



# 三维模型最优视图生成技术研究

答辩人：元培学院 2016级 黄道吉

指导老师：连宙辉

本科生毕业论文



# 目录

- 1 / 引言
- 2 / 相关工作
- 3 / 方法介绍
- 4 / 实验结果

# 01

## *Part One* 引言





# 研究背景

三维模型被广泛应用，但

- 数据集可视化并不理想
- 检索模型仍很费力

The screenshot shows the ShapeNetCore website interface. On the left, there is a sidebar titled "Choose taxonomy:" with a dropdown menu set to "ShapeNetCore". The list includes categories like "display,video display(3,100)", "earphone,epiece,headphone,phone(0,73)", and "motorcycle,bike(2,337)". Below this is a search bar and a navigation menu with links for Home, Browse, Resources, Download, API, Challenges, About, and Q/A Forum. The main content area has tabs for "Synset Models", "TreeMap", "Stats", and "Measures". It displays a grid of 160 motorcycle models from Synset 1, with labels such as "motorcycle" repeated multiple times. A navigation bar at the bottom shows pages 1, 2, 3, and >.

→生成最优视图

- 现有方法只基于几何信息
- 没有考虑到材质的影响

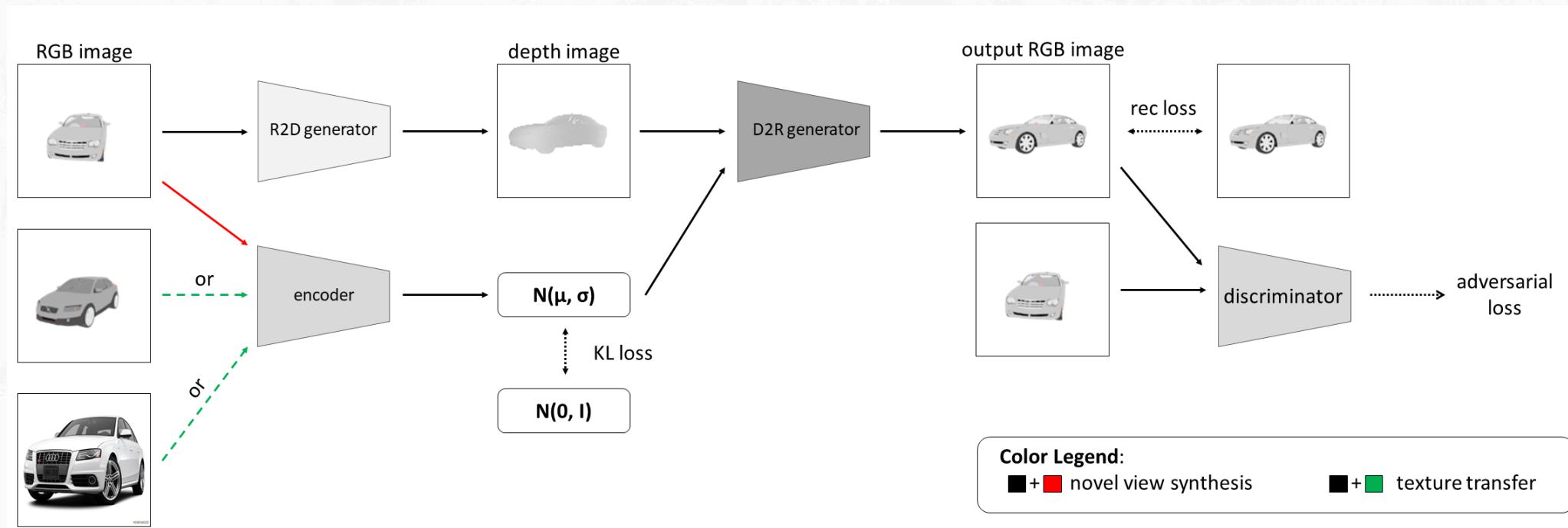




# 研究内容

## 生成三维模型最优二维视图

- 基于新视角生成模型，分离材质与内容隐变量
- 能条件/非条件生成真实性材质的二维视图
- 用重构其他视图误差衡量视图的信息，选取最优视角



# 02

*Part Two*

## 相关工作

---





## 相关工作

### 最优视角选择

#### 图形学方法求解最优视角

- 最优视角被定义为信息最多的视角
- 二维视图：视角熵
- 网格顶点：Mesh Saliency
- 没有考虑到材质信息的影响，对建模方式敏感



(a)



(b)



(c)



(d)

