

# 3月论文阅读计划

## EVA3D 涉及的相关论文

下面的列表对 EVA3D 论文中介绍的目前做 3D 人体重构的 NeRF 和 GAN 相关的主要研究, EVA3D 方法中使用的主要模块和工具, 论文对比使用的 Baseline, 数据集和评估指标等方面相关的论文做列举;某些人体 NeRF 或者 3DGAN 同时也是其他研究的 baseline 或则是在某些条件下的 SOTA

计划使用一个月的时间将列表中的论文阅读完.

对于略读的论文, 在读完之后, 主要思考回答下面几个问题(源自于 readpaper的十问,根据习惯做了一些简化):

0. 作者团队来自哪里?有哪些经典研究?
1. 研究主要解决什么样的问题?
2. 相关问题出色的研究有哪些?
3. 问题产生的原因?or解决问题的动机?
4. 方法的主要思想和pipeline?
5. 定性定量实验有哪些?Baseline 有哪些?结果如何?如何验证设计有效性(消融实验)?
6. 研究没有解决或者引入的问题?or下一步要解决的问题

论文阅读清单随论文阅读进行不断更新,已经阅读完的论文会通过在前面的☐打勾,如下所示:

☒ [EVA3D:Compositional 3D Human Generation from 2D Image Collections](#)

### 论文阅读进展

每隔三天通过邮件汇报一次论文阅读进展?

在搜集相关的论文过程中发现就时尚人物生成和 3D 人物生成, NTU 的 S-Lab, MPI, Standford 等机构有稳定的研究产出并且往往是该领域的 SOTA的主要提出者

## 综述

☐ [Deep Generative Models on 3D Representations: A Survey](#)

## 精读

- ☐ [SMPL: a skinned multi-person linear model](#)描述人体姿势和形状变化相关的参数模型
- ☐ [NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis](#)使用神经网络从一组稀疏 2D 图像产生图像的新视角场景
- ☐ [StyleSDF:High-Resolution 3D-Consistent Image and Geometry Generation](#)仅在单视图 RGB 数据上进行训练, 使用 styleGAN2 生成图像,基于 SDF 的 3D 建模定义了详细的 3D 表面,从而实现一致的体积渲染(Baseline, 也是子网络摸扩的 backbone, 使用 SDF 作为隐式几何表示)
- ☐ [Garment4D: Garment Reconstruction from Point Cloud Sequences, Fangzhou Hong, NTU](#)使用穿着人体的 3D 点云序列进行服装重建, 引入了一种新的建议引导的层次特征网络和迭代图卷积网络, 它们集成了高级语义特征和低级几何特征以进行精细细节重建,用于平滑服装动作捕捉的时间变换器

- ☐ [AvatarCLIP: Zero-Shot Text-Driven Generation and Animation of 3D Avatars, FanZhou Hong, NTU](#)用于 3D 头像生成和动画的零样本文本驱动框架,在自然语言描述的驱动下,使用形状 VAE 网络初始化 3D 人体几何生成,利用体积渲染模型进一步做表面微调 and 纹理生成
- ☐ [2d gans know 3d shape? unsupervised 3d shape reconstruction from 2d image gans, ZiWei Liu, NTU](#)接从仅在 RGB 图像上训练的现成 2D GAN 中挖掘 3D 几何线索,预训练的 GAN 确实包含丰富的 3D 知识,因此可以用于以无监督的方式从单个 2D 图像恢复 3D 形状,探索和利用 GAN 图像流形中的不同视点和光照变化

