

音乐可视化设计中的多通道映射模式研究

作者：张屹南

发表于《设计艺术研究》.2014 年第 4 期：58-63

摘要：

本文的研究对象是可适用于演出现场的实时音乐可视化设计。由于音乐本身具有丰富而微妙的情感化信息，在音乐可视化过程中，如何通过多通道映射模式以增强可视化作品的人性化特质是本文的研究目的。通过对视听映射基本形式的梳理分析、对多通道映射方法的研究以确定实验方法。通过建立实验模型的手段具体探讨不同模式的多通道映射。结论中明确了将表演者或操控者对音乐主观理解所转化成的动作和姿势作为多通道介入因素的视听映射设计方法。

关键词：

数字媒体设计；音乐可视化；新媒体；视听映射；视觉音乐；

The multichannel mapping systems in music visualization design

Abstract:

The research background is realtime music visualizations for performative environments. The aim is to setup multi-channel audiovisual mapping systems to enhance the emotional and intuitive aspects of music visualization work. The experimental methods come from analysis of audiovisual mappings. Several experiments have been carried out to set the researching prototypes. As conclusion, the multichannel mapping systems rely on involvements of physical gestures sensing by the performers or manipulators as representations of intuitive methods.

Keywords:

digital media design; music visualization; new media; audiovisual mapping ; visual music

引言

音乐可视化（Music Visualization）是指将音乐中传达的信息以视觉形式进行演绎的过程。在数字媒体设计领域对于音乐可视化的研究有助于搭建视听媒介间的桥梁，以跨界的方式拓展动态媒体表现型式。由于音乐本身具有丰富而微妙的情感化信息，在音乐可视化过程中，很难以单一的法则去机械地将其转换成视觉，如何通过多通道映射模式以增强可视化作品的人性化特质是本文的研究目的。

音乐可视化设计中的多通道映射模式研究

1. 音乐可视化设计中的视听映射

1.1 音乐可视化设计的概念与应用

音乐可视化设计的应用领域较为宽泛，在视觉研究、声音研究、表演与展示、交互设计、音乐软件界面设计、游戏设计等方面均有重要的理论与应用价值。瓦西里·康定斯基（Wassily Kandinsky）是研究绘画中的音乐特性的代表人物。康定斯基由绘画语言的分析入手，提出了绘画中的“内在声音”（Inner Sound）这一概念。^[1]而在现代视觉设计领域，音乐可视化属于信息可视化的范畴。由于音乐是一种特殊的声音信息，如何将这些信息以视觉图形图像形式进行呈现以准确演绎出音乐独有的特征，是研究的关键所在。音乐可视化设计通过艺术与技术的融合以增强人类的视听感官体验，是一门与时俱进的学科。如今设计师们不再一味地追求娱乐体验，而是从音乐信号的内部结构开始分析，使得作品像是连接视觉表达与音乐符号的桥梁，使生涩的音乐波形图变得生动有趣。^[2]

在现代舞台演出中，用现场影像或动画来伴随音乐进行可视化艺术表现能创造出综合视听媒介。在 20 世纪 70 年代诞生了 VJ 这一名词，用以称呼那些专在聚会、主题活动或演出现场提供影像的人。VJ 是“影像骑师”（Visual Jockey）的缩写。VJ 艺术在广义上可用来描述任何实时操控的视频表演，主要目的是为现场观众提供与音乐表演相同步的视觉体验。本文的研究对象主要是可适用于 VJ 演出现场的实时音乐可视化设计。

1.2 视听映射的概念与类型

在音乐可视化设计中，声音与视觉信息之间需建立起某些属性上的关联，以实现信息的转化与生成，这一关联手段称为映射（Mapping）。映射是建立在拥有可供对位匹配的数据源的基础上，先将这些数据源以某些预设的法则进行对应，再将其连通以实现信息之间单向、双向或多向的传输和干预。映射的对象是指完成映射活动的信息载体，一般将特定媒介中的可量化属性作为映射对象，这些属性需要具有线性的可变区域，由此每个对象之间在变化范围内可以进行对位匹配。