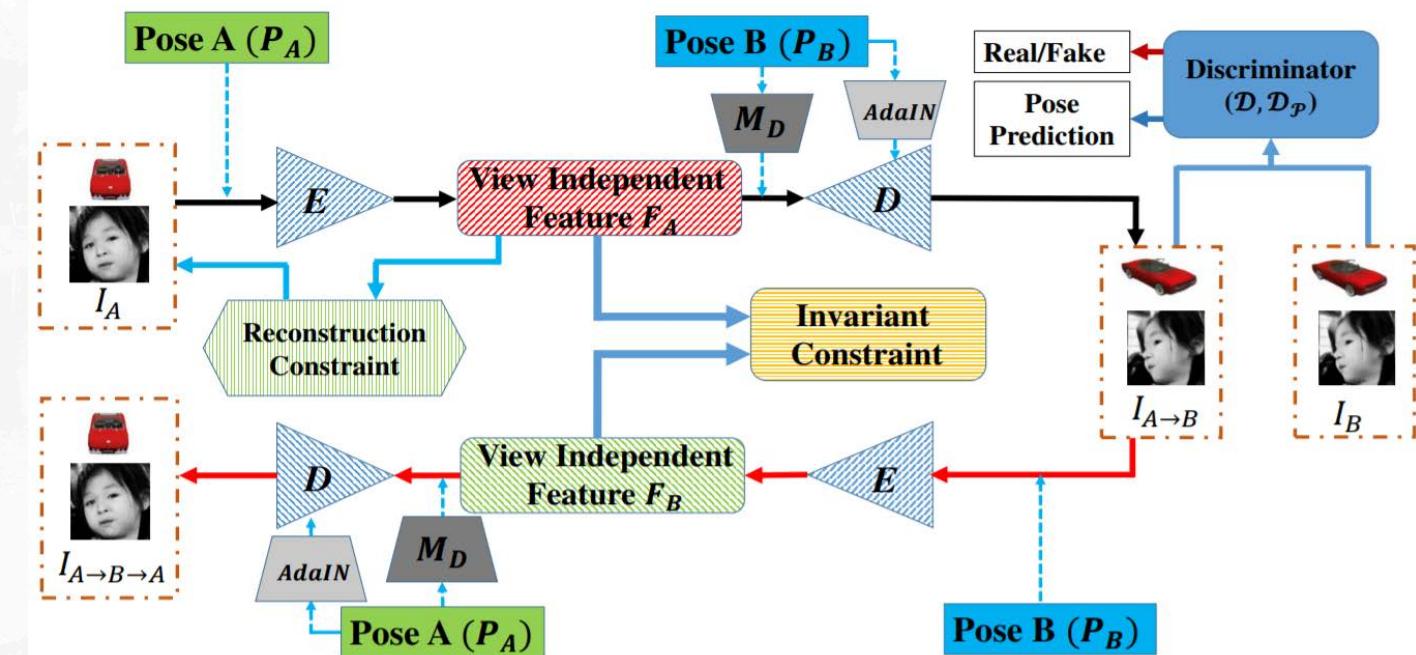


# 相关工作

## 新视角生成

给定一个/多个视角生成新视角下视图

- 非良定义问题，需要先验知识和正则化项
  - 感知损失函数
  - GAN损失函数
  - 视角无关三维信息
- 基于学习方法：看成图像生成任务
  - 预测流：appearance flow
  - 直接生成每一个像素：mv3d