## 略读

### NeRF + GAN

- ☐ [EG3D\(Baseline\):Efficient Geometry-aware 3D Generative Adversarial Networks, Eric R. Chan, Stanford & NVIDIA](#)分离特征生成和神经渲染,显式-隐式混合表示,姿势感知卷积生成器,双向判别器,StyleGAN
- ☐ [ENARF-GAN\(Baseline\):Unsupervised Learning of Efficient Geometry-Aware Neural Articulated Representations, Atsuhiko Noguchi, Tokyo & MSRA](#)提出基于三平面的关节物体的有效神经表示,使用 GAN 的框架进行无监督训练
- ☐ [pi-GAN:Periodic Implicit Generative Adversarial Networks for 3D-Aware Image Synthesis, Eric R. Chan, Stanford](#)利用具有周期性激活函数和体积渲染的神经表示来将场景表示为视图一致的辐射场
- ☐ [GRAF:Generative Radiance Fields for 3D-Aware Image Synthesis, Katja Schwarz, MPI & Tübingen](#)用于单个场景的新视图合成的辐射场生成模型,基于 patch 的多尺度的判别器
- ☐ [GIRAFFE:Representing Scenes as Compositional Generative Neural Feature Fields, Michael Niemeyer, MPI&Tübingen](#)将合成 3D 场景表示合并到生成模型中会导致更可控的图像合成,将场景表示为组合生成神经特征场,使我们能够从背景中分离出一个或多个对象以及单个对象的形状和外观,同时从非结构化和非摆姿势的图像集合中学习,而无需任何额外的监督,将此场景表示与神经渲染管道相结合,可以生成快速逼真的图像合成模型.
- ☐ [GRAM:Generative Radiance Manifolds for 3D-Aware Image Generation, Yu Deng, Tsinghua&MSRA&USTC](#)提出了一种调节 2D 流形上的点采样和辐射场学习的新方法,体现为 3D 体积中的一组学习隐式曲面,对于每条观察光线,我们计算光线与表面的交点并累积它们由网络生成的辐射率
- ☐ [GNARF:Generative Neural Articulated Radiance Fields, Eric R. Chan, Stanford](#)开发一个 3D GAN 框架来解决这些挑战,该框架学习以规范姿势生成人体或面部的辐射场,并使用显式变形场将它们扭曲成所需的身体姿势或面部表情
- ☐ [HumanGen: Generating Human Radiance Fields with Explicit Priors](#)

### GAN + 其它

- ☐ [HoloGAN: Unsupervised Learning of 3D Representations From Natural Images](#)HoloGAN 通过学习的 3D 特征的刚体变换提供对生成对象姿态的明确控制,从未标记的二维图像进行端到端训练
- ☐ [Escaping Plato's Cave: 3D Shape From Adversarial Rendering](#)
- ☐ [StylePeople: A Generative Model of Fullbody Human Avatars](#)
- ☐ [3D-Aware Semantic-Guided Generative Model for Human Synthesis](#)

## 人体 NeRF

- ☐ [HumanNeRF: Efficiently Generated Human Radiance Field from Sparse Inputs, Fuqiang Zhao, SHTU](#)具有高效泛化能力的神经表征——用于动态人类的高保真自由视角合成, 在多视图输入中采用聚合像素对齐功能, 以及用于处理动态运动的姿势嵌入式非刚性变形场
- ☐ [Animatable Neural Radiance Fields for Modeling Dynamic Human Bodies, Sida Peng, ZJU](#)引入神经混合权重场来产生变形场, 基于骨架驱动的变形, 混合权重场与 3D 人体骨骼一起使用, 以生成观察到规范和规范到观察的对应关系
- ☐ [NARF: Neural Articulated Radiance Field, Atsuhiko Noguchi1, Tokyo&MSRA](#)在制定 3D 关节物体的隐式表示时, 我们的方法仅考虑最相关物体部分的刚性变换, 以求解每个 3D 位置的辐射场, 通过这种方式, 所提出的方法可以在不显著增加计算复杂性的情况下表示与姿态相关的变化, NARF 是完全可微的, 可以从带有姿势注释的图像中进行训练, 通过使用自动编码器, 它可以学习一个对象类的多个实例的外观变化。
- ☐ [H-NeRF: Neural Radiance Fields for Rendering and Temporal Reconstruction of Humans in Motion, Google](#)用于运动中人类 (H-NeRF) 的渲染和时间 (4D) 重建的神经辐射场, 结合了神经场景表示、新视图合成和隐式统计几何人体表示的思想, 并使用新的损失函数进行耦合
- ☐ [A-NeRF: Articulated Neural Radiance Fields for Learning Human Shape, Appearance, and Pose](#)隐式模型需要显式表面模型中使用的正向运动学的逆。我们的重新参数化定义了相对于身体部位姿势的空间潜在变量, 从而克服了具有过度参数化的不适定逆运算
- ☐ [StyleNERF: A STYLE-BASED 3D-AWARE GENERATOR FOR HIGH-RESOLUTION IMAGE SYNTHESIS, MPI](#)
- ☐ [SelfRecon: Self Reconstruction Your Digital Avatar from Monocular Video, USTC](#)结合隐式和显式表示穿衣人体重建的方法, 利用显式网格的微分掩模损失来获得连贯的整体形状, 而隐式表面上的细节则通过可微神经渲染进行细化
- ☐ [ARAH: Animatable Volume Rendering of Articulated Human SDFs, MPI](#)将铰接式隐式表面表示与体积渲染相结合, 同时进行射线-表面相交搜索和对应搜索
- ☐ [DANBO: Disentangled Articulated Neural Body Representations via Graph Neural Networks](#)
- ☐ [Mixture of Volumetric Primitives for Efficient Neural Rendering](#)
- ☐

