

进展汇报

黄道吉

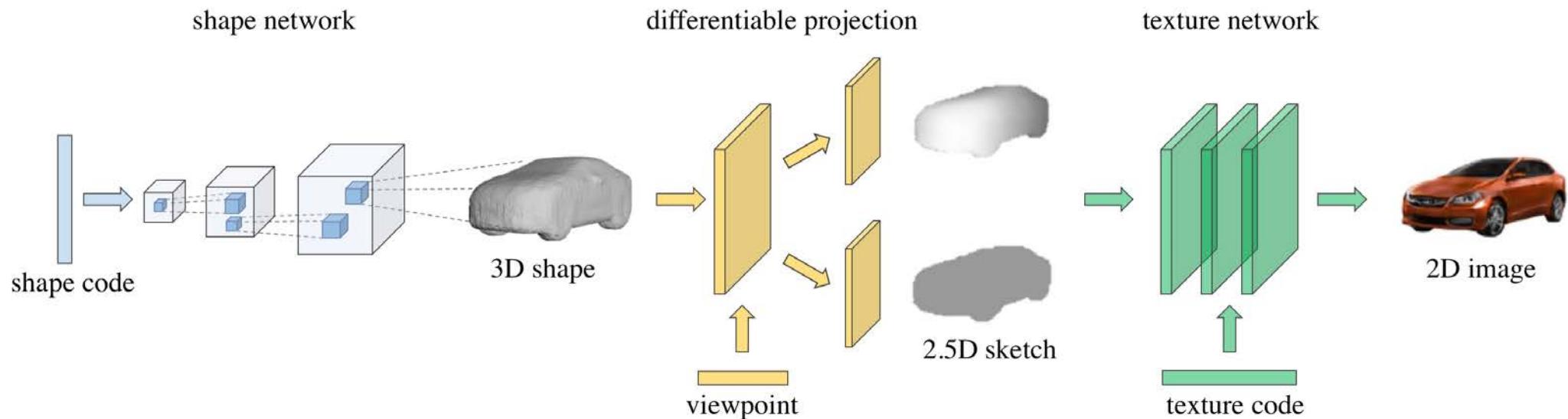
问题描述

- 任务：三维模型最优视图生成
 - 输入：三维模型(的深度图/轮廓图…)
 - 输出：渲染好的视图
 - 三维模型的材质
 - 可以看作图片到图片转换的问题
 - VON, VIGAN…

相关工作

- VON

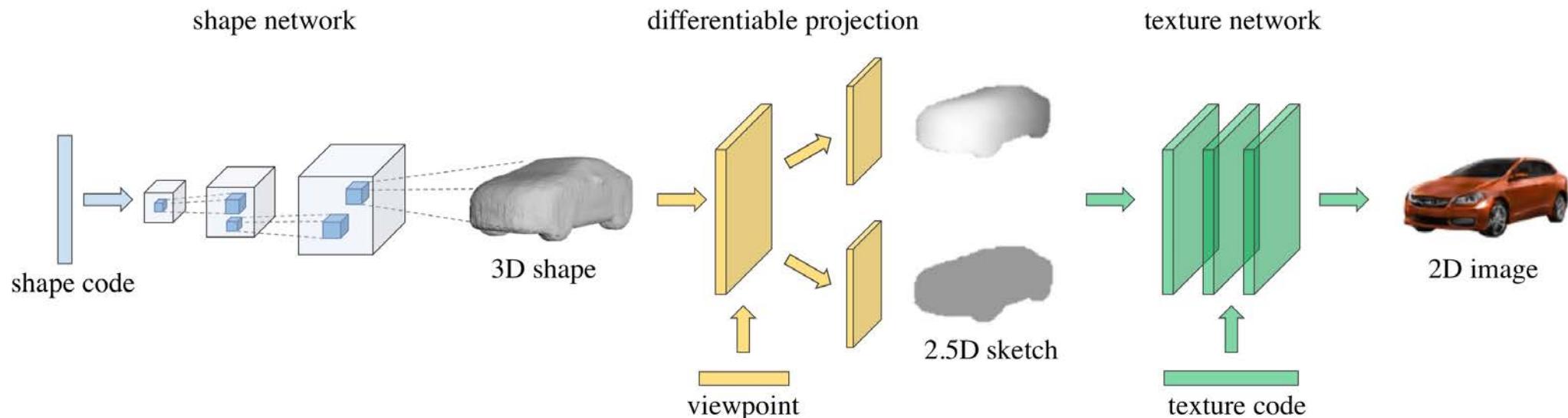
- Texture network: conditional CycleGAN
- 训练数据：合成的深度图和2k张真实图片
- Texture code随机生成，不支持条件生成