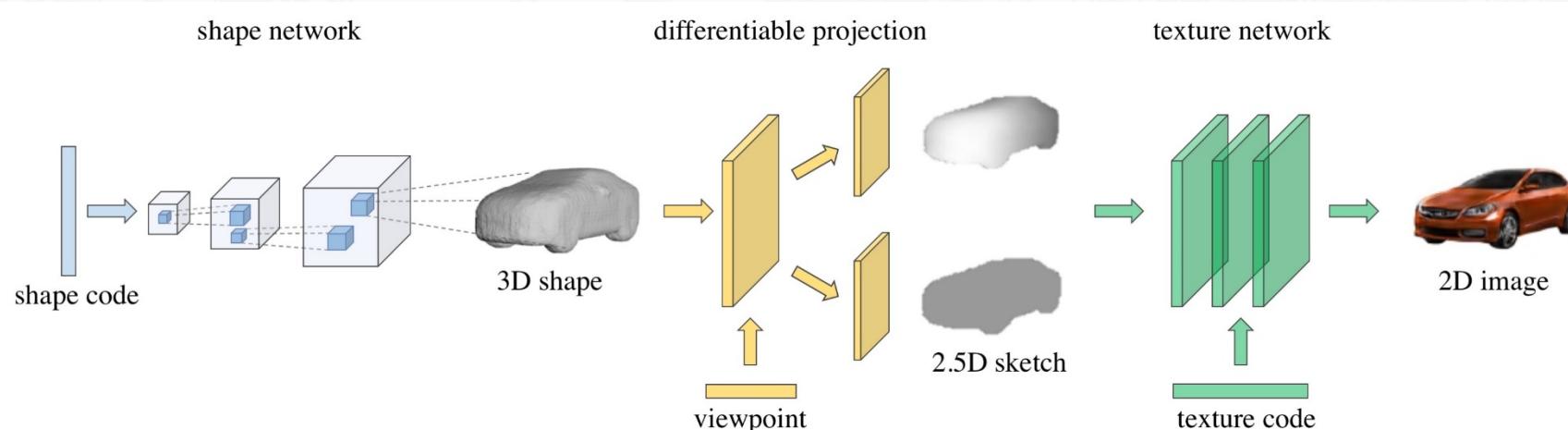


# 相关工作

## 材质迁移

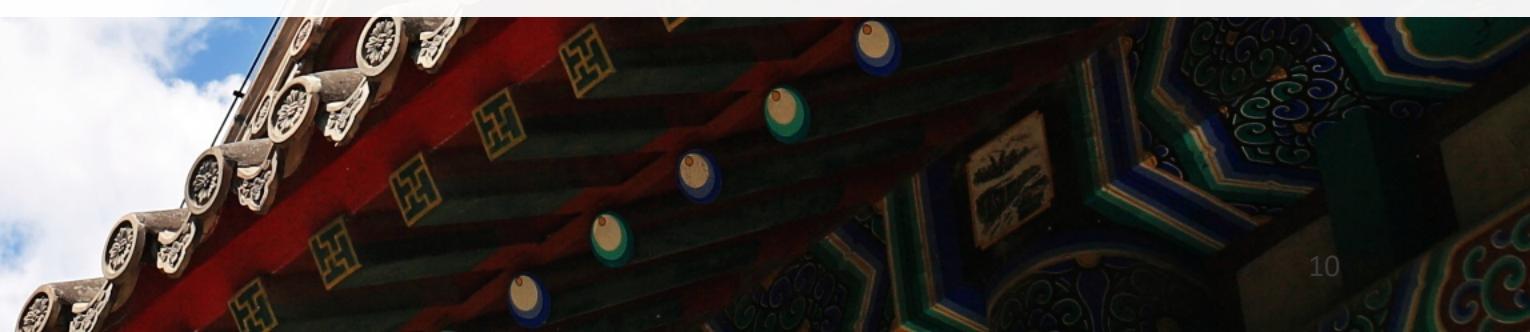
给定内容图片和材质信息，生成融合二者的图片

- 材质信息提取自图片/随机采样
- 没有成对监督数据：需要正则项分离材质和内容
  - 预训练VGG：低层内容 高层材质
  - 循环损失函数
  - 设计独特的判别器
- 应用在三维模型
  - 直接生成面片颜色：预定义拓扑结构，可微分渲染器
  - 深度图上色：可利用图像领域方法，保留三维信息



# 03 *Part Three*

## 方法介绍





# 方法介绍

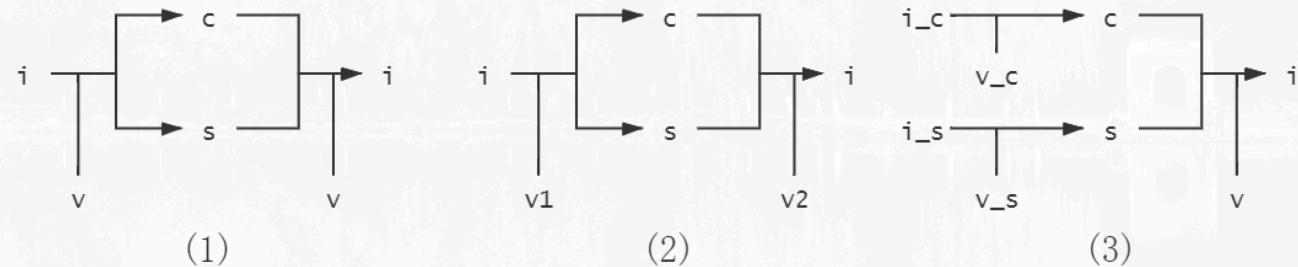
## 实验原理

### 优化目标

- $i_{RGB} \sim P(i|c, s, v)$
- $i_d \sim P(i|c, v)$
- 仿照VAE，变分下界

$$\begin{aligned}\log p(i) &= \log p(c) + \log p(i|c) - \log p(c|i) \\ &\geq \log p(c) + \log \int p(i, s|c) ds \\ &\geq \log p(c) + E_q \log \frac{p(i, s|c)}{q(s|i, c)} \\ &= \log p(c) + E_q \log \frac{p(i|c, s)p(s)}{q(s|i)} \\ &= \log p(c) + E_q \log p(i|c, s) + KL(q(s)||p(s)) \\ \log p(i|v) &= \log p(c) + E_q \log p(i|c, s, v) + KL(q(s)||p(s))\end{aligned}$$

- 隐变量之间互相独立
  - 仅重构图像并不足够
  - 新视角生成
  - 材质迁移



- 最优视角选择
  - 信息最多的视角：在重构视图的意义下

$$vp_{best}^m = \operatorname{argmin}_{vp} E_{vp'} |i_{vp'} - f(i_{vp}, vp, vp')|$$