自然映射（Natural Mapping）是通过自然界中物理手段实现的视听媒介传递与转换，又称物理映射。比如 17 世纪德国音乐家恩斯特·克拉尼（Ernst Chladni）发现，如果用小提琴的弓去敲打铺满沙粒的铜盘，根据弓在铜盘边缘敲打位置的不同，能产生不同形状的几何图形，这些图形现在被命名为克拉尼图形。^[3]也有不少人尝试用声波的振动来驱动水面或沙粒形成共振以生成图案化视觉效果，在西方这种声音的物理可视化学科被称为“Cymatics”。

直觉映射（Intuitive Mapping）是指设计者根据自身的感受和主观判断制定信息间的对位方式，体现了创作者自身情感与经验的主导。研究直觉映射规律时常需要从心理学的角度来分析。“通感联觉”（Synesthesia）是指人们的感官器官，即视觉、听觉、嗅觉、味觉、触觉之间的相互感通，对一种感官器官的刺激可以引起其他感官的感觉。基于视觉和听觉的通感方式，或被称为“视听通感”。1927 年，心理学家沃尔夫冈·科勒（Wolfgang Köhler）做了相关的测试，成为“语音符号学”（Phonetic Symbolism）的经典案例——他在给出两张抽象图片于测试者后问：“两张图中哪一张代表声音 Maluma、哪一张代表声音 Takete？”（图 1）几乎所有的被测者都能给出相同的答案，将两张图分别与其中一个声音进行匹配。^[4]这些理论说明了人类在处理视听语言之间的关联性时具有某些共通的心理倾向。直觉映射有较为人性化、情感化的特质，能较好地体现出人对于音乐的理解。

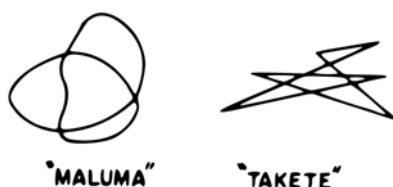


图 1 语音符号学图片案例

逻辑映射 (Logical Mapping) 是指根据预先设置的逻辑关系来判断信息间的匹配规则。在艺术表现中,每一种媒介都有自身发展的逻辑规律,而不同媒介在起承转合间能寻找到蛛丝马迹的相似点以相互匹配。比如音乐中声音的音高、音强、音色、旋律、节奏等要素,与视觉感受中的颜色、亮度、形状、体积等特质间能建立起丰富的对位关系。逻辑映射虽然没有“放之四海而皆准”的法则,但针对人类一些共通的感受而言是可以总结出规律的。比如:人们对音乐的频率高低之于颜色的亮暗之间具有本能的对位判断,频率较高的声音让人产生较亮的视觉意象,频率较低声音产生较暗的视觉意象。再比如,根据人的情态变化进行基于运动关系的对位映射:“高涨”、“低落”,“强烈”、“缓和”,“紧张”、“松弛”等等心理活动都可以物化为相应的视听语言特征。

数字化映射 (Digital Mapping) 是指利用数字技术以及计算机算法实现的视听对位方式。数字化映射并非独立于之前几种映射方式的全新模式,而是将数字媒体所特有的信息处理、储存和运算方式运用于不同的映射方法中。数字音乐可视化常见的例子是数字音频播放器中的可视化动画演绎,比如使用 Nullsoft 公司的 Winamp Advanced Visualization Studio 以及苹果公司的 iTunes Visualizer 等软件插件,可通过软件算法实时生成音乐动画。动态的数字化映射其基本原理是首先提取出数字视听媒介中关键信息的参数数据,并将这些数据的变化范围进行动态匹配:比如将声音的音量的数据变化范围(0—127)分别匹配到图像透明度的数值变量(0—255),当音量为 0 时透明度为 0,音量递增至 127 时透明度递增至 255。音乐可视化中的数字映射包括数字音频信号分析与数字图形图像生成两个步骤。