相关工作

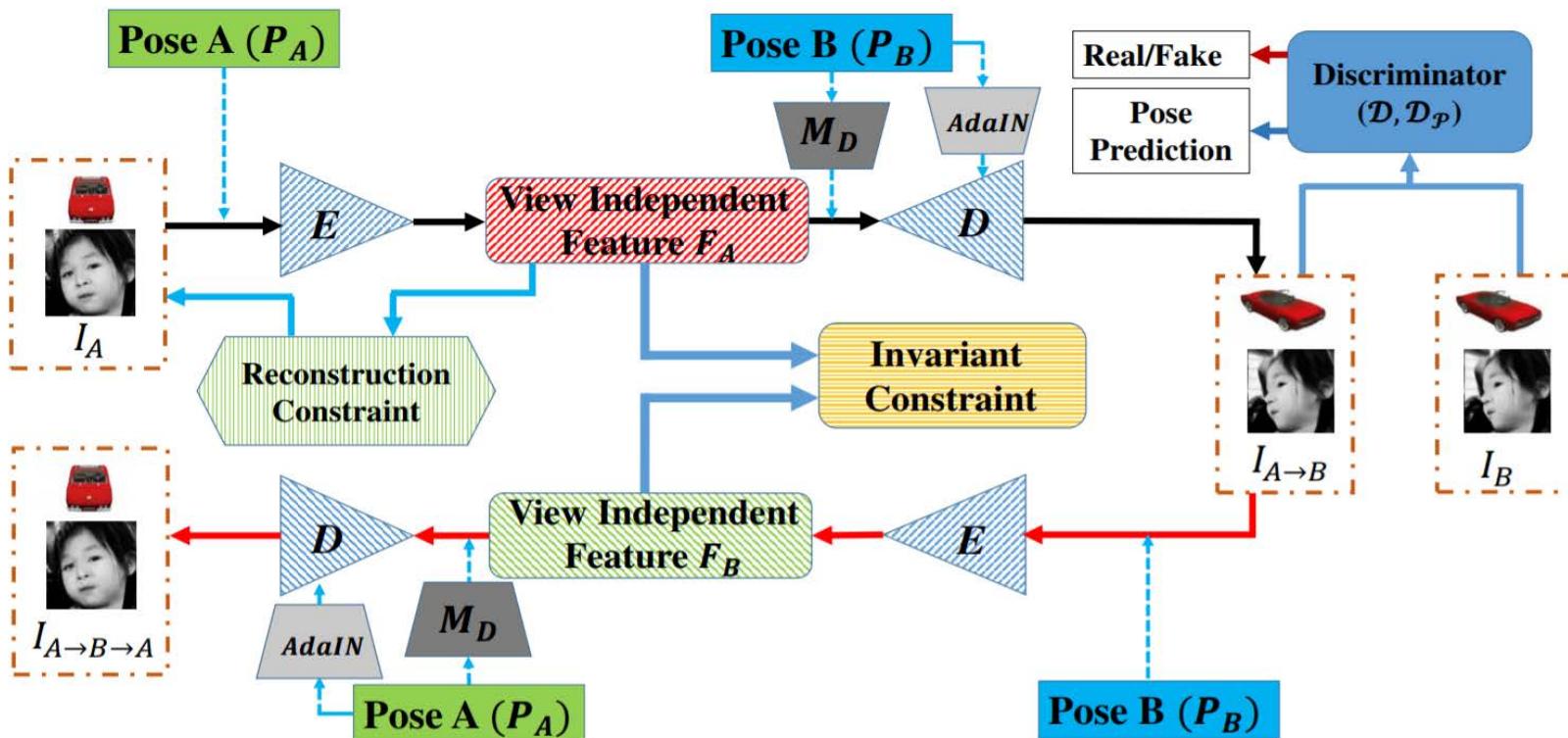
- VON

- Texture network: conditional CycleGAN
- 训练数据：合成的深度图和2k张真实图片
- Texture code随机生成，不支持条件生成
- 复现结果
 - 训练过程不稳定
 - 不同视角下生成图片材质差异大 -> 没有约束条件



相关工作

- VIGAN
 - 将视角和视角无关的信息分离
 - 视角变换只考虑azimuth的变化
 - 参考了训练方式



相关工作

- VIGAN
 - 将视角和视角无关的信息分离
 - 视角变换只考虑azimuth的变化
 - 参考了训练方式
 - Shapenet上的实验结果
 - 材质相比car并不复杂
 - 光影