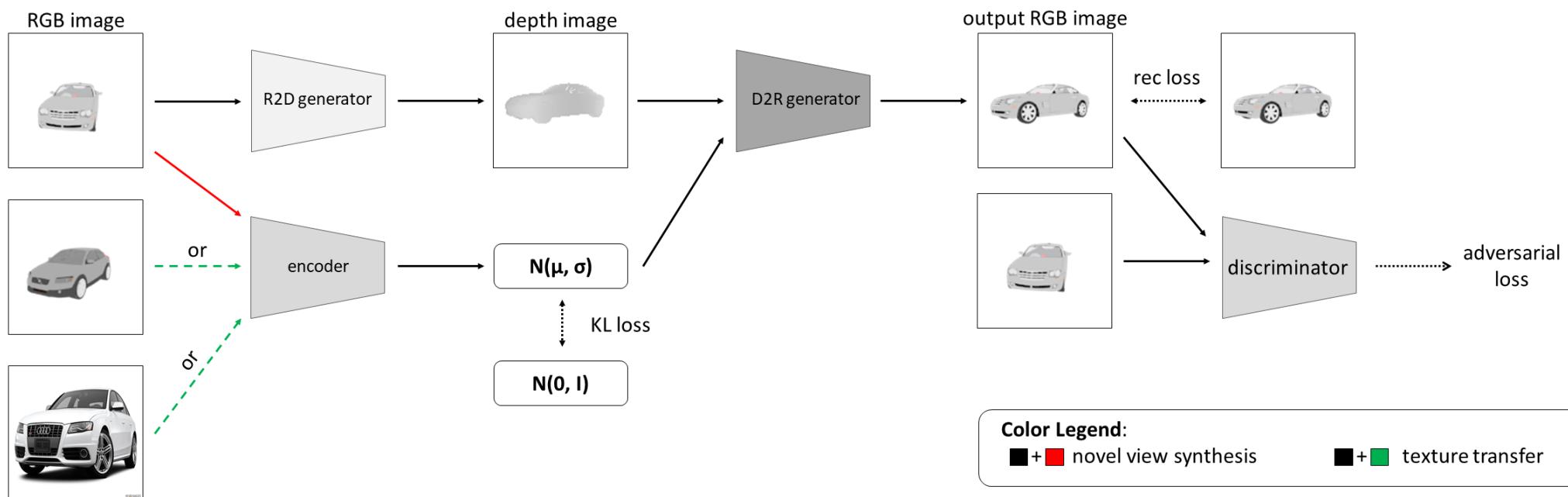


# 方法介绍

## 实现细节

需要实现的映射  $c = f(i_{RGB}, v), i_d = g(c, v), s = h(i_{RGB}, v), i_{RGB} = m(c, s, v)$

- 编码器
  - 提取视角无关的材质信息
  - 视角拼接在图片后
  - 全局池化，全连接  $\rightarrow$  全局特征
- 判别器
  - 弥补缺失的监督数据
  - 拟合符合训练目标的分布
    - 高质量/真实性图片
    - 具有特定材质图片：成对输入





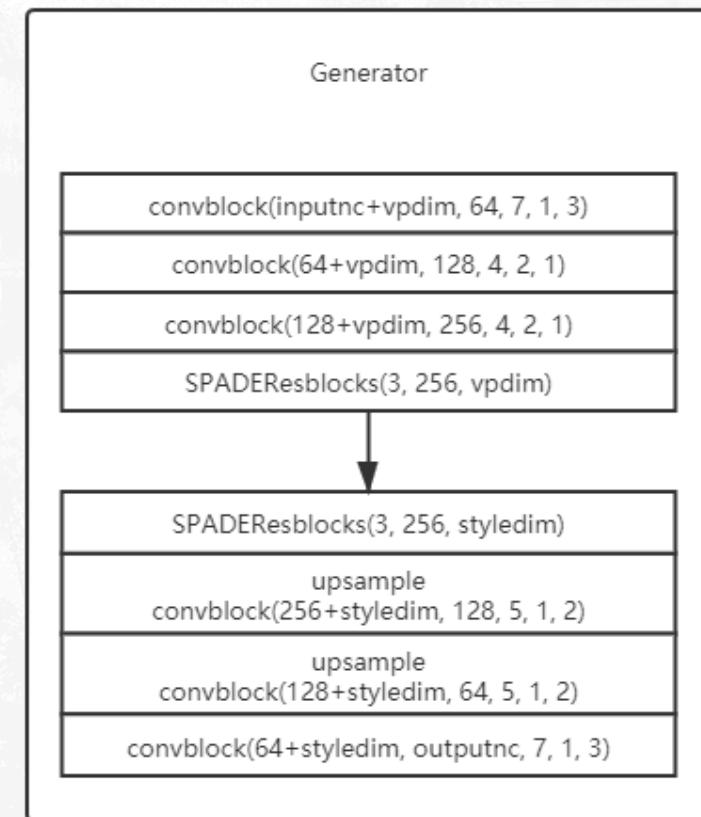
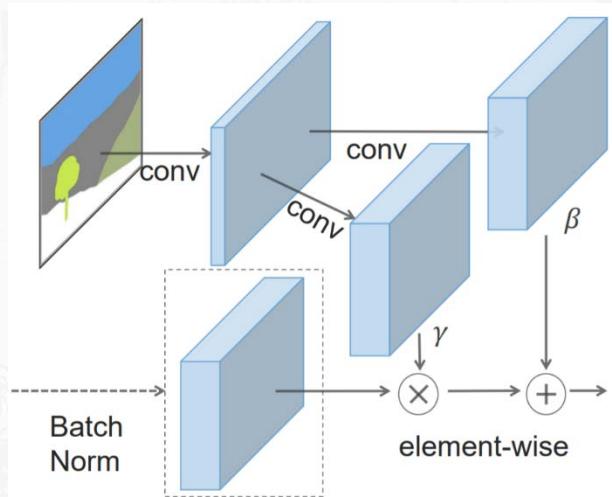
# 方法介绍

## 实现细节

需要实现的映射  $c = f(i_{RGB}, v), i_d = g(c, v), s = h(i_{RGB}, v), i_{RGB} = m(c, s, v)$

- 生成器

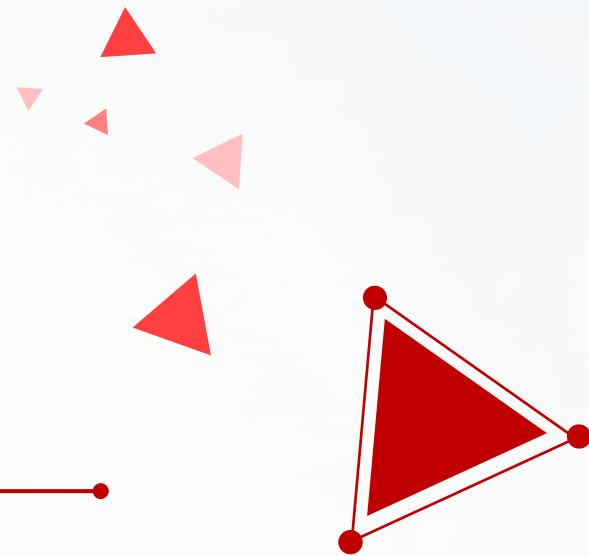
- 观察  $i_{RGB} = m(c, s, v) = m'(i_d, v, s) \quad i_d = g(f(i_{RGB}, v_1), v_2) = g'(i_{RGB}, v_1, v_2)$
- 用编码-解码网络实现
  - 编码部分提取出内容信息
  - 解码部分渲染目标图像： SPADE
  - 位置信息： CoordConv



