

同行专家业内评价意见书编号: 20250854377

附件1

**浙江工程师学院（浙江大学工程师学院）
同行专家业内评价意见书**

姓名: _____ 宜逸昕

学号: _____ 22260380

申报工程师职称专业类别（领域）: _____ 电子信息

浙江工程师学院（浙江大学工程师学院）制

2025年03月14日

填表说明

一、本报告中相关的技术或数据如涉及知识产权保护、军工项目保密等内容，请作脱密处理。

二、请用宋体小四字号撰写本报告，可另行附页或增加页数，A4纸双面打印。

三、表中所涉及的签名都必须用蓝、黑色墨水笔，亲笔签名或签字章，不可以打印代替。

四、同行专家业内评价意见书编号由工程师学院填写，编号规则为：年份4位+申报工程师职称专业类别(领域)4位+流水号3位，共11位。

一、个人申报

(一) 基本情况【围绕《浙江工程师学院（浙江大学工程师学院）工程类专业学位研究生工程师职称评审参考指标》，结合该专业类别(领域)工程师职称评审相关标准，举例说明】

1. 对本专业基础理论知识和专业技术知识掌握情况(不少于200字)

在研究生阶段，我系统深入地学习了人脸肖像合成相关的基础理论和专业技术知识。首先，我对计算机视觉、数字图像处理、机器学习以及深度学习的基本原理有了扎实的掌握，深入学习了神经辐射场以及三维高斯泼溅模型。其次，我通过大量的实验和项目实践，熟悉了TensorFlow、PyTorch等深度学习框架的使用，能够独立完成模型的设计、训练和调优，并针对人脸肖像合成的复杂特点提出针对性的解决方案。此外，我还关注最新的前沿技术与研究动态，能在传统方法基础上结合新兴算法实现更高效的肖像合成，为解决实际问题提供了坚实的理论与技术支撑。总的来说，我在本专业的的基础理论和专业技术方面均具备扎实的知识储备和丰富的实践经验，为今后的科研和应用开发打下了坚实的基础。

2. 工程实践的经历(不少于200字)

在为期一年的工程实践中，我的主要实验内容是人脸肖像合成以及驱动。这个实践的背景是在今年早些时候，苹果出了个MR眼睛，上面有一个Persona这个软件，通过扫描人脸肖像的视频能够实现肖像驱动和重建。实践目的是用单目视频拍摄一段视频，然后重建出一个完整的3D人脸，并且用另一个人的视频去驱动这个人的人脸。实践的结果是我完成预处理代码的编写，实现从数据采集到预处理到数字人构建到驱动完整流程，并且基于这个方向产出了一些成果。

3. 在实际工作中综合运用所学知识解决复杂工程问题的案例(不少于1000字)

实际工作和学校中的科研是有很大的区别的。对于这个软件流程，都是需要从零开始做，我首先需要在网上查询相关的资料和文献，找到与我项目类似的产品，并查找他们的技术文档。对于最核心的重建以及驱动步骤，我结合已有论文GaussianAvatars来展开我们的项目。但是这个项目有两个难点：第一是这个论文是基于多目视角的数据集，我们想做的就是基于单目采集的三维重建；第二就是这个论文没有公开数据预处理代码，就是从第一步的视频拍摄，到最后预处理到网络输入这部分都并没有。

对于第一个问题，顾名思义多目视频包含更多视角的人脸视频，其包含更多的信息，对于重建问题，视角越多，其重建效果越好。但是对于我们目前这个场景，其只包含一个正面视角，其相比于多目视频的信息更少，所以为了能在有限视角内重建出更好的效果，我提出了多个创新点来解决这个问题。首先，我提出了一个基于点的可学习初始化场，该初始化场的作用是来初始化高斯核模型，以便网络能够更好地拟合人脸集合结构。其次我设计了一个转换网络，用于更好地拟合人脸表情变化。最后为了提升肖像头发和肩部等边缘区域的渲染性能，我引入了Alpha渲染损失和结构损失。通过上述提出的几个方法，我设计的方法即使在单目视角下，也能有很好的渲染结果。在这个过程中我使用到了三维高斯泼溅模型相关的知识，以及神经网络结构等知识。