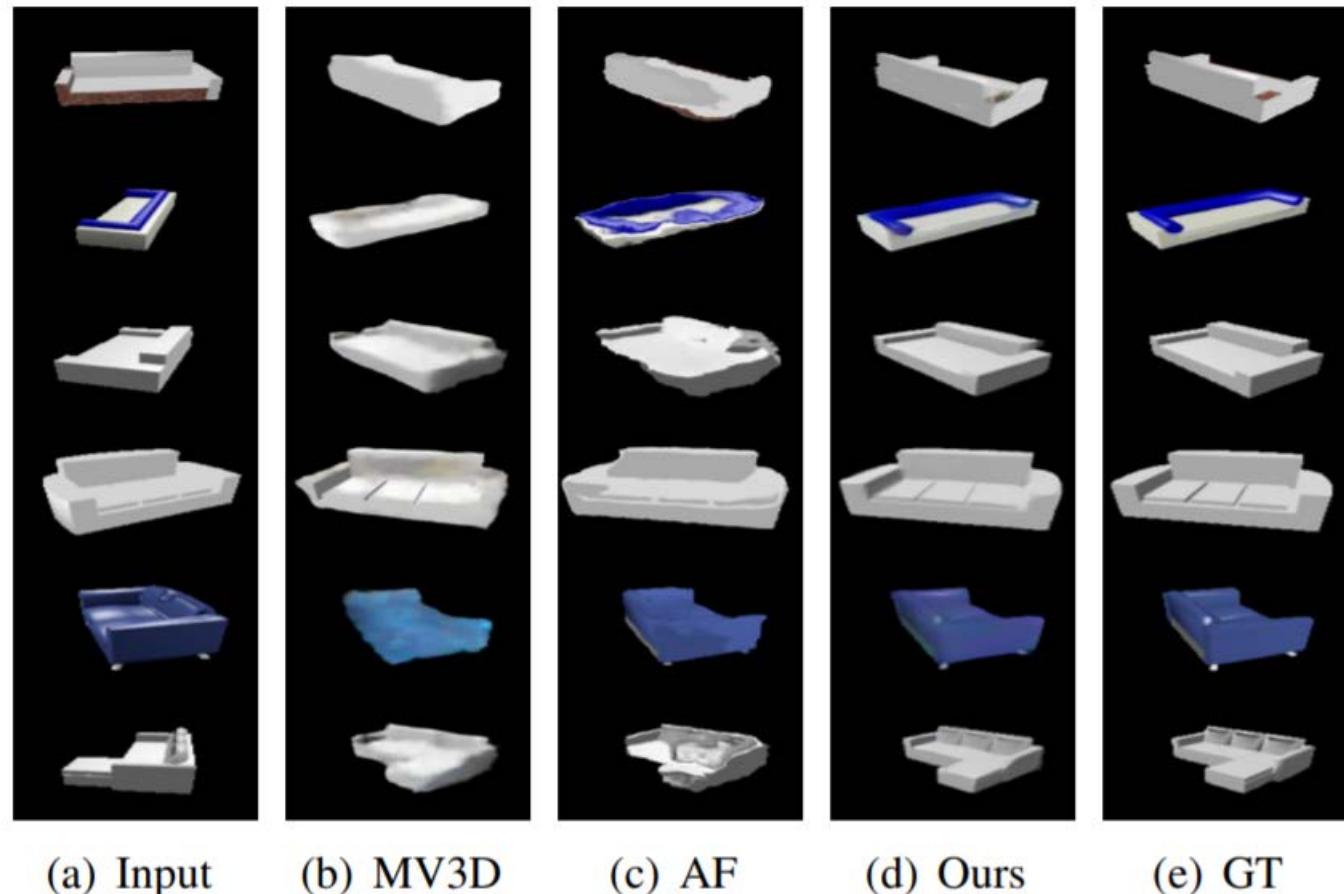
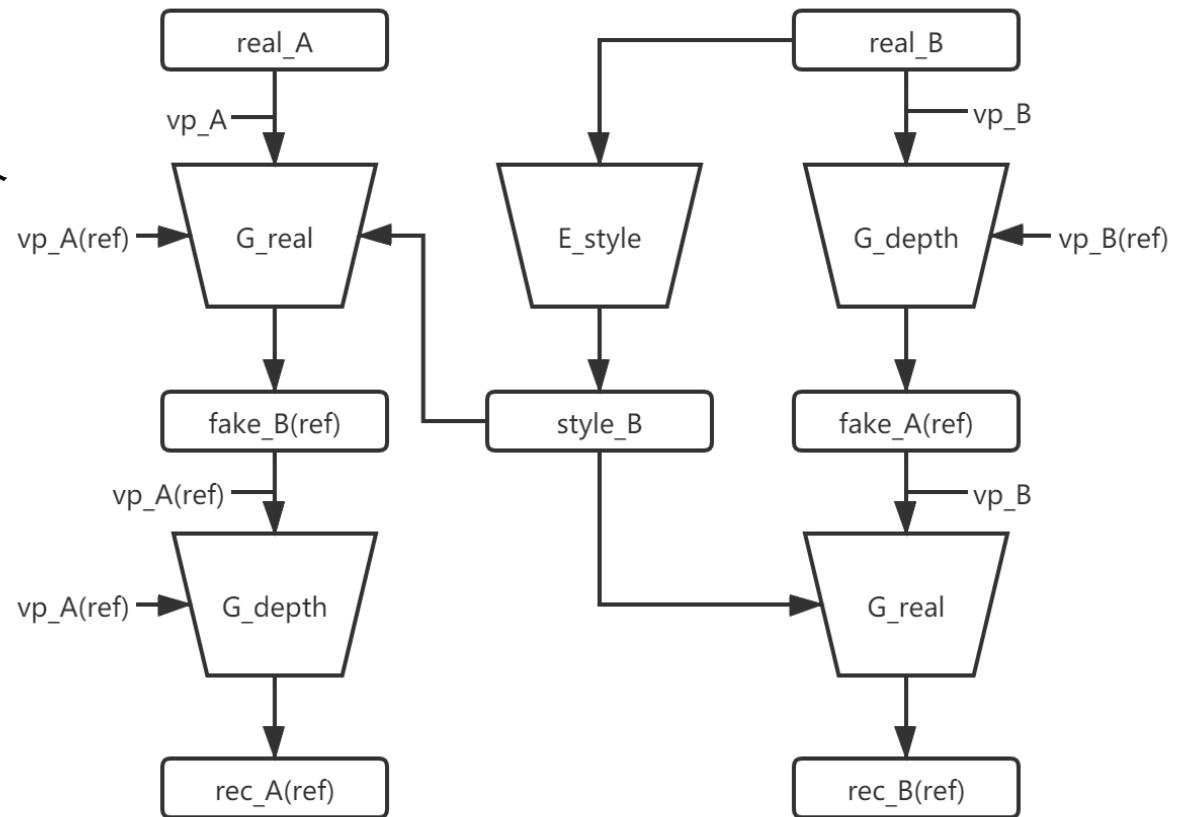


Figure 9. Results in “sofa” category. Order of each column is the same as that of Fig. 7.

