

Image2StyleGAN

这是ICCV2019的文章，从这篇论文中可以学习到如何将一张图片转换成潜在编码，只是这个方法不是实时的，后面有一些新的思路可以实时的操作。

Image2StyleGAN

2022年5月18日 19:41

将 Image 嵌入到 styleGAN 的 Encoding 空间当中。

1. 选取了在人脸上预训练的 styleGAN，选择了（人脸、猫、狗、绘画、汽车）各 5 张共 25 张图片。



在人脸上训练的 styleGAN，在其他物体上也可以使用 \Rightarrow 生成器的通用性。

2. styleGAN 嵌入的性质对仿射变换 非常敏感。



GAN 的泛化能力对仿射变换是敏感的，其学得的表示在一定程度上仍然是尺度和位置相关的。



但是对于图像的平移是很鲁棒的。

① 有利于图像的编辑

② 说明潜在空间没有要求必须生成完整的面。

③ 不同面部特征的潜在是相互独立的。

3. 潜在空间的选择

\rightarrow 扩展的 w 空间

$w, w \rightarrow$ 不容易 $\rightarrow w^+ (bn, 18, 512)$

在 styleGAN 的原始代码中 w 是通过 $(bn, 512)$ 广播到 $(bn, 18, 512)$ 的，也就是相同的 18 份 $(bn, 512)$ 。

但是这里学习的潜在空间 w^+ 是不同的。

4. 可以通过线性插值、交叉以及添加向量和比例差向量来实现 图像变换，风格迁移和特征迁移。

5. 方法

想法：从一个初始化 w 开始，^{根据} w^* 对应的 image 与原始 image 的相似性来反向插值。

Algorithm 1: Latent Space Embedding for GANs

Input: An image $I \in \mathbb{R}^{n \times m \times 3}$ to embed; a pre-trained generator $G(\cdot)$.

Output: The embedded latent code w^* and the embedded image $G(w^*)$ optimized via F' .

```
1 Initialize latent code  $w^* = w$ ;
2 while not converged do
3    $L \leftarrow L_{\text{percept}}(G(w^*), I) + \frac{\lambda}{N} \|G(w^*) - I\|_2^2$ ;
4    $w^* \leftarrow w^* - \eta F'( \nabla_{w^*} L )$ ;
5 end
```

https://blog.csdn.net/waixin_43135178

6. 损失函数

采用的 VGG-16 的感知损失和 MSE 损失的加权组合。

$$w^* = \min_w L_{\text{percept}}(G(w), I) + \frac{\lambda_{\text{mse}}}{N} \|G(w) - I\|_2^2 \quad (1)$$

$$L_{\text{percept}}(I_1, I_2) = \sum_{j=1}^4 \frac{\lambda_j}{N_j} \|F_j(I_1) - F_j(I_2)\|_2^2 \quad (2) \quad \Rightarrow \text{其中 } F_j \text{ 是 VGG-16 的 conv1-1, conv1-2, conv3-2, conv4-2 的输出特征}$$

$$L_{\text{percept}}(I_1, I_2) = \sum_{j=1}^4 \frac{\lambda_j}{N_j} \|F_j(I_1) - F_j(I_2)\|_2^2 \quad (2) \quad \Rightarrow \text{其中 } F_j \text{ 是 VGG16 的 conv1-1, conv1-2, conv3-2, conv4-2 的特征输出.}$$