1.3 多通道映射之研究方法

由于音乐本身具有丰富而微妙的情感化信息,在音乐可视化过程中,很难以单一的法则去机械地将其转换成视觉。况且人的感官对于视觉的敏感程度又超过听觉,因此如何创造出生动且具有表现力的视听转换模式是音乐可视化设计的难点所在。瓦尔特·本雅明 (Walter Benjamin) 在论文《机械复制时代的艺术》中提出过,艺术作品特殊性的“韵味”(aura)在机械复制的时代消失了。^[5]这很容易让人联想到许多媒体播放器的音乐可视化动画,其无法准确地体现出不同音乐的独特审美特性。

音乐可视化设计中的多通道视听映射是指在声音与视觉信息之间建立起多重的映射通道,使不同映射属性之间相互影响而实现综合的视听效果,并发挥出不同映射手段间的互补性。多通道映射的匹配对象在数量上可分为三类:一对多映射、多对一映射和多对多映射。一对多映射(发散映射)是指一个控制变量影响多个参数;多对一映射(收敛映射)是一个属性参数被多个不同的控制器输入参数所影响;多对多映射是指匹配对象之间具有较为复杂的对应关系,一种属性不仅能影响多种属性参数,其自身同时被多种不同属性控制。

本文所探讨的重点问题是如何将计算机预设的可视化程序与人的情感因素相结合,如利用动态捕捉、电脑界面操控、电子乐器界面操控等数字交互技术实现多通道视听映射的设计方法。将人所具有的情感化因素有机融入数字化作品进程中,能增强作品的自然性和灵活性,实现理性因素与感性因素的平衡。新媒体艺术先驱罗伊·阿斯科特 (Roy Ascott) 认为,不应低估化学在创作中发挥的作用,大脑的化学组织为艺术传递了丰富的含义,这种作用可能不亚于认知学以及超出生物边界的领域的意义。^[6]因此,将人工操控作为一重或多重通道介入预设的软件程序中,形成具有部分人工介入因素的半自动系统,对于研究多通道视听映射具有启发意义。

笔者近年来利用开源编程平台openFrameworks制作了一系列基于多通道映射的音乐可视化交互模型，并与多位音乐演奏家、肢体表演艺术家展开跨界合作，以探讨这些案例所带来的衍生意义，分析新媒体交互界面是如何在人工与非人工因素间建立起信息传递的桥梁，以促进人性化艺术表现。在研究中多通道映射的设置方式为：将第一层映射通道设置为现场音频对视觉的直接影响，比如基于声音音量和频率的数据采集，并实时与视觉元素的属性（比如亮度、运动速度等）进行映射；而第二层映射通道是基于对现场人员的主观操控数据的采集，其中主要分为三个类型：（1）电脑界面操控；（2）乐器演奏触后姿势操控；（3）舞蹈等肢体动作操控。这些人工操控数据对于可视化生成具有干预作用，操控数据的不同属性可以同时介入控制多重的视觉属性。

2. 基于多通道映射的实验模型

2.1 电脑界面操控与音乐共同参与映射

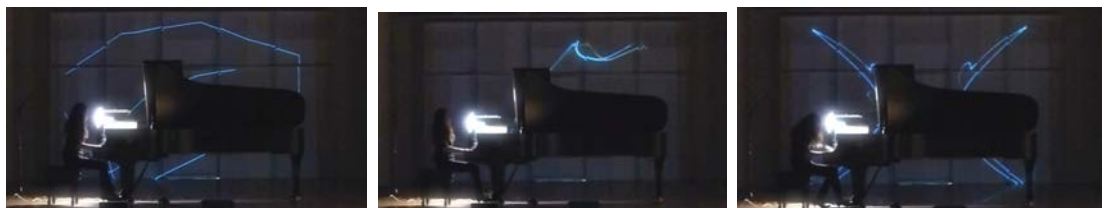
在本组实验中采用钢琴演奏的现场音乐作为可视化对象，在提取现场音频中的音量参数同时通过使用鼠标、键盘、触摸屏等人机交互界面进行介入，共同实现实时可视化动画效果。