## 人体表示

- ☐ [NPMs: Neural Parametric Models for 3D Deformable Shapes, Pablo Palafox, TUM&MPI](#)提出了神经参数模型 (Neural Parametric Models, NPM), 这是一种新颖的、可学习的替代传统参数化 3D 模型的方法, 它不需要手工制作的、特定于对象的约束
- ☐ [Real-time Deep Dynamic Characters, MPI](#)提出了一种参数化和可微分的字符表示, 它允许我们对粗略和精细的动态变形进行建模, 例如, 衣服皱纹, 作为明确的时空相干网格几何形状, 并根据运动和视点增加了高质量的动态纹理
- ☐ [Neural body: Implicit Neural Representations with Structured Latent Codes for Novel View Synthesis of Dynamic Humans, Sida Peng, ZJU](#)假设在不同帧中学习的神经表示共享同一组锚定到可变形网格的潜在代码, 以便可以自然地整合跨帧的观察结果, 可变形网格还为网络提供几何指导, 以更有效地学习 3D 表示
- ☐ [Style and Pose Control for Image Synthesis of Humans from a Single Monocular View, MPI](#)
- ☐ [HumanGAN: A Generative Model of Human Images, MPI](#)
- ☐ [gDNA: Towards Generative Detailed Neural Avatars](#)

- ☐ [SPAMs: Structured Implicit Parametric Models](#)
- ☐ [Coap: Compositional articulated occupancy of people](#)

## 其它

- ☐ [Which training methods for gans do actually converge?](#)
- ☐ [AMASS: Archive of Motion Capture As Surface Shapes](#)
- ☐ [Expressive body capture: 3d hands, face, and body from a single image.](#)
- ☐ [GEOMETRY-CONSISTENT NEURAL SHAPE REPRESENTATION WITH IMPLICIT DISPLACEMENT FIELDS](#)
- ☐ [Vibe: Video inference for human body pose and shape estimation](#)
- ☐ [Implicit Neural Representations with Periodic Activation Functions](#)
- ☐ [PaMIR: Parametric Model-Conditioned Implicit Representation for Image-based Human Reconstruction](#)
- ☐ [MotionDiffuse: Text-Driven Human Motion Generation with Diffusion Model](#)(作者主页上一篇关于扩散模型的多模态动作生成报告)

## 数据集

- ☐ [StyleGAN-Human: A Data-Centric Odyssey of Human Generation](#)
- ☐ [InsetGAN for Full-Body Image Generation](#)
- ☐ [DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations](#)
- ☐ [HuMMan: Multi-Modal 4D Human Dataset for Versatile Sensing and Modeling](#)
- ☐ [AIST: AIST DANCE VIDEO DATABASE: MULTI-GENRE, MULTI-DANCER, AND MULTI-CAMERA DATABASE FOR DANCE INFORMATION PROCESSING](#)
- ☐ [UBCFashion: DwNet: Dense warp-based network for pose-guided human video generation](#)

## 评估标准

- ☐ [FID: Gans trained by a two time-scale update rule converge to a local nash equilibrium](#)
- ☐ [KID: Demystifying mmd gans](#)
- ☐ [PCKh@0.5: 2D Human Pose Estimation: New Benchmark and State of the Art Analysis](#)
- ☐ [Depth: Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer](#)

## 下游任务

- ☐ ...

## PIFU/ICON相关的

---

另外一条可能的研究路线, 沿着修宇亮在知乎上这篇文章:[ICON: 提高三维数字人重建的姿势水平](#) 描述的研究路线阅读论文, 了解研究进展

- ☐ [PIFU](#)
- ☐ [PIFUHD](#)

☐ [ICON](#)

☐ [ECON](#)

☐ .....

☐