# 2.2 三维场景生成

三维场景生成最关键的部分是场景理解与场景合成。前者需要考虑大量的场景信息，选出合理的对象类别，后者则需要适当的将选出的对象布置在场景中。三维场景可以说是虚拟现实的重要基础，一个良好的三维场景能够极大的提升虚拟现实技术带给使用者的体验。因此，三维场景的生成技术在不断的推陈出新，效果逐渐提升，但其本身的复杂性也带来了许多的挑战，如仿真场景的真实性、虚拟场景的合理性以及场景本身的美化程度，都需要大量的人力、物力投入。目前，Chat Gpt及其带来的AIGC技术给场景理解和场景合成都带来了新的发展。

## 2.2.1 传统技术

传统的三维场景生成技术中，人机交互方法占据了主要的地位，场景理解主要由开发者完成，借助3D建模软件或是一些专用软件进行自动或是半自动的场景合成工作。

传统技术在三维场景生成方面有着广泛的应用。美国伊利诺伊大学的Fard、Mani等人将BIM技术和摄影影像技术相结合，利用软件处理数据资料以完成3D施工的模拟效果[1]；基于三维软件sketchup可建立小型建筑场景，实现建筑场景的三维可视化表达；许多大学，也采用全景技术搭建校园场景，方便新生及其他社会人士，在任何地点通过虚拟现实技术了解校园场景。

但是，我们可以看到，传统技术下，开发者需要承担许多重复、繁琐的场景搭建工作，且场景搭建的效果与开发者对场景的理解有着相当大的关联，有着极大的潜在风险。与此同时，通过这些方式生成的三维场景大都是静态的，与用户存在的交互很少且体验效果较为机械。目前的虚拟现实市场对于更优的交互、更真实、动态变化的场景有着较高的需求。而动态化的三维场景搭建，需要新的技术或方法来提供更好的解决方案。

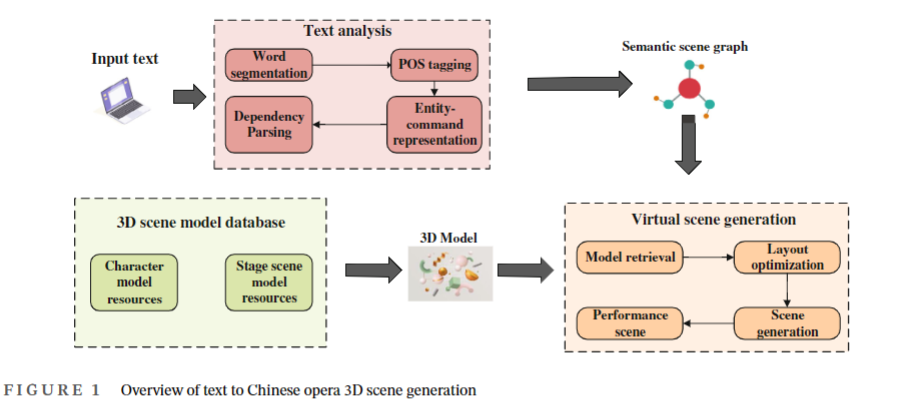
## 2.2.2 AIGC应用于三维场景生成

ChatGpt 表现出的强大学习能力、语义理解能力、上下文记忆对话能力带来极大震撼的同时，也让我们看到了其所带来的AIGC技术的强大生产能力。三维场景生成中的场景理解可由AIGC完成，且同步自动进行后续场景合成等，完成三维场景生成工作。AIGC技术给三维场景合成带来了新的发展动力。

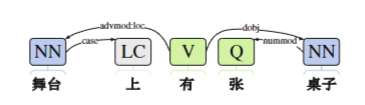
### 2.2.2.1 一种基于语义的戏曲表演生成方法[2]

来自郑州轻工业大学、英国伯恩茅斯大学的梁辉等人，将AIGC技术应用于三维场景生成，提出了一种基于语义的中国戏曲表演生成方法。该方法可依据输入的描述性脚本自动生成戏曲表演的三维场景，首先通过分析场景的描述脚本了解戏曲虚拟场景的构成要素，然后利用先验概率学习场景模型中的模型放置规则，最后生成出与脚本对应的戏曲表演的数字场景。该方法被证明可以生成自然的可接受的歌剧数字表演场景。

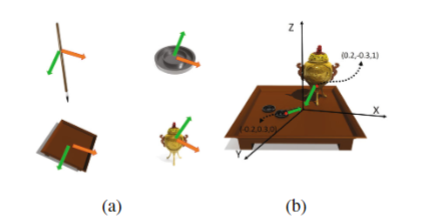
该方法由三部分组成：由戏曲数字资产三维模型本体库支撑的数据库；由用户输入脚本理解出具有空间关系的语义场景地图；由语义场景地图从数据库中提取模型并进行场景搭建。流程图如下：