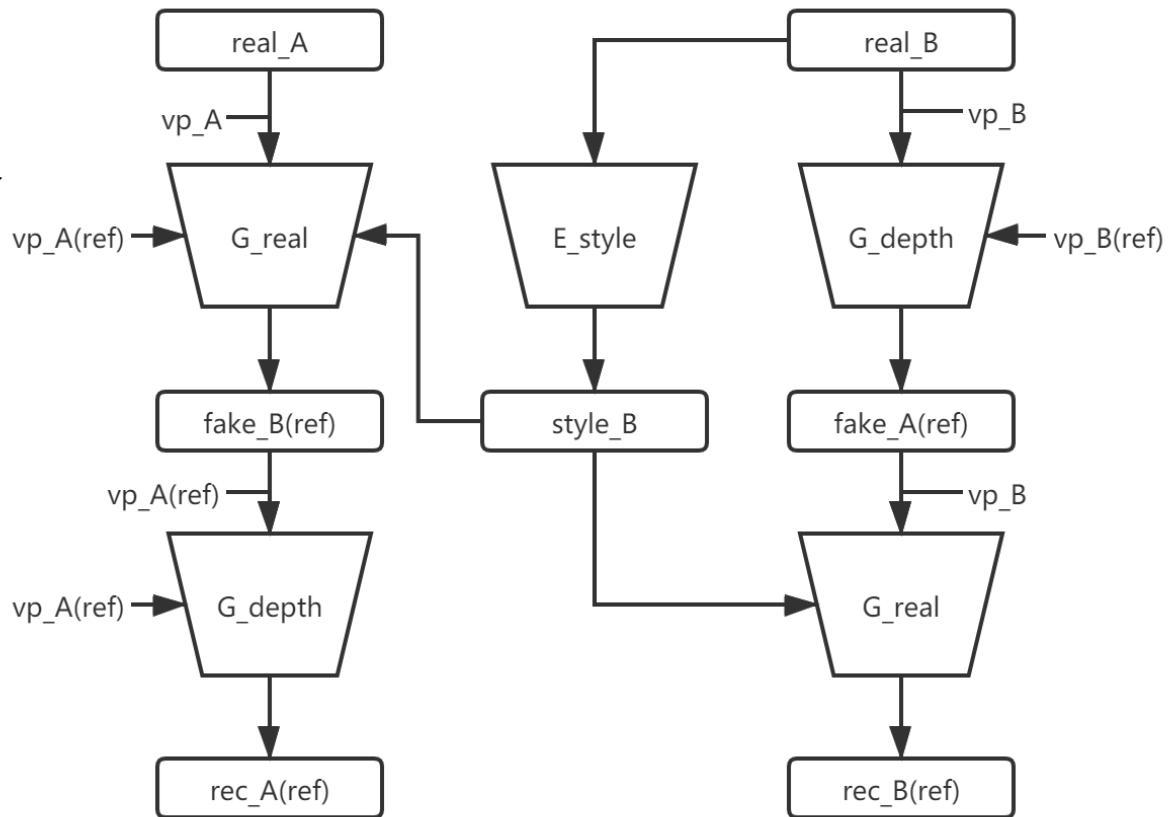
模型结构

- 以VON代码为基础修改模型
 - 用深度图表征内容信息
 - 隐变量表征材质信息
 - Rgb图片是上述两种信息的混合
 - Generator
 - 编码时每层concat视角
 - 解码时每层concat材质



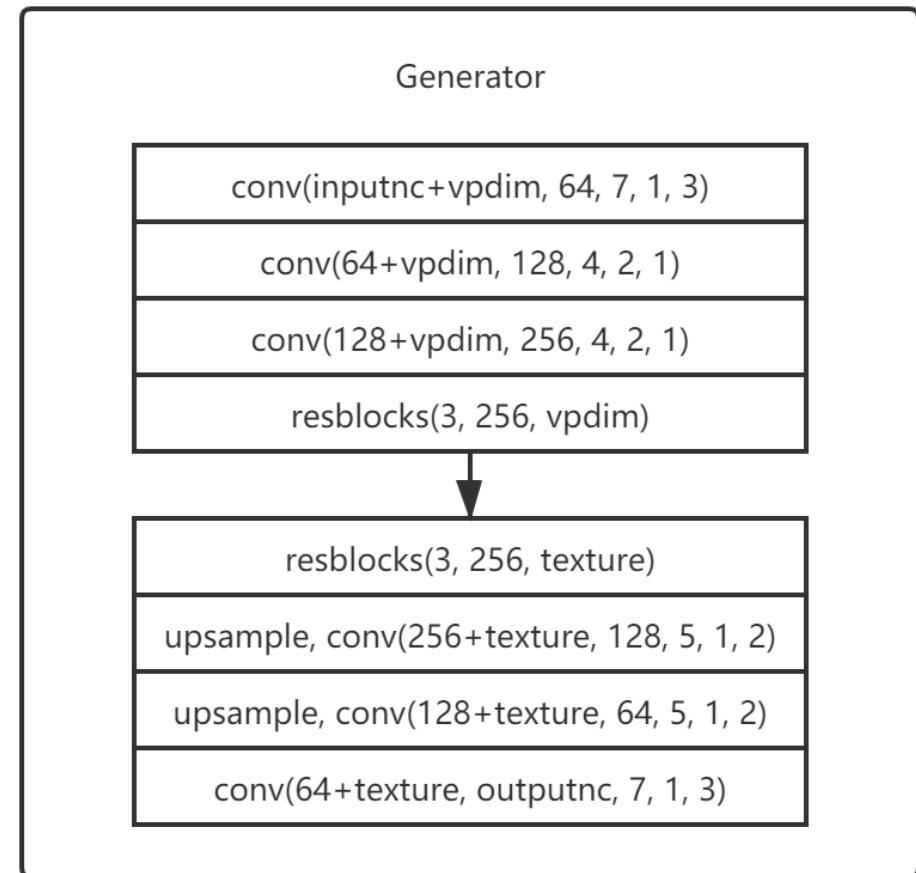
模型结构

- 以VON代码为基础修改模型
 - 用深度图表征内容信息
 - 隐变量表征材质信息
 - Rgb图片是上述两种信息的混合
 - Generator
 - 编码时每层concat视角
 - 解码时每层concat材质
 - Loss
 - 重建loss: L1
 - GAN loss: 多分辨率的判别器
 - VAE loss