# 04

*Part Four*

## 实验结果





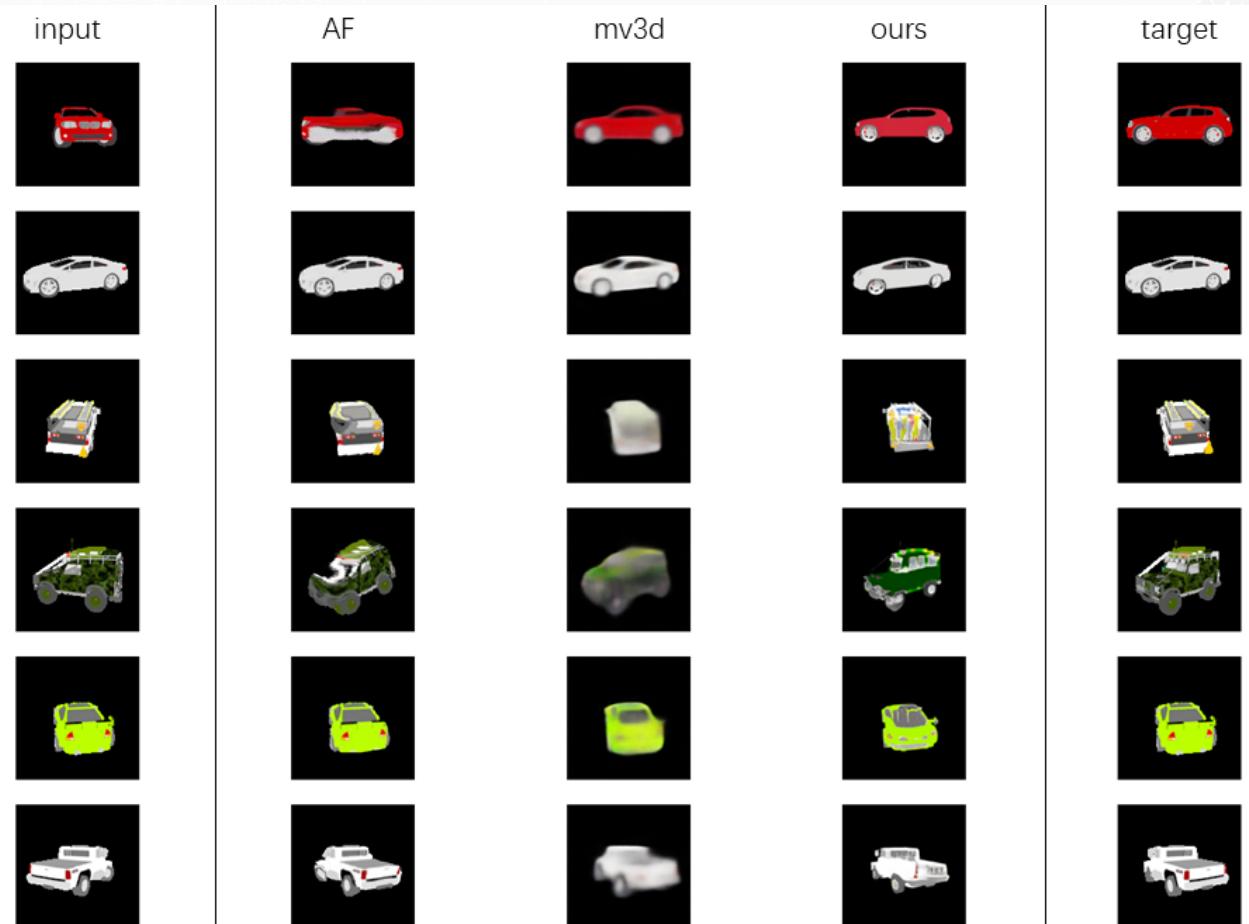
# 实验结果

## 视角生成

### 与视角生成经典方法比较

- AF: 预测原图到目标图片的流
- mv3d: 直接生成像素
- 定性定量结果有优势
  - 能够生成原图中缺失的像素
  - 生成图像更加清晰

方法	L1	FID
AF [25]	0.354	151.66
mv3D [47]	0.303	195.38
ours	0.208	13.05





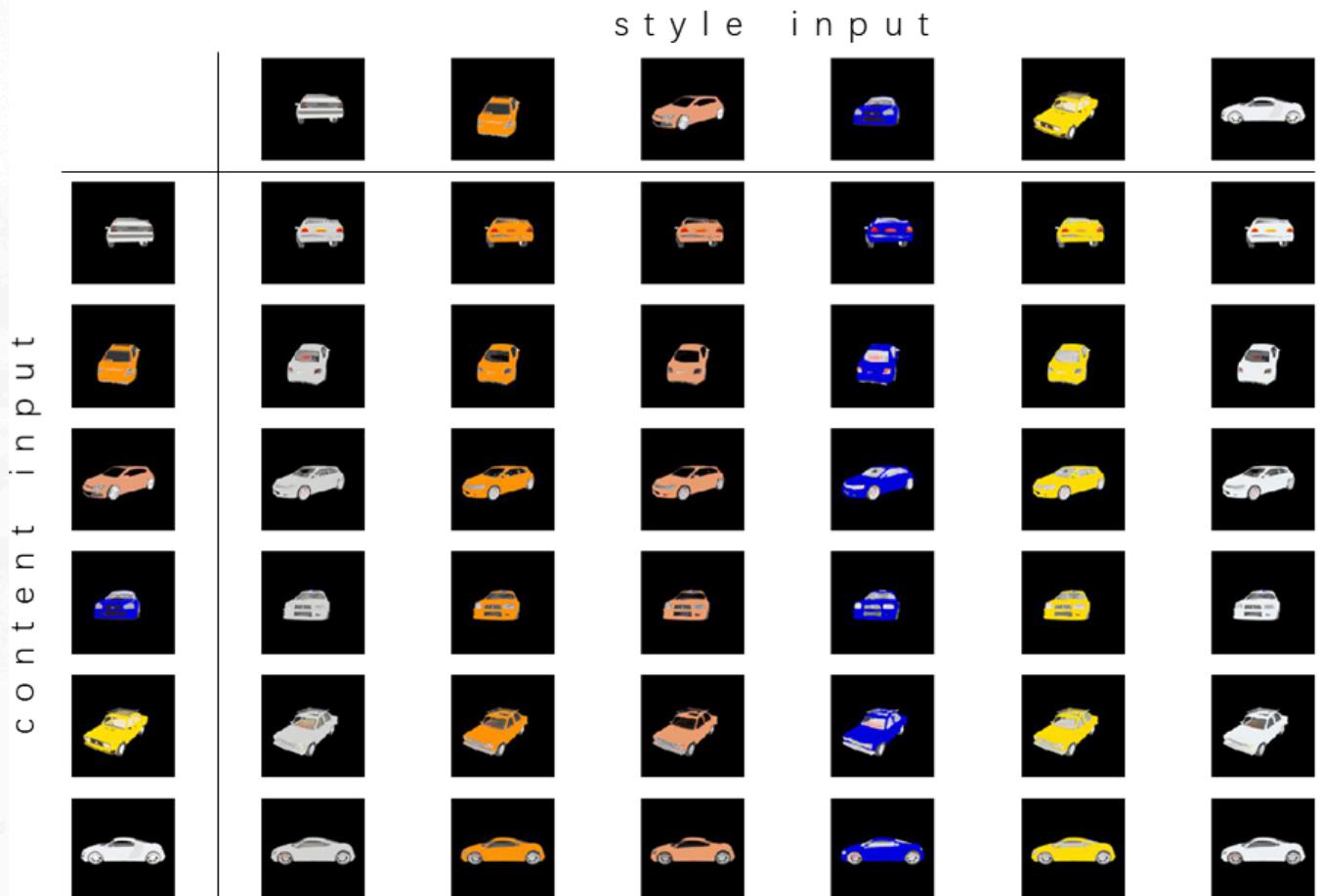
# 实验结果

## 材质迁移

### 材质迁移

- 能够迁移基本的材质
- 细节有不足
- 各种改进有效性

方法	L1	FID-视角生成	FID-材质迁移
baseline	0.671	205.109	206.636
above+depth map generation	0.586	96.953	104.857
above+SPADE+coordconv	0.293	53.818	89.826
above+realmng finetune	0.208	13.05	29.314