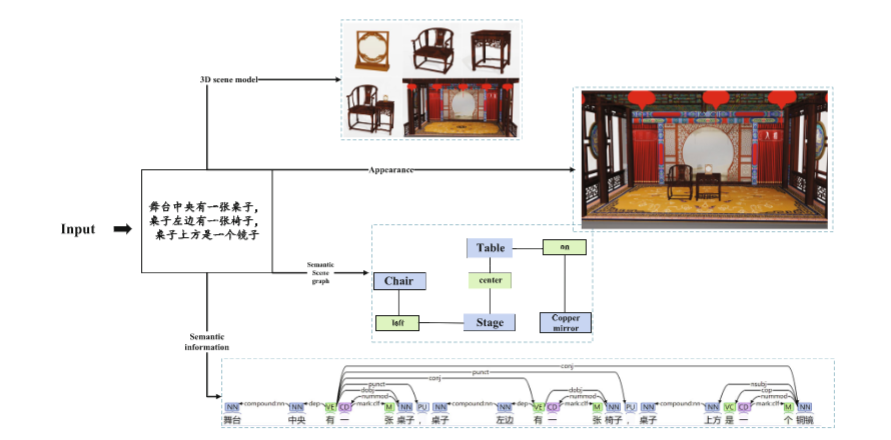
场景理解方面，该方法的输入脚本处理分为两步完成。第一步是利用隐马尔可夫模型将输入的描述脚本划分为独立的虚拟场景元素，然后进行顺序标注，实现元素分割。第二步是进行语义统计及独立元素分类，将输入的文本信息理解成空间语义信息，如下图：



场景合成方面，该方法对数据库中的每个模型都添加了名称、属性、向前方向、大小、旋转和质心坐标等语义标记，将语义标签与数据库中的模型相关联。在此基础上，就可基于前述的的空间语义信息生成具体的三维模型。



一个较为详细的三维场景搭建示例如下：



通过此方法，我们可以看到，良好的场景理解可以为自动的场景合成提供优良的基础。

### 2.2.2.2 SkyBox AI

基于生成对抗网络模型的 Skybox AI 在自动三维场景生成方面也进行了积极的探索，Blockade Labs 通过基于Skybox AI 的在线网站为用户展现了Skybox的能力，用户可通过输入关键字快速获取目标场景的360°全景图像。

生成对抗网络是一种由生成模型和判别模型组成的深度学习模型，在图像处理方面有着强大的作用、广泛的应用。目前最新出现的DragGan模型在AI绘图场景下表现出了强大的能力，其可以让用户通过拖动图片中的点来操纵图像中如动物、汽车等元素使之表现出不同的姿态、朝向等。

我们可以看到，深度学习技术在场景合成方面有着巨大的潜力。

## 2.2.3 展望

可以期待，如ChatGpt 的大型语言模型的语义理解能力会不断增强，可以准确、高效的根据三维场景的需求完成场景理解工作，深度学习技术的发展会让场景合成不再需要大量的人力投入。AIGC技术将会使三维场景生成工作高度自动化，用户只需输入他的需求，便可得到与之对应的精细的三维场景。开发者的工作将由具体的场景搭建，转变为建立三维场景自动搭建的框架。

但同时，自动化的生成工作也势必带来不可预知的问题。一方面，由于生成内容完全依赖于用户输入，就有可能导致结果同质化，缺乏多样性，也可能导致生成场景的效果无法达到用户需求，缺乏可控性。另一方面，用户的输入决定输出，也意味着自动生成场景的结果在很大程度上是不可控的，当下已经出现利用AI技术生成假视频的现象，如何将自动生成的结果控制在社会道德乃至法律范围之内将是一项必要工作。

言而总之，ChatGpt及其带来的AIGC技术给三维场景生成技术的发展带来了新的动力，自动化的三维场景生成已成为一种发展趋势。未来，如何实现更精细的场景理解，如何利用深度学习技术实现更好的动态化场景搭建，以及如何进行有效的监管，将有很大机会成为热点话题。

**参考文献**

[1] 鲁鑫鑫. 虚拟三维场景的搭建与可视化动态展示[D]. 山东大学软件工程, 2021.

[2] Liang H, Dong X, Liu X, et al. A semantic-driven generation of 3D Chinese opera performance scenes[J]. Computer animation and virtual worlds, 2022(3/4):33.

**地址：**

虚拟三维场景的搭建与可视化动态展示

[**https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMzAxMTISCFkzODE0OTQyGghvamQ2OHY2OA%3D%3D**](https://d.wanfangdata.com.cn/thesis/ChJUaGVzaXNOZXdTMjAyMzAxMTISCFkzODE0OTQyGghvamQ2OHY2OA%3D%3D)

Liang H, Dong X, Liu X, et al. A semantic-driven generation of 3D Chinese opera performance scenes[J]

[**A semantic‐driven generation of 3D Chinese opera performance scenes - Liang - 2022 - Computer Animation and Virtual Worlds - Wiley Online Library**](https://onlinelibrary.wiley.com/doi/10.1002/cav.2077)