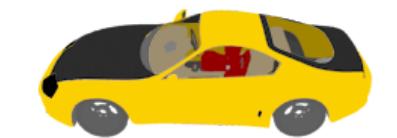
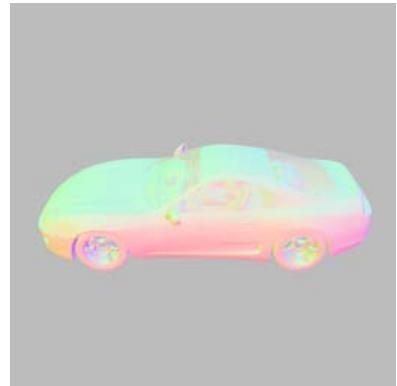
模型结构

- 以VON代码为基础修改模型
 - 用深度图表征内容信息
 - 隐变量表征材质信息
 - Rgb图片是上述两种信息的混合
 - Generator
 - 编码时每层concat视角
 - 解码时每层concat材质
 - 超参数
 - vpdim=2
 - texture=64
 - n_downsample=2



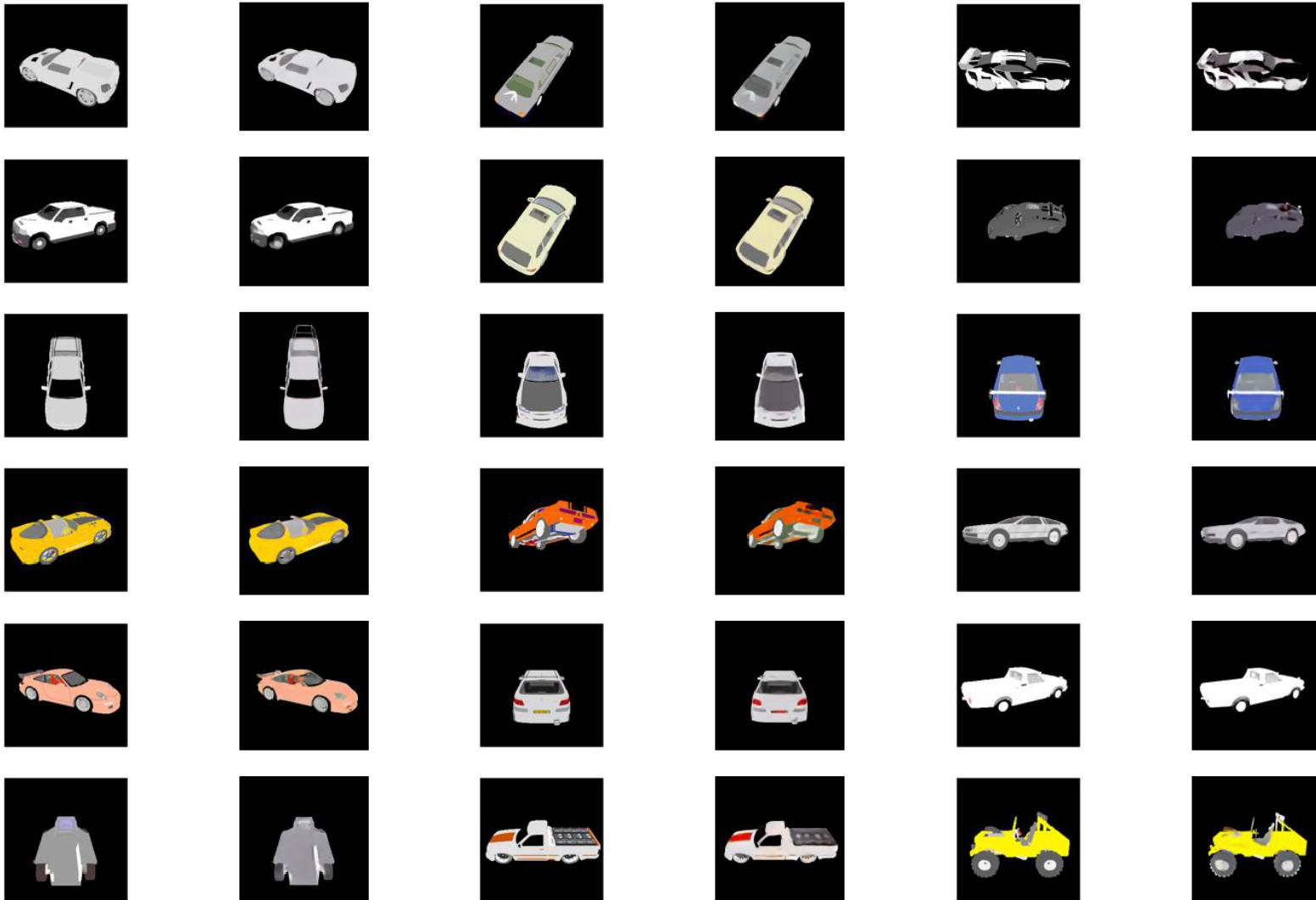
训练数据

- Blender渲染shapenet-car分支下的模型
 - 训练-测试集随机划分
 - Train: 1667 model, >140k images
 - Test: 152 model, > 10k images
 - 每一个模型渲染 $12 * 6$ 视角下的图片
 - az: 0 -> 330, el -30 -> 45
 - 用depth map和albedo训练
 - normal map在CycleGAN上面没有性能提升