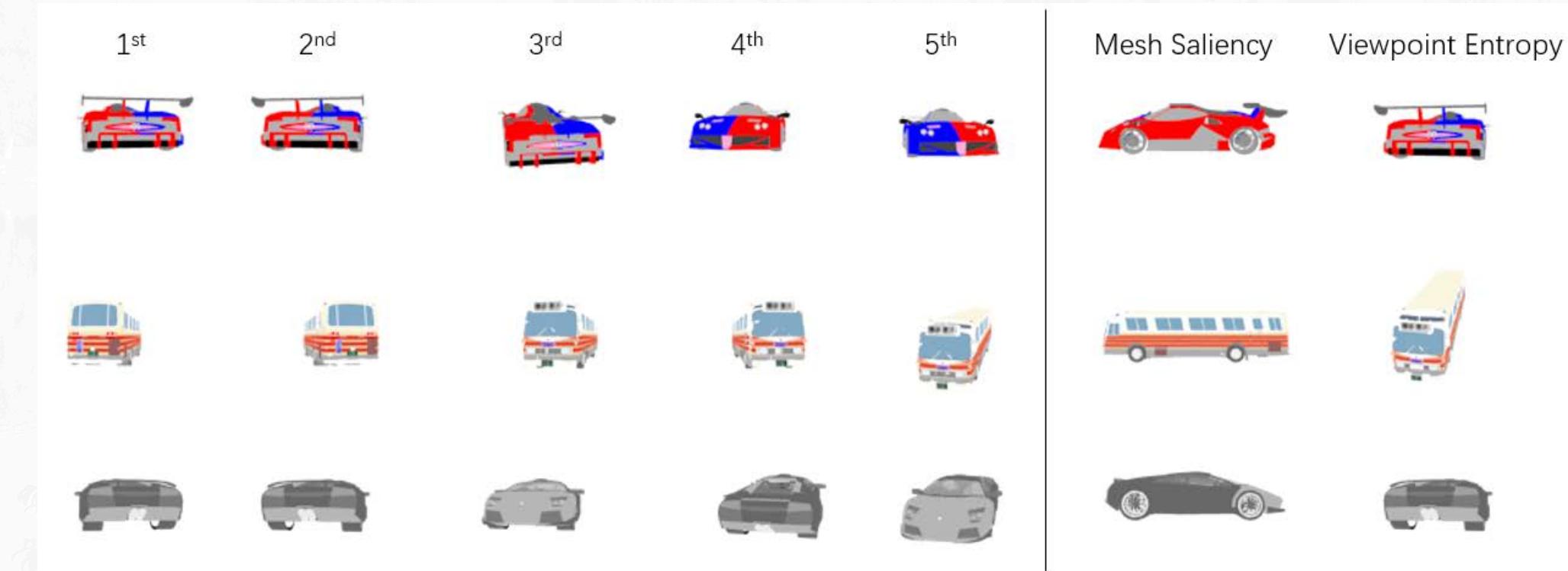


# 实验结果

## 最优视角选择

最优视角：包含信息最多的视角

- 重构其他视图误差最小的视角