实验模型一：鼠标与音量共同驱动的路径跟随动画

本实验是为钢琴家现场演奏的巴赫第五号英国组曲（English Suite No.5 in E Minor）所制作的实时可视化动画（图2）。笔者使用开源编程工具openFrameworks开发了一套利用鼠标方位控制与音乐音量控制结合生成的路径跟随动画原型。该动画的基本视觉形态为几条几何曲线所构成的图形，每条曲线分别由一系列互相跟随的长方形节点构成。

动画路径生成算法如下：现场钢琴音量的分贝值通过麦克风接收输入电脑，将音量变量同时控制线条顶点运动的加速度变化以及线条的亮度和宽度；鼠标的移动位置代表着曲线顶点移动的方向；曲线顶点始终会跟随着鼠标所在的位置做减速运动。

在演出中可以由后台操作者跟随着现场音乐带来的感受操控曲线的跟随路径，这种操控过程接近于舞蹈者跟着事先排演过的音乐进行舞蹈动作一般，由手动操控路径的舞动。相比单纯跟随鼠标路径运动的动画而言，这一动画原型将声音的反馈与可控的路径相结合。由于音量变化实时控制着路径运动的加速度，因此在对钢琴等乐器发出的颗粒感较强的声音效果能产生比较理想的视觉变化，尤其当演奏者弹奏某一单音时，几何曲线会骤然加速并慢慢减速，这一运动速度上的骤变与渐变的结合能创造出比较自然生动的动态效果。



