将他们的评价方法引入到音乐厅音响效果的感知评价中呢？这样或许能够更好地理解音乐会上听众的不同感受。一个多世纪以前，Wallace Sabine对音

乐厅的音响效果进行了开创式的研究，科学家尝试去理解，哪些感知特性会影响人们对音响效果的看法。为了理解人们对封闭空间内复杂声场的反应，研究人员对室内声场进行了客观和主观方式的研究，这与葡萄酒业的方法类似。

客观的方法是用分析声能的衰减来研究音乐厅。特别的是，研究人员在观众席的某个位置测量脉冲响应，接收发自舞台的一个简短声信号，记录其随时间的变化。从这些测量中，可以推导出一系列声学参数，如混响时间、声强、先后传来的声能间的平衡、以及所测量声音的空间分布，虽然国际标准化组织对这些参数均有描述，但它们仅来自于在几个测量座位的空间平均值。对这项标准的批评说：“用来计算参数的算法不精确，所应用的频率范围太窄，另外，一个单一的、无定向的声源不能代表在一个真实交响乐队中几十个声源的情况。”进一步，客观测量的参数不能细致描述听众所感知到的音响效果。

对不同音乐厅进行主观比较并非易事。音乐本身、指挥和交响乐队的水准都会影响听众的感觉，Leo Beranek提出一种对世界上一流音乐厅排序的方法。德国哥廷根和柏林、英国巴斯大学等团队也做出了有影响的研究。主观评价也可以在实验室条件下进行，如运用虚拟音响技术，将音乐信号与脉冲响应

卷积，这种实验可以在真实音乐厅进行，或者通过室内音响模拟方式进行。

### 采样比对

葡萄酒品尝中所使用的感知评价方法的基本特点是，评价员可以对不同样品进行比较。将倒满不同品种酒的玻璃杯排成一排，提供给品酒员，这样他们可以任何顺序对这些酒品尝任意次。每品完一杯酒，他们用语言描述他们所品尝的酒的特点，用一个词语来描述不同酒之间的感知区别。感知评价可以通过“共识词汇法”进行，即品酒员们与同组人员讨论，以获得共识的评价，使用形容词来描述酒。而与此对应的“单一词汇法”是品酒员们独立工作。每一个人用不同的属性来描述，比如酒的甜度或者香醇程度，然后用这些属性对酒进行排序。统计分析表明，所有品酒员给出的排序都能表明这些酒最主要的区别，整个组的品酒员都能够觉察到这种区别。

“共识词汇法”需要评价员经验丰富，他们能够理解所使用的词汇的复杂内涵。而“单一词汇法”则没有这个要求。比如，一位品酒员将一种酒描述为“水果味”，而另外一位描述为“浆果味”。如果他们按照水果味或者浆果味排序，很有可能他们排出的顺序相同。在某些场合，个体感知或许很不相同，但是实际上，当较多的评价员，如10—

15人或者更多时，他们对酒的重要特征的看法一致。

很显然，当将感知评价方式用于音乐厅的音响效果评价时，所遇到的挑战是对大厅即时效果的比较。人类的听觉记忆很短，对于听音者来说，在不同的音乐厅中，现场听取音乐来对音乐厅进行比较是不可靠的。必须将不同音乐厅及不同座位的音响差异记录下来，然后在实验室重放，这样听音者可以在不同音响之间进行切换。进一步的，所录制的音乐必须由同一交响乐团在每个音乐厅中以相同的力度和节奏进行演奏。事实上，即使是专业的交响乐团也难以做到，这是因为音乐家们会按照音响效果自动调整他们的演出风格。为了在录制中区别音响效果，我的研究团队的解决方法是用扬声器构建自己的交响乐队，如图1(a)所示。

这个模拟交响乐队是由34个经过校正的扬声器组成，按照真实乐队的方式排列。扬声器的摆放和方向与乐器位置尽可能相同。比如，每个小提琴通道包括两支扬声器，

一支朝向听众，另外一支方向朝上，这是因为小提琴发出的声音方向主要向上，特别是高频声音。舞台上每支扬声器都必须尽可能发出其代表的乐器的声音，不受其他乐器或者乐器所在录音房间的声音响应的干扰。我们的解决方法是，在一个消音房间中单独为每位乐师录音。当在演奏乐师自己的部分时，他们在一个小屏幕上观看指挥的录像，同时听着音乐的钢琴声部。令人惊讶的是，交响乐队的音乐家能够准确地按照音调和节奏演奏，而不需要周围音乐家提供的视觉和听觉的提示。我们录制了不同年代不同交响乐队的四个乐段，每段2—4分钟，由规模不同的乐团演奏。