- 考虑到材质信息也没有忽略内容的影响



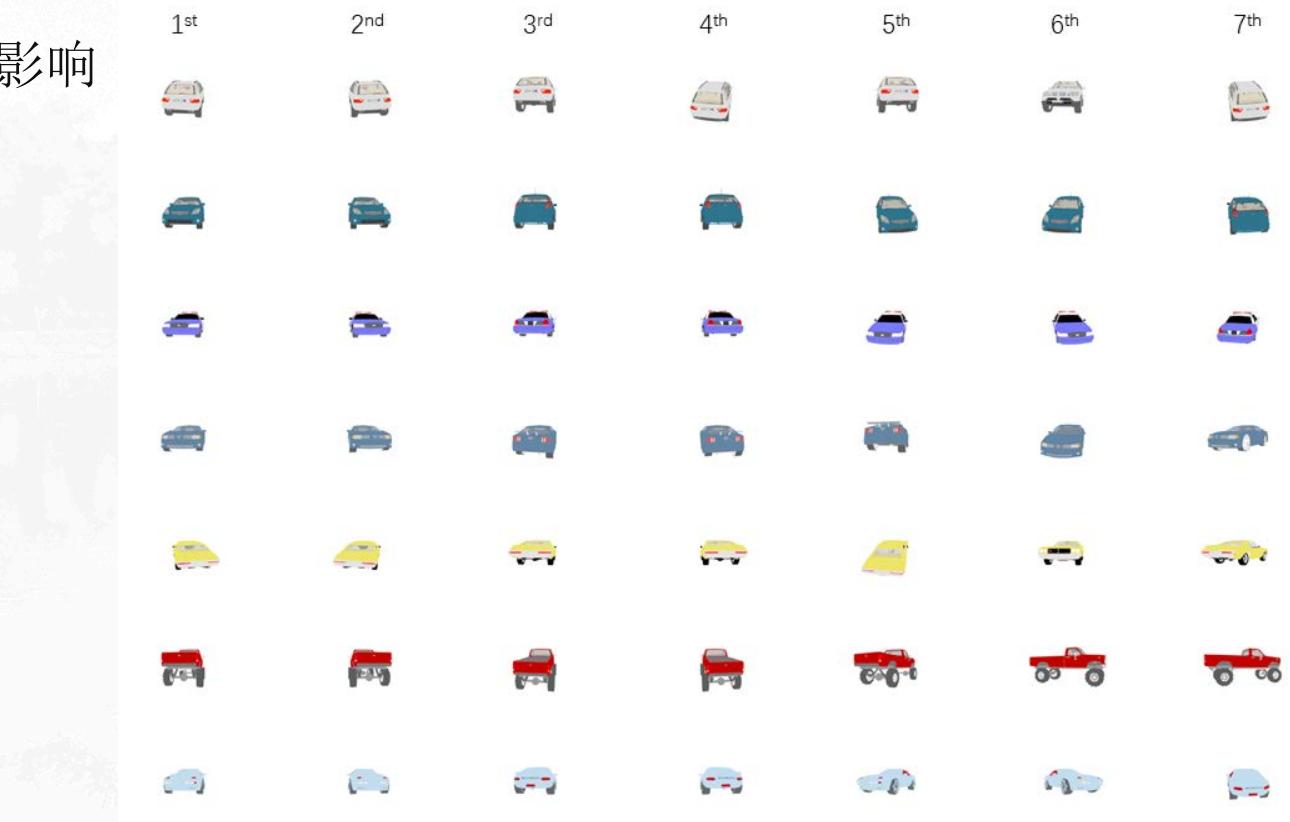
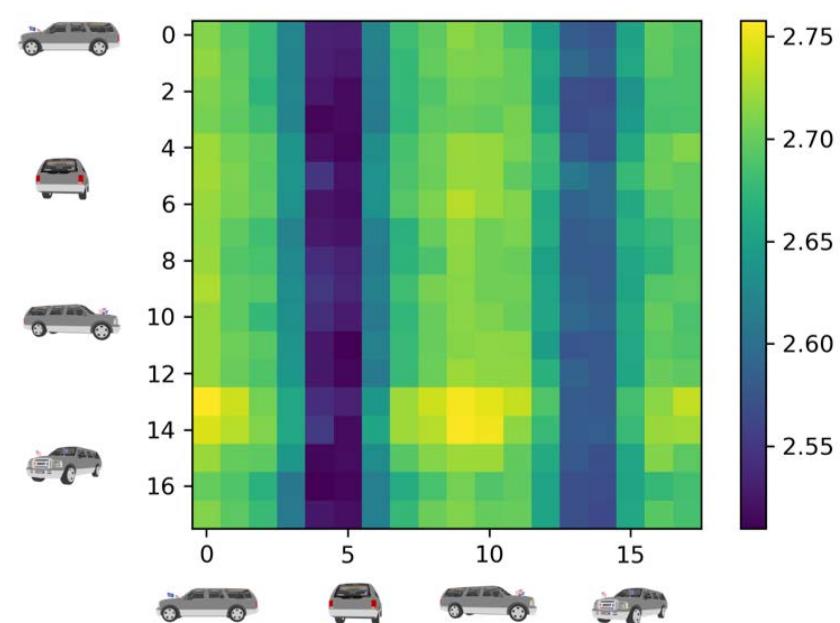