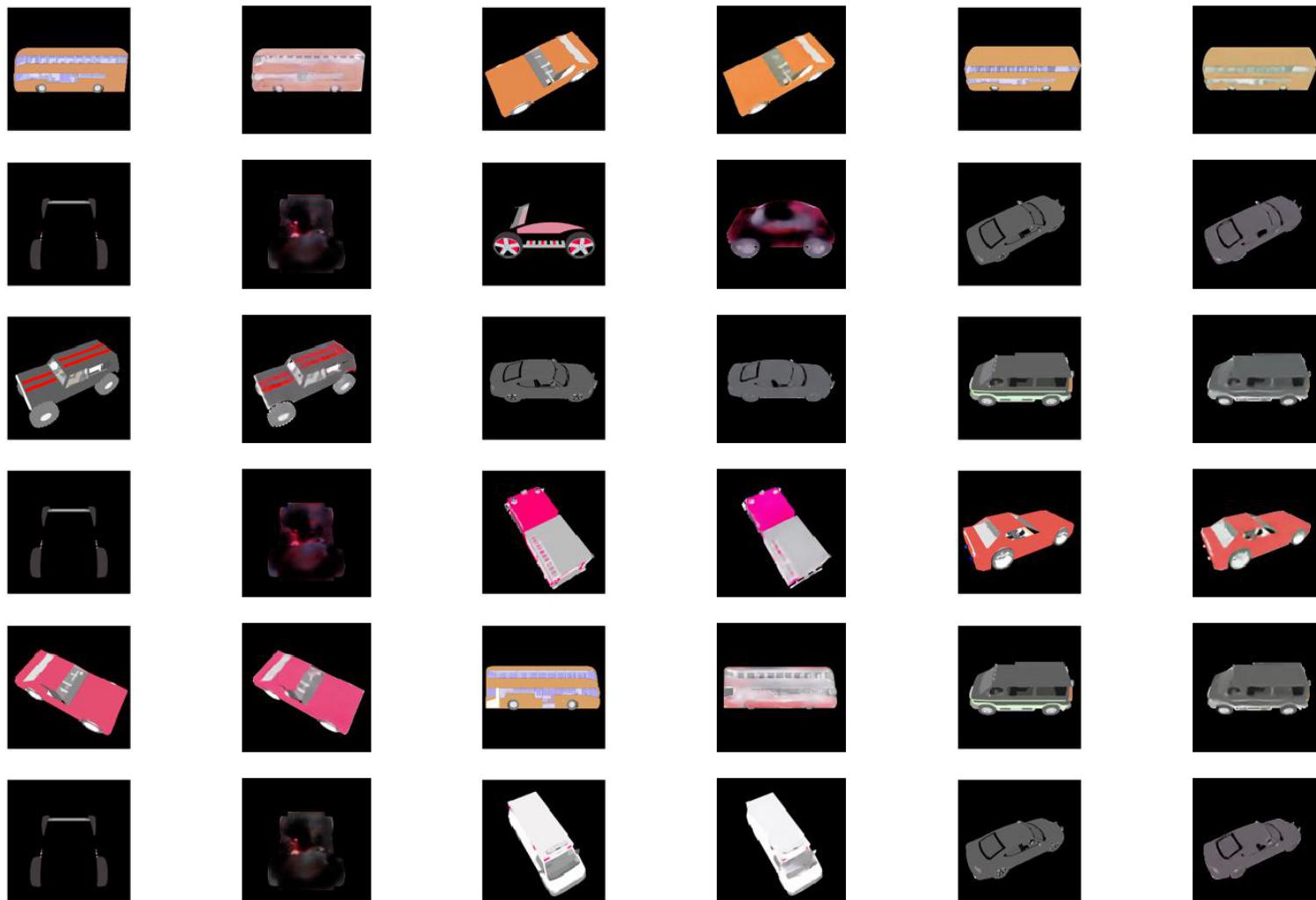
测试结果

- 重建任务
 - 测试集上的结果
 - FID: 40.64
 - 135列为GT
 - 细粒度的部分还原有问题
 - 车灯/车牌/窗框



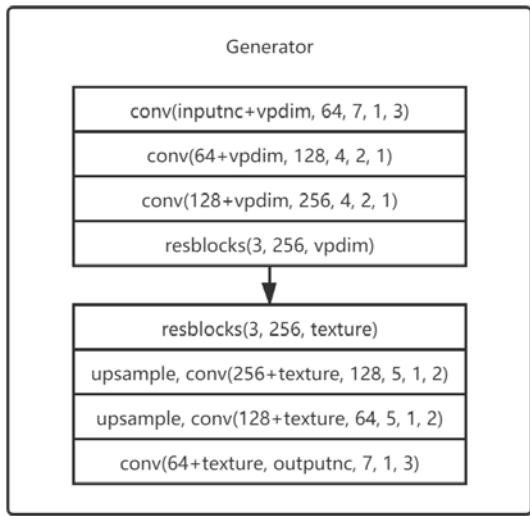
测试结果

- 重建任务
 - 误差最大的测试结果