对于第二个问题，我独立设计了一整套预处理流程。首先是数据采集，我使用一个支架来固定单目摄像头，使其在采集的过程中保持相机位置不动；然后让一个人坐在摄像头前面，做出各种各样的表情；然后使用裁剪算法，将原来半身人像的视频裁剪成600*600像素的头部视频；使用背景分割算法分割原有视频的背景，并且获取数据的掩码图；然后使用基于FLAME框架的人脸身份提取算法去提取第一帧无表情的人脸的身份信息；最后使用表情追踪器来提取每一帧人脸的表情参数以及动作参数（眼睛方向，下巴方向，脖子方向），在这个过程中保持人身份系数不变。通过上述方法，我打通了从数据采集，到预处理，到输入网络训练

的整套流程。在这个过程中，我使用了三维人脸可形变模型，人像分割等方面知识。对于具体问题，例如我在讲预处理后的数据放入模型训练时发现训练结果很差，甚至都达不到原方法的效果。我通过分析训练过程发现是表情追踪器提取出来的参数与真实图片的对齐上有问题，我根据原版表情追踪器提取出来的参数发现其和图片对不齐，甚至有时候提取出来的参数构建的三维Mesh在高斯模型相机参数外面了。我仔细阅读了表情追踪器的代码，发现其使用的相机模型和三维高斯泼溅模型的渲染中的相机模型不太一样，表情追踪器原来使用相机模型是焦距相机，而高斯中使用的是正交相机，也就是使用了正交矩阵，我将原来的相机模型换成了使用正交坐标系的相机模型，并且将其中的参数都保持和高斯模型一致，这样解决了这个对不齐的问题。

(二) 取得的业绩(代表作)【限填3项,须提交证明原件(包括发表的论文、出版的著作、专利证书、获奖证书、科技项目立项文件或合同、企业证明等)供核实,并提供复印件一份】

1. 公开成果代表作【论文发表、专利成果、软件著作权、标准规范与行业工法制定、著作编写、科技成果获奖、学位论文等】

成果名称	成果类别 [含论文、授权专利(含发明专利申请)、软件著作权、标准、工法、著作、获奖、学位论文等]	发表时间/授权或申请时间等	刊物名称/专利授权或申请号等	本人排名/总人数	备注
基于三维高斯泼溅的人脸高保真度和可驱动的重建方法	发明专利申请	2025年10月01日	申请号: 202410723951.5	2/1	
Cvae-layout: automatic furniture layout with constraints	一级期刊	2024年03月14日	The Visual Computer	1/4	SCI期刊收录

2. 其他代表作【主持或参与的课题研究项目、科技成果应用转化推广、企业技术难题解决方案、自主研发设计的产品或样机、技术报告、设计图纸、软课题研究报告、可行性研究报告、规划设计方案、施工或调试报告、工程实验、技术培训教材、推动行业发展中发挥的作用及取得的经济社会效益等】

(三) 在校期间课程、专业实践训练及学位论文相关情况	
课程成绩情况	按课程学分核算的平均成绩： 85 分
专业实践训练时间及考核情况(具有三年及以上工作经历的不作要求)	累计时间： 1 年(要求1年及以上) 考核成绩： 83 分
本人承诺	
个人声明：本人上述所填资料均为真实有效，如有虚假，愿承担一切责任，特此声明！	
申报人签名： <u>宣逸昕</u>	