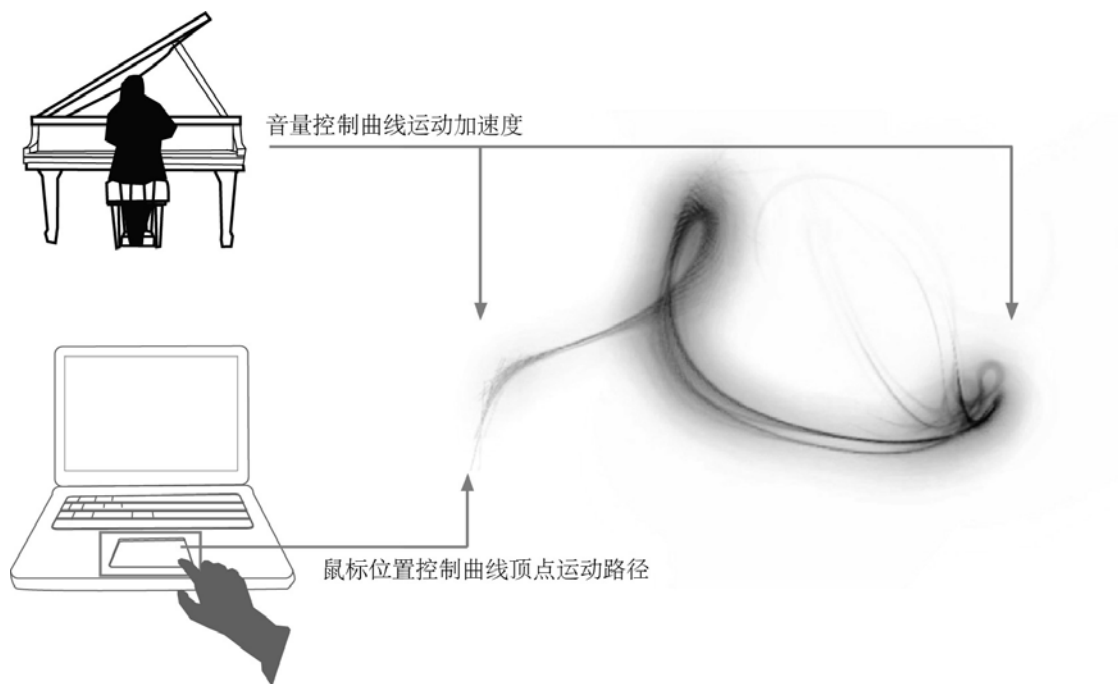


图 2 实验模型一：现场效果及映射模式

实验模型二：以按键实时切换群集动画的整体运动属性

在另一组钢琴演奏的可视化实验中，笔者将先前的控制模式进行了调整。动画的视觉主体仍是可以感应音乐音量的动态曲线，多条曲线呈现群集状出现在画面中，为体现该群集动画的生动性，该程序由人工通过电脑键盘切换线条整体的运动属性和线条的出现数量（图 3）。比如当按下按钮 A 键，会有两条线条根据现场音量从画面底部游动出现，而当再次按下 A 键时会有另外两条曲线出现，由此重复便可根据音乐进程逐渐增加曲线数量。而另两个按钮 D 和 F 用来切换曲线整体的运动模式，当按下 D 时每条曲线各自遵循不同的路径运动，而按下 F 时所有曲线同时向一个固定坐标运动，此效果也适合根据音乐的具体进程来操作，比如在弹奏到某一个强音的时候，将曲线的运动忽然切换成整齐划一，会产生较为强烈的视觉冲击。在这一交互原型的应用演示中，将投影影像直接投射在了钢琴表面以及背景的白色幕布上。当人工切换配合音乐音量的时候，能产生线条动画从琴身内部钻出的效果。