 - 上到下, 左到右
 - 细节纹理还原不好
 - 车窗



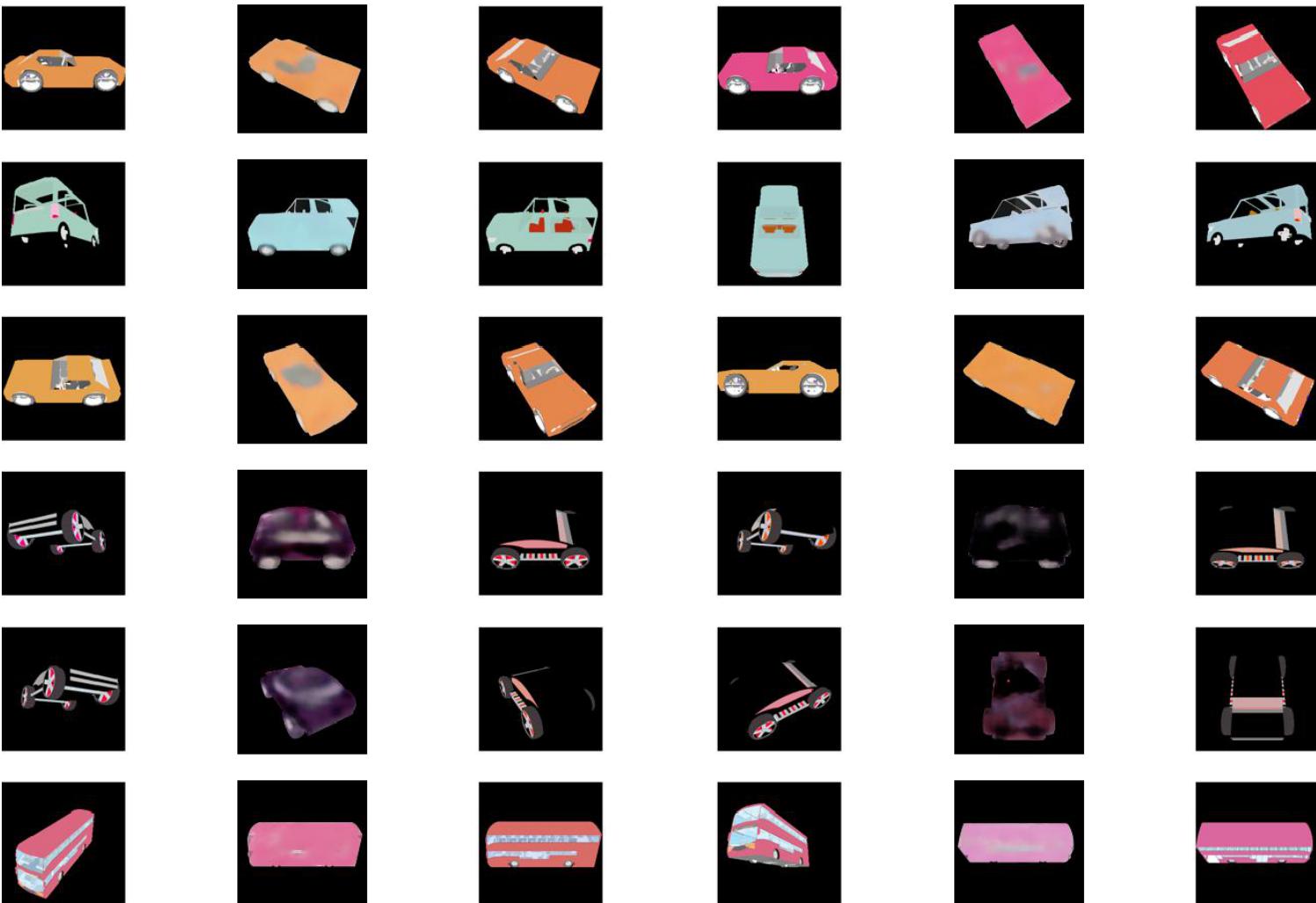
测试结果

- Novel view synthesis
 - 随机结果
 - Input | output | target
 - FID: 109.29
 - GT颜色误差
 - 加大下采样次数?
 - 感受野
 - 约束隐变量