浙江大学研究生院
攻读硕士学位研究生成绩单

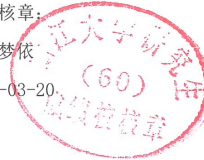
学号: 22260380	姓名: 宣逸昕	性别: 男	学院: 工程师学院	专业: 控制工程	学制: 2.5年						
毕业时最低应获: 25.0学分	已获得: 27.0学分			入学年月: 2022-09	毕业年月:						
学位证书号:			毕业证书号:		授予学位:						
学习时间	课程名称	备注	学分	成绩	课程性质	学习时间	课程名称	备注	学分	成绩	课程性质
2022-2023学年秋季学期	工程技术创新前沿		1.5	90	专业学位课	2022-2023学年冬季学期	产业技术发展前沿		1.5	84	专业学位课
2022-2023学年秋季学期	创新设计方法		2.0	通过	专业选修课	2022-2023学年秋冬学期	高阶工程认知实践		3.0	87	专业学位课
2022-2023学年秋季学期	工程伦理		2.0	85	专业学位课	2022-2023学年秋冬学期	研究生论文写作指导		1.0	78	专业选修课
2022-2023学年秋季学期	工程数值分析		2.0	88	专业选修课	2022-2023学年冬季学期	自然辩证法概论		1.0	98	公共学位课
2022-2023学年秋季学期	智能无人系统及应用实践		2.0	80	专业选修课	2022-2023学年春季学期	研究生英语基础技能		1.0	免修	公共学位课
2022-2023学年冬季学期	智能控制技术		2.0	88	专业学位课	2022-2023学年春季学期	研究生英语		2.0	免修	专业学位课
2022-2023学年冬季学期	新时代中国特色社会主义思想理论与实践		2.0	91	专业学位课		硕士生读书报告		2.0	通过	
2022-2023学年冬季学期	机器视觉及其应用		2.0	80	专业学位课						

说明: 1. 研究生课程按三种方法计分: 百分制, 两级制 (通过、不通过), 五级制 (优、良、中、及格、不及格)。
2. 备注中 "*" 表示重修课程。

学院成绩校核章:

成绩校核人: 张梦依

打印日期: 2025-03-20





国家知识产权局

310013

浙江省杭州市西湖区古墩路 701 号紫金广场 B 座 1103 室 杭州求是
专利事务所有限公司
傅朝栋(0571-87911726-812) 张法高(0571-87911726)

发文日:

2024 年 10 月 01 日



申请号或专利号: 202410723951.5

发文序号: 2024100100306670

申请人或专利权人: 浙江大学

发明创造名称: 基于三维高斯泼溅的人脸高保真度和可驱动的重建方法

发明专利申请进入实质审查阶段通知书

上述专利申请, 根据申请人提出的实质审查请求, 经审查, 符合专利法第 35 条及实施细则第 113 条的规定, 该专利申请进入实质审查阶段。

提示:

1. 根据专利法实施细则第 57 条第 1 款的规定, 发明专利申请人自收到本通知书之日起 3 个月内, 可以对发明专利申请主动提出修改。

2. 申请文件修改格式要求:

对权利要求修改的应当提交相应的权利要求替换项, 涉及权利要求引用关系时, 则需要将相应权项一起替换补正。如果申请人需要删除部分权项, 申请人应该提交整理后连续编号的部分权利要求书。

对说明书修改的应当提交相应的说明书替换段, 不得增加和删除段号, 仅只能对有修改部分段进行整段替换。如果要增加内容, 则只能增加在某一段中; 如果需要删除一个整段内容, 应该保留该段号, 并在此段号后注明: “此段删除” 字样。段号以国家知识产权局回传的或公布/授权公告的说明书段号为准。

对说明书附图修改的应当以图为单位提交相应的替换附图。

对说明书摘要文字部分修改的应当提交相应的替换页。对摘要附图修改的应当重新指定。

同时, 申请人应当在补正书或意见陈述书中标明修改涉及的权项、段号、图、页。

审查员: 自动审查

联系电话: 010-62356655

审查部门: 初审及流程管理部



210307
2023.03

纸件申请, 回函请寄: 100088 北京市海淀区蓟门桥西土城路 6 号 国家知识产权局专利局受理处收
电子申请, 应当通过专利业务办理系统以电子文件形式提交相关文件。除另有规定外, 以纸件等其他形式提交的文件视为未提交。



国家知识产权局

310013

浙江省杭州市西湖区古墩路 701 号紫金广场 B 座 1103 室 杭州求是
专利事务所有限公司
傅朝栋(0571-87911726-812)张法高(0571-87911726)

发文日:

2024 年 06 月 05 日



申请号: 202410723951.5

发文序号: 2024060501321790

专利申请受理通知书

根据专利法第 28 条及其实施细则第 43 条、第 44 条的规定, 申请人提出的专利申请已由国家知识产权局受理。现将确定的申请号、申请日等信息通知如下:

申请号: 2024107239515

申请日: 2024 年 06 月 05 日

申请人: 浙江大学

发明人: 潘宇, 宣逸昕

发明创造名称: 基于三维高斯泼溅的人脸高保真度和可驱动的重建方法

经核实, 国家知识产权局确认收到文件如下:

权利要求书 1 份 3 页, 权利要求项数: 10 项

说明书 1 份 12 页

说明书附图 1 份 3 页

说明书摘要 1 份 1 页

专利代理委托书 1 份 2 页

发明专利请求书 1 份 4 页

实质审查请求书 文件份数: 1 份

申请方案卷号: 傅-241-129

提示:

1. 申请人收到专利申请受理通知书之后, 认为其记载的内容与申请人所提交的相应内容不一致时, 可以向国家知识产权局请求更正。

2. 申请人收到专利申请受理通知书之后, 再向国家知识产权局办理各种手续时, 均应当准确、清晰地写明申请号。

审查员: 自动受理

联系电话: 010-62356655

审查部门: 初审及流程管理部



200101
2023.03

纸件申请, 回函请寄: 100088 北京市海淀区蓟门桥西土城路 6 号 国家知识产权局专利局受理处收
电子申请, 应当通过专利业务办理系统以电子文件形式提交相关文件。除另有规定外, 以纸件等其他形式提交的文件视为未提交。

Home > The Visual Computer > Article

CVAE-LAYOUT: automatic furniture layout with constraints

Original article | Published: 22 December 2023

Volume 40, pages 7731–7745, (2024) [Cite this article](#)



The Visual Computer

[Aims and scope](#) →

[Submit manuscript](#) →

Download PDF ↓

Access provided by Zhejiang University Library

Yixin Xuan, Chao Song, Jianqiu Jin & Bailin Yang

1022 Accesses 1 Citation [Explore all metrics](#) →

[Use our pre-submission checklist](#) →

Avoid common mistakes on your manuscript.



Abstract

We propose an automatic layout method for indoor scenes that effectively satisfies specific constraints. Our approach involves enhancing the existing scene representation method to accommodate complex constraints, including the precise placement of doors, windows, and user-specified furniture. To achieve this, we construct a conditional vector that encapsulates the necessary constraints. Moreover, our automatically constrained layout approach is implemented by training a conditional variational autoencoder model. Given the constraints and randomly sampled vectors, the decoder module can generate diversified reasonable indoor layout results. Evaluations show that our model outperforms the existing methods. Furthermore, our model exhibits a lower parameter count and faster execution speed compared with the existing approaches.

Sections Figures References

- [Abstract](#)
- [Introduction](#)
- [Related work](#)
- [Methods](#)
- [Results and discussion](#)
- [Conclusion](#)
- [References](#)
- [Funding](#)
- [Author information](#)
- [Additional information](#)



Cvae-layout: automatic furniture layout with constraints

Yixin Xuan¹ · Chao Song² · Jianqiu Jin² · Bailin Yang²

Accepted: 16 November 2023
© Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

We propose an automatic layout method for indoor scenes that effectively satisfies specific constraints. Our approach involves enhancing the existing scene representation method to accommodate complex constraints, including the precise placement of doors, windows, and user-specified furniture. To achieve this, we construct a conditional vector that encapsulates the necessary constraints. Moreover, our automatically constrained layout approach is implemented by training a conditional variational autoencoder model. Given the constraints and randomly sampled vectors, the decoder module can generate diversified reasonable indoor layout results. Evaluations show that our model outperforms the existing methods. Furthermore, our model exhibits a lower parameter count and faster execution speed compared with the existing approaches.

Keywords Conditional variational auto-encoder · Interior design · Automatic layout · Deep learning

1 Introduction

The demand for automated indoor scene generation has witnessed a significant surge in recent decades, driven by applications, such as general 3D content creation [1] and virtual reality technology [2]. This technology not only serves as a valuable tool for artificial decoration design but also has the potential to replace traditional manual design processes. The task of scene generation holds a crucial position in the domains of computer vision and graphics.

It is the diverse distribution of furniture that make the task challenging. Furniture, characterized by variations in location, size, and orientation [3], forms a multitude of semantic relationships within a given scene. Additionally, there exist generic geometric constraints, such as maintaining appropri-

ate gaps between furniture pieces, further complicating the scene generation task.