# 实验结果

## 最优视角选择

最优视角：包含信息最多的视角

- 重构其他视图误差最小的视角
- 考虑到材质信息也没有忽略内容的影响
- 分离隐变量使得这种定义有意义





# 实验结果

## 真实图片输入

真实图片：迁移视角/材质

- 提供自定义最优视角的选项
- 真实图片精修提高效果





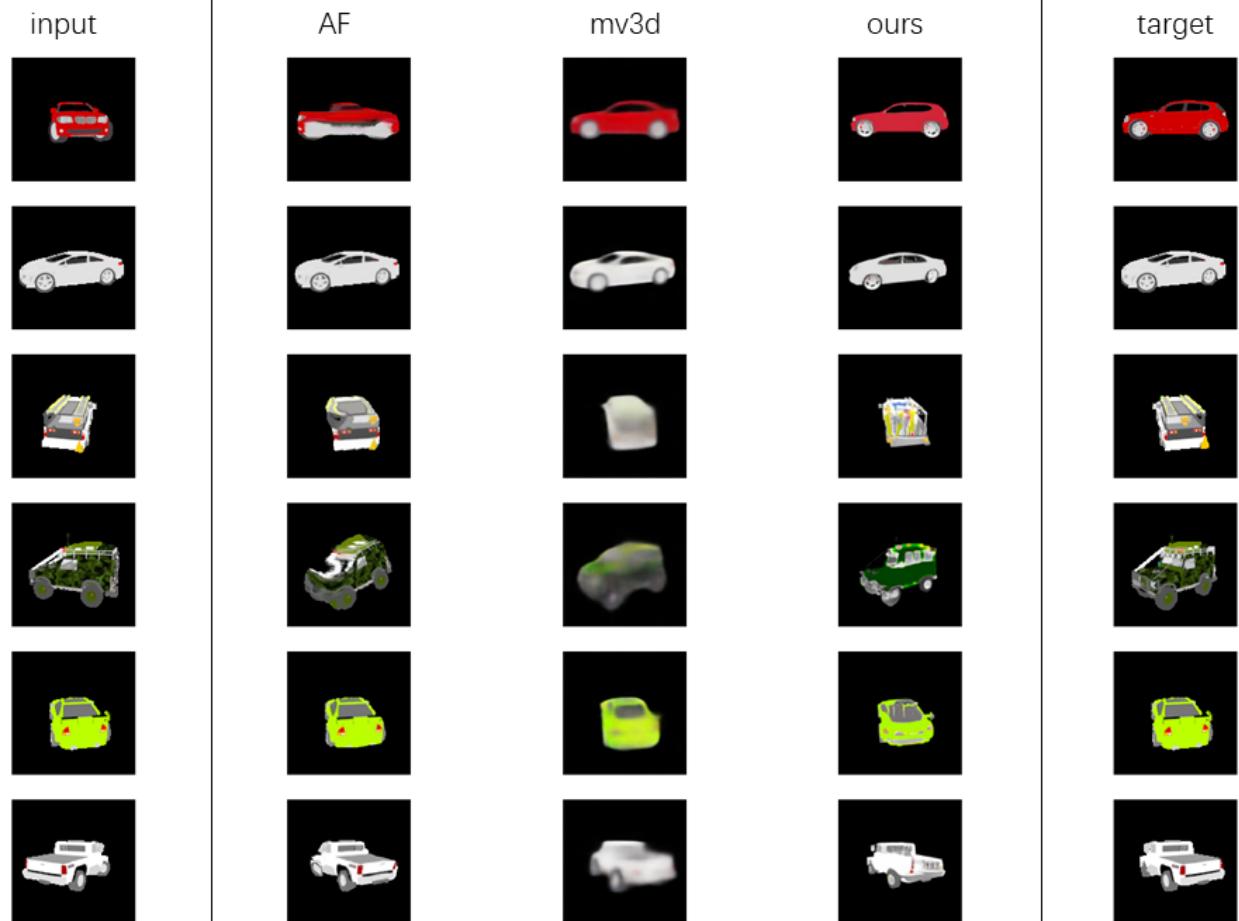
## 总结与展望

### 本文的贡献

- 提出了新的生成三维模型视图的方法
- 提出了新的选择最优视角的方法
- 实验证明这种方法的有效性

### 本文的不足：只提取全局材质

- 像素生成方法的通病
- 细粒度特征停留在浅层
  - 可能的改进：结合流方法
- 模型结构有优化空间
  - 图片上色结果辅助模型上色



**请您批评指正， 谢谢**