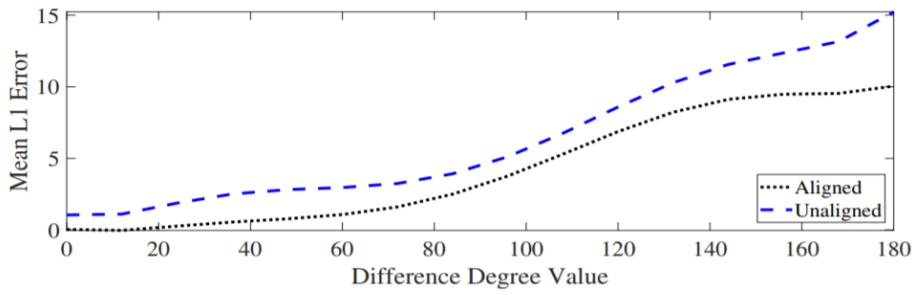
测试结果

- Novel view synthesis
 - 误差最大的结果
 - 上到下, 左到右
 - 原视角遮盖住的细节
 - 色差 hue



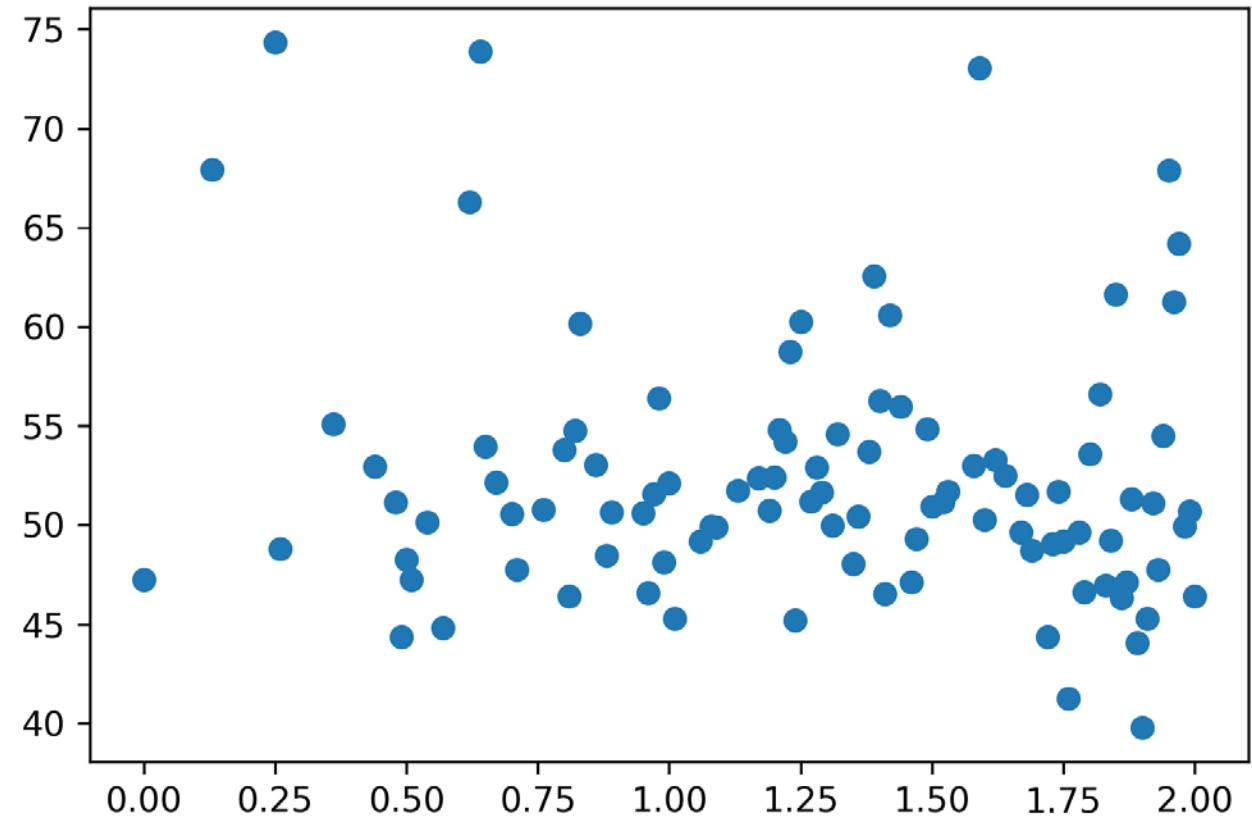
测试结果

- Novel view synthesis
 - L1 loss与视角间距的关系
 - 中间高两边低 -> 对称



(b) Pose Sensibility Analysis of Decoder

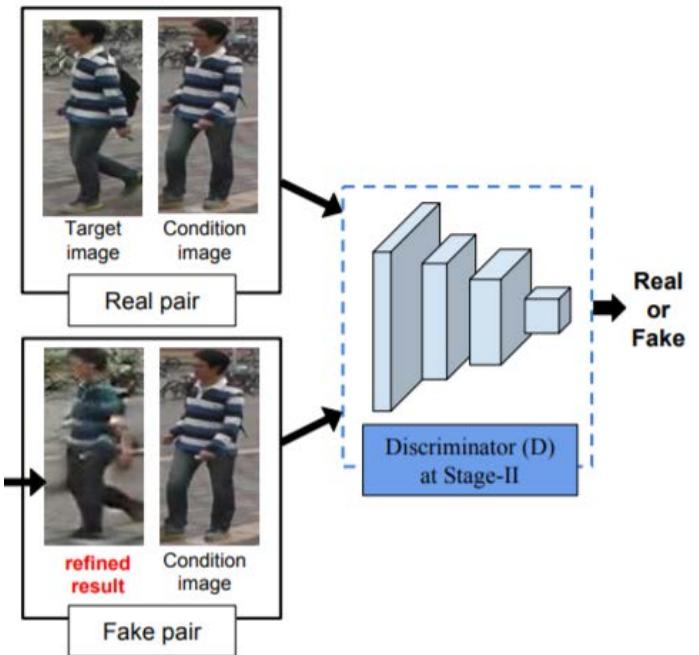
Figure 3. Pose sensibility analysis for the encoder and decoder.



测试结果

- Texture transfer

- 随机结果
 - FID: 59.28
- 只能大致将色调迁移过去
- 可能需要引入额外的判别器
 - 需要获取real的材质信息



测试结果

- 视角无关性

- Content encoder
 - Novel view synthesis
- Style encoder
 - Real/realref/fake/fakeref



测试结果

- 视角无关性

- Content encoder
 - Novel view synthesis
- Style encoder
 - Real/realref/fake/fakeref
- Decoder
 - Real/fake/fakeref



测试结果

- 无条件生成
 - FID: 78.15



改进方向

- 对content 隐变量加限制
 - 视角无关性
 - 加深下采样次数-> +fc?
- 改进判别器
 - 成对的判别
- 整合各个视角的信息
 - 要先做好novel view synthesis
 - confidence map?
- 高分辨率生成
 - 真实图片: 不成对 数据少