当我们在音乐厅中通过扬声器播放这些录制的音乐时，听起来很真实。但是，为了录制大厅中听众所感知声音的空间效果，并且能够在实验室重放，我们必须录制来自不同入射方向的复杂声场。事实证明这是一种无法逾越的挑战。虽然已经有了许多空间声音的录制与重

放的技术，但是它们的目标是制作艺术家式的优质录音，而不是声音的真实复制。

不同的是，我们采用了脉冲响应的方式进行空间声音的录制。从由34个扬声器组成的交响乐团的每一支扬声器中，我们播放了一个频率对数的扫描信号，然后在一个听众座位上安放了由6个定向麦克风组成的阵列，用来测量其响应，如图1(b)所示。通过分析这些脉冲响应，可以估计声能入射方向随时间与频率的改变。将消音室内的乐器录音与空间脉冲响应数据进行卷积，我们计算出我们所需要的信号，使得听音室内通过24个扬声器发出的声音能够真实地重现在音乐厅中通过麦克风交响乐团发出的空间效果。

#### 属性与偏好

至今为止，我们录制了20个欧洲音乐厅的音响用作感知评价。由于交响乐队和录音系统都是经过标定的，这样可以区分由音乐厅建筑结构所造成的音响效果。如同在一排葡萄酒杯前评酒一样，我们可以对不同音乐厅的音响效果进行直接比较。

运用感知评价方法，我们可以很好地评价不同音乐厅或同一大厅内不同座位的音响效果。特别是，我们发现“单一词语法”是一种可靠的方法，可以用来提取不同音乐厅之间的感知差别。首先，听音者提供给我们的词语可以描述音乐厅之间的最显著的差别。其次，在多维空间中将这些属性归类，可以对所研究的音乐厅给出共识和感知印象。进一步，可以将这些音乐厅按照这些属性归类进行排序。如果我们请听音者按照演出效果对音乐厅进行排序，他们整体的偏好倾向可

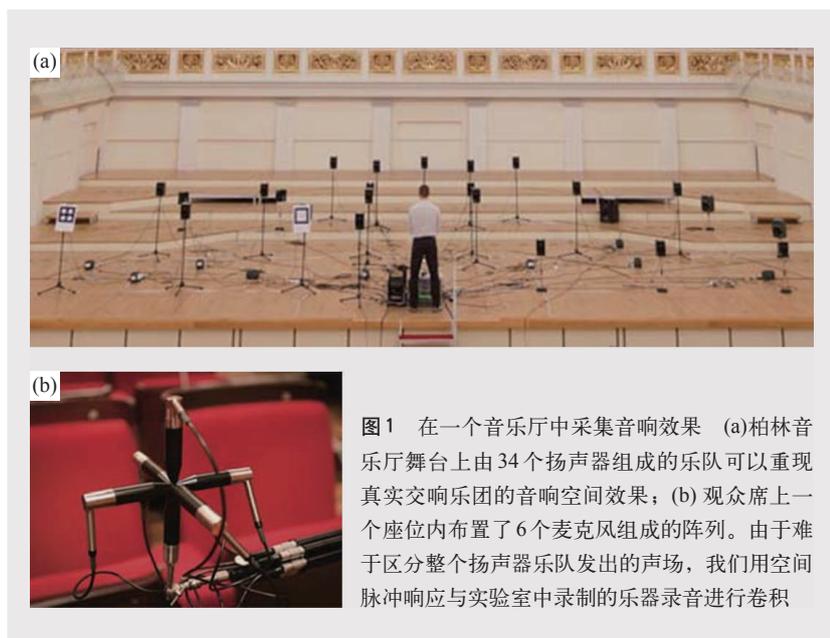


图1 在一个音乐厅中采集音响效果 (a)柏林音乐厅舞台上由34个扬声器组成的乐队可以重现真实交响乐团的音响空间效果；(b)观众席上一个座位内布置了6个麦克风组成的阵列。由于难于区分整个扬声器乐队发出的声场，我们用空间脉冲响应与实验室中录制的乐器录音进行卷积