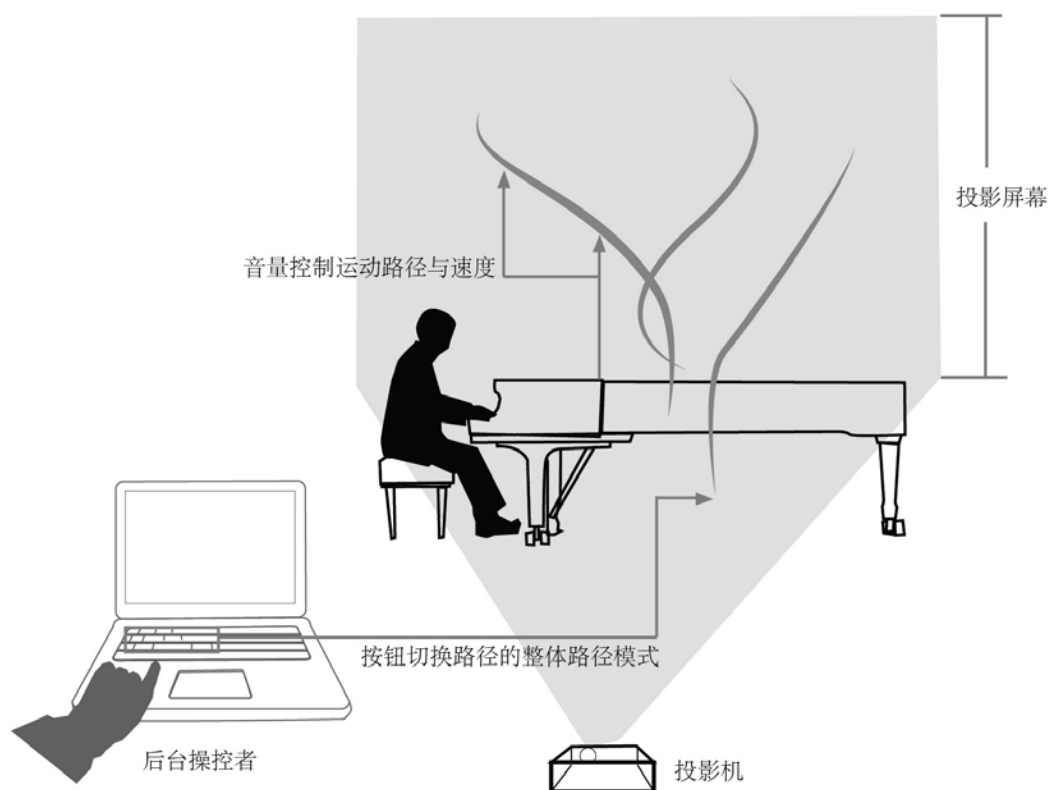


图 3 实验模型二：现场效果及映射模式

2. 2 乐器演奏姿势与音乐共同映射

实验模型三：MIDI 键盘与 Kinect 摄像头结合的交互实验

从演奏乐器的角度，身体姿势行为是创造声音和视觉的重要手段，该实验模型试图以演奏者演奏键盘乐器的触控与身体姿势作为信息源共同决定可视化效果。

在演奏乐器时演奏姿势可分为有效姿势和辅助姿势，有效姿势指触发声音的身体运动，如手指触碰琴键、手臂牵引琴弓拉动琴弦等等。辅助姿势是指用以支持和辅助有效姿势的一系列身体运动，比如摇摆躯干、摇头、脸部肌肉运动、深呼吸等等。本实验试图利用演奏姿势中的辅助姿势来作为改变可视化效果的信息参数。乐器演奏中的“触后姿势”是指触碰了琴键或其它乐器界面的下一个瞬间出现的肢体行为，比如弹压钢琴琴键之后抬手的动作就是自然的触后动作。将触后姿势作为信息参数动态捕捉入电脑中可以用来进一步操控先前由于弹奏而生成的动画。

该交互原型使用 MIDI 键盘的演奏与实时手势共同操控音乐与可视化动画的生成。这一系统可以让参与者在演奏音乐音符的同时在前方的屏幕上生成动态符号，并用手势操控这些符号并以此继续改变音乐的进程。（图 4）

在该实验中大致做了如下几种步骤的逻辑设定：首先触发琴键生成音乐，同时生成视觉几何形态，然后以触后姿势中手臂的抬起高度来控制音符及其可视化图形的后续进程。这套原型是由 Kinect 红外线摄像头捕捉手部的运动路径，同时将 MIDI 键盘的信号传输至电脑中。