In the early research, the approaches to indoor scene generation could be broadly categorized into rule-based methods and learning-based methods. Rule-based methods relied on pre-constructed generative grammars and formulated mathematical optimization models with rule constraints. However, with the advancement of neural networks and the availability of annotated 3D scene datasets, learning-based methods have gained increased reliability and robustness in scene analysis. Our approach belongs to the latter.

Our work draws inspiration from the GRAINS framework proposed by Li et al. [4], which introduces a hierarchical representation for indoor scenes and enhances a generative recursive autoencoder. This representation encompasses object representation, relative positions, and semantic relations between scene objects. Building upon this foundation, Li and her colleagues trained a variational recursive autoencoder specifically for indoor scene generation. This model exhibits the ability to generate 3D indoor scenes with ease and high efficiency.

However, there are three limitations: (1) The generated indoor scene is relatively random, which means the size of the room is unfixed; (2) GRAINS removes the windows and doors when pre-processing the dataset, which is not reasonable in reality; (3) The generated indoor scenes cannot meet the requirements included type and quantity of furniture and other constraints given by users.

✉ Chao Song
csong@zjsu.edu.cn

Yixin Xuan
xyx18061932@gmail.com

Jianqiu Jin
jqjin@zjsu.edu.cn

Bailin Yang
ybl@zjgsu.edu.cn

¹ Polytechnic Institute, Zhejiang University, Hangzhou 310015, Zhejiang, China

² School of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou 310018, Zhejiang, China

经检索《Web of Science》、《Journal Citation Reports (JCR)》及《中国科学院文献情报中心期刊分区表升级版》(中科院期刊分区)数据库,《Science Citation Index Expanded (SCI-EXPANDED)》收录论文及其期刊影响因子、分区情况如下。(检索时间:2025年3月12日)

第 1 条, 共 1 条

标题:CVAE-LAYOUT: automatic furniture layout with constraints

作者:Xuan, YX(Xuan, Yixin);Song, C(Song, Chao);Jin, JQ(Jin, Jianqiu);Yang, BL(Yang, Bailin);

来源出版物:VISUAL COMPUTER 卷:40 期:11 页:7731-7745 提前访问日期:DEC 2023

DOI:10.1007/s00371-023-03204-2 出版年:NOV 2024

入藏号:WOS:001129743300002

文献类型:Article

地址:

[Xuan, Yixin] Zhejiang Univ, Polytech Inst, Hangzhou 310015, Zhejiang, Peoples R China.

[Song, Chao; Jin, Jianqiu; Yang, Bailin] Zhejiang Gongshang Univ, Sch Comp Sci & Technol, Hangzhou 310018, Zhejiang, Peoples R China.

通讯作者地址:

Song, C (corresponding author), Zhejiang Gongshang Univ, Sch Comp Sci & Technol, Hangzhou 310018, Zhejiang, Peoples R China.

电子邮件地址:xyx18061932@gmail.com; csong@zjsu.edu.cn; jqjin@zjsu.edu.cn; ybl@zjgsu.edu.cn

IDS 号:L2M10

ISSN:0178-2789

eISSN:1432-2315

期刊《VISUAL COMPUTER》2023 年的影响因子为 3.0, 五年影响因子为 3.0。

期刊《VISUAL COMPUTER》2023 年的 JCR 分区情况为:

Edition	JCR® 类别	类别中的排序	JCR 分区
SCIE	COMPUTER SCIENCE, SOFTWARE ENGINEERING	37/132	Q2

期刊《VISUAL COMPUTER》2023 年的中科院期刊分区情况为:

刊名	VISUAL COMPUTER		
年份	2023		
ISSN	0178-2789		
	学科	分区	Top 期刊
大类	计算机科学	3	否
小类	COMPUTER SCIENCE, SOFTWARE ENGINEERING 计算机: 软件工程	3	-

注:

1. 期刊影响因子最新数据以 JCR 数据库最新数据为准。
2. 以上检索结果来自 CALIS 查收查引系统。
3. 以上检索结果均得到委托人及被检索作者的确认。