以与音乐厅的感知印象相关联。

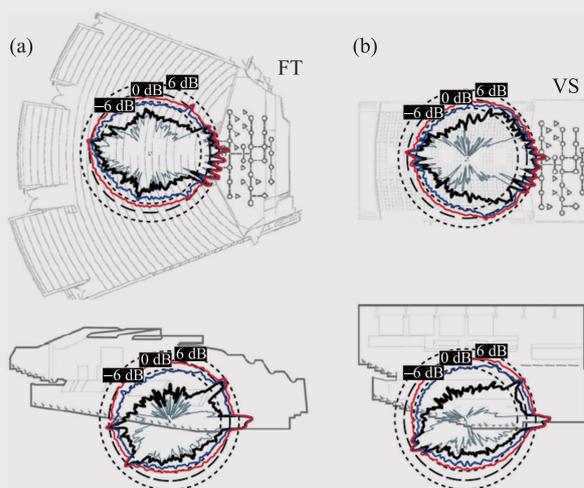
在我们首次进行的主要研究中，我们请20位听音者评价了三个芬兰的音乐厅中各三个录音位置的音效。一些录音点所在的座位是在包厢，其他位于交响乐队附近。评价者辨认并确定了102个特征。我们的分析表明，仅仅一类特征与总音量和听音距离有关，由于录音位置的物理距离有很大不同。另外，我们发现不同的听音者可能使用相同的描述词语，但是他们要表达的内容却完全不同。比如，不同听音者描述的属性“洪亮”分别属于两个不同组。一些听音者描述的特征与空间大小有关，另一些人将这个特征与来自各个方向的所包围的音效相关联，而不仅仅来自前方。这个结果如果是在由研究人员所确定属性的听音试验中是不可能得到的。感知评价和单个词汇法在提供丰富的认知信息时显现出重要作用。

在下一个研究中，在每个音乐厅中录音距离固定为12 m，在9个音乐厅中，17位评估员给出了60种不同的特性，除了三种以外，其他特性分为7个大组。最大的一组包含的特性有声音强度、包围感、洪亮程度。第二大组包含低音表现力和亲和力。第三大组包括清晰度和明快程度。

我们还要求听音者将9个音乐厅按照偏好排序。每一种按照偏好的排序或者按照单个属性的排序可以用多维空间的一个点代表。我们可以将属性、偏好、音乐厅之间最重要的区别可视化。根据偏好，评价者可分为两组。第一组偏爱的音乐厅是使亲和力的声音具有清晰和明快的特点。换言之，他们喜欢相对亲密的声音，在其中可以清楚地辨别每种乐器和美好旋律的乐句。第二组倾

图2 某位听音者在一个音乐厅中感受的随着方向和时间变化声能的显示。折线代表在直传声音到达后5 ms(淡灰)、30 ms(黑色)、100 ms(蓝色)以及2000 ms(红色)内接受的累积声能。曲线与音乐厅建筑图的顶部和侧面图相叠加，用来显示大厅建筑设计

之间的关系 (a)音乐厅FT的形状类似于扇形，低屋顶产生很强的顶部早期反射，但是来自侧面的反射很少。这两个特点造成听音者有音效来自遥远的感觉。听音者们并不喜欢；(b)VS是最让人喜爱的音乐厅之一，所属的情况正好相反



向于更高强度及洪亮的声音，更有鼓动性和强烈的低音。所有的评价者不喜欢音乐厅的声音柔弱而飘渺。令人吃惊的是，这两组评价者最感兴趣的均是感知的亲和力。

总的来说，最不受人喜欢的音乐厅的特点是声音柔弱而飘渺。具有高强度及洪亮的声音的音乐厅表现出缺乏清晰度和明快度，因此，它们不能表现最深情的声音。这些音乐厅只能排列在平均偏好排名的中间。而最令人欣赏的音乐厅具有足够的音效强度和包容性，良好的清晰度，以及能够表现深情音效的低音。

在每个音乐厅中，扬声器交响乐队的效果都严格地在12 m位置录制。那么，为什么感知的距离随不同音乐厅有如此大差异？我们最近发展了一种技术，能够对累积的声能进行时空可视化，或许可以帮助我们理解这一差异。使用我们在每个音乐厅中所测量的扬声器乐队数据，运用相同的空间脉冲响应数据，我们将声能随方向和时间变化

的到达时间绘成图表。如果将这个图表与音乐厅的建筑平面图叠加，可以确定哪些反射面使得声音的到达比从音乐厅前方到达得晚些。

图2为两个音乐厅的时空分布图。音乐厅FT的建筑结构新颖，产生的声音感觉遥远，是最不受偏爱的音乐厅之一。音乐厅VS则不同，它能够传递最为亲情的音响，是听音者最喜欢的音乐厅之一。图2中的粗黑线代表累积声能在接受了直接发出声音之后30 ms内到达。这说明FT大厅中绝大部分声音来自于向前传播的声音，而VS大厅中存在墙壁强烈反射的声音。另外，FT大厅具有明显的来自低屋顶的反射，而VS大厅则仅有少量早期垂直反射(这是来自于舞台上方的反射，不是来自于屋顶)。我们发现这两座大厅所代表的重要类型：早期横向反射提供更为温柔的音效，而强烈的早期屋顶反射使人感到声源来自远方。原因是来自两侧的反射被放大得比来自顶部反射更强，特别是高频段。