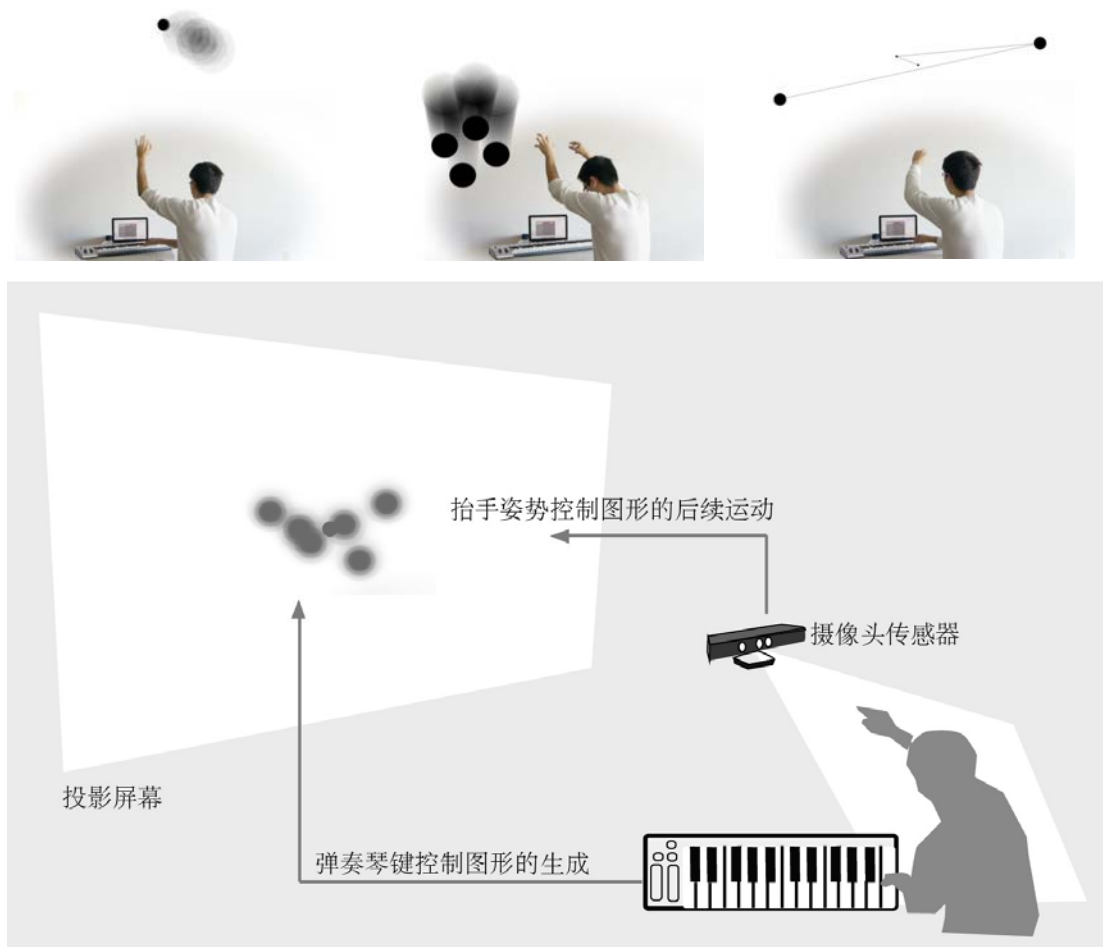


图 4 实验模型三：现场效果及映射分析

2. 3 肢体表演与音乐共同映射

在舞台演出中，表演者无论是在台上演奏乐器还是表演舞蹈或进行语言交流，都具有比日常生活中更为丰富和生动的身体运动信息，可以说身体是表演过程中最为重要的信息传播媒介。

《Messa di Voce》是新媒体艺术家 Golan Levin 和 Zach Lieberman 在 2003 年创作的作品，该作品较好地诠释了表演者肢体运动与声音配合的多通道映射。整个作品以舞台表演的形式展现，两位抽象人声艺术家在现场发出的讲话、叫喊声和歌声等被实时可视化合成为平面动态图形作为场景中的视觉增强介质。并且演员的肢体动作也会被摄像头实时捕捉以确保动画出现的位置与演员身体的位置重合，这样能产生可视化图形是直接从演员体内发散出来的效果。

笔者的研究侧重于探索人的肢体语言所具有丰富的情感信息，将演员在听到音乐的同时所做出的肢体直觉反馈作为映射的内容，并将这些信息与音乐结合以增强可视化视觉的层次感与表现力。

实验模型四：

笔者于 2013 年 7 月与两位奥地利艺术家和合作的作品《Art of Reading Map》，其中使用了将表演者的身体动作与现场音乐共同作为可视化影像内容的交互模式。画面中整体视觉元素是由密布着的亮线条组成。现场的电子音乐表演者操控电子合成器发出高频音与低频音的组合，音频的音高变化由 FFT（快速傅立叶变换）算法识别，由此来控制画面中线段的抖动频率和幅度。画面同时被表演者的肢体剪影所干预，表演者进入空间通过肢体表演在画面中形成一个空洞的剪影。

实验模型五：

《对话》是笔者与上海音乐学院音乐工程系合作的新媒体音乐剧演出，作为第十五届上海国际艺术节扶持青年艺术家计划委约作品。其中的新媒体影像部分通过将舞者实时剪影形象的异化动画投射在悬幕上，并附加延迟等视觉效果。（图5）通过现场古琴演奏音乐的音量来控制舞者剪影形态的模糊程度。从映射通道而言，舞者根据古琴音乐起舞，作为第一层通道控制视频内容，而音乐的音量与频率对视频内容的亮度、模糊度和延迟数值产生多重介入。



图5 实验模型五：现场效果

3. 结论

本文简要概括了音乐可视化设计中视听映射的基本类型，并针对其中的直觉映射和逻辑映射类型以数字化的型式探索两者的融合方式，在可视化过程中以多通道模式共同介入作品的生成。

以操控电脑界面来介入实时的音乐可视化呈现，其操作目的可分为确定动画运动方位和切换动画生成模式两类。前者适合用鼠标或触摸屏等指向性输入设备来操作，其本质是将人对音乐的主观理解转化为手指的肌肉运动，并将运动信息数字化映射为图形的动态参数；后者适合用键盘按键切换来操作，是人为选择和改变预设程序的方法。设计过程中需特别注意的是关于映射参数间的影响范围，比如多强的音量变化能产生多明显的效果改变，这都需要根据实际情况进行精确的现场校准。将乐器演奏姿势与音乐共同实现视听映射，该模式能发挥出乐器演奏中肢体运动信息的复杂性，并将这些触发出音乐的运动信息进行再度利用，使身体姿势和音乐之间的因果循环关系在视觉上得到延伸。表演者针对音乐的直觉反馈通过身体的运动与姿势反映出其内心的变化，在这层通道的基础上附加数字化特效，是一种以人工因素为主导的可视化模式。其作品的进程具有即兴表演的不确定性，正如本雅明所述：新媒体舞台艺术的复制却非传统艺术意义上的“复制”或“表现”，互动使得每一次复制都是一次全新的创作。

多通道映射模式的具体设计方法源于设计者自身建立的审美体系，只有将设计师对音乐的理解融入视觉创作中，才能真正将艺术与技术融会贯通。有一些转换模式并不能很好地反应出原始音乐信息的动态特征，在某些映射模式中原始信息甚至是无关紧要或是根本不会被观众察觉。很多情况下，可转换数据本身并不是作品中首要的问题所在，而是通过这些数据的传输使得一些数据流的趣味点更好地被人们理解和体验，或是用一种新的方式被感知到。^[7]因此创作者需要额外注意作品得美学因素以及交互过程的可识别性。另外，创作中应避免对视觉特效的过度解读。比如有时候视觉的存在反而削弱或干扰了音乐本身的表现力。一些优秀的VJ艺术家被称为“动态影像的炼金术士”（Moving-image alchemists）^[8]，因为他们能使两种媒介间产生化学效应，给原本已令人产生审美疲劳的后现代文化注入新的生命力。

- [1] 康定斯基. 康定斯基论点线面 [M]. 北京:中国人民大学出版社, 2003: 5
- [2] Golan Levin. Audiovisual Software Art: A Partial History[EB/OL]
www.flong.com/texts/essays/see_this_sound_old/.2009
- [3] 维基百科. “恩斯特·克拉德尼” 词条.
http://en.wikipedia.org/wiki/Ernst_Chladni. 2014
- [4] Niall Moody. Ashitaka: An Audiovisual Instrument[D]. 英国. 格拉斯哥大学. 2009
- [5] 瓦尔特·本雅明. 机械复制时代的艺术[M]. 重庆: 重庆出版社. 2006
- [6] 罗伊·阿斯科特. 未来就是现在: 艺术, 技术和意识[M]. 北京: 金城出版社, 2012:271
- [7] Golan Levin. Audiovisual Software Art: A Partial History[EB/OL]
www.flong.com/texts/essays/see_this_sound_old/.2009
- [8] Michael Faulkner. audio-visual art+vj culture [M]. 英国 :Laurence King Publishing. 2006:70