Unsupervised Person Image Synthesis in Arbitrary Poses

**摘要**

本文提出一种利用GAN合成任意姿势的人的写实照片的新方法。给出一张人的照片（某种姿势），以及一个2D的骨架（类似火柴人），综合能看到的部分，对于不能看到部分进行猜测性的构建，最终给出这个人其它姿势的照片。其它相似的问题已经以有监督学习的方式被完成。我们主要采用无监督学习的方式来完成这样一个任务。本文创新点有二，一是构建了双方向生成器，映射回原始姿势，可以直接与输入的图像进行比较。二是设计了一种新的损失函数。最终得到的结果与有监督学习接近。

**关于GAN**

GAN是非常强大的基于博弈论的生成模型。同时训练一个生成器网络和一个鉴别器网络。GAN已经被证明可以产生非常逼真的图像，并具有高度的细节，同时GAN已经成功用于渲染人脸，室内场景和衣服。

**Method**

**关于Pose joint的位置**

这里有一个式子，表示Pose joint在点的概率。

**网络结构**

是生成器。目标是根据给定的图像I生成特定pose **p**的图像。

是2D骨骼关节的检测网络，用ResNet实现。

**三种损失函数**

1. Image Adversarial Loss（generator和discriminator的误差）
2. Conditional Pose Loss（生成Pose joint与给出的Pose p的误差）
3. Identity Loss（Identity损失）

最后的估计，需要完成。这个式子意为把产生的图片重新以最初输入的pose转换回去，以期达到的效果。

Identity Loss有两个子损失函数：

1. 内容损失

代表了z层的激活函数。

1. **图像风格损失（\*）**

这里使用Gram matrix来表示图像风格。（对纹理风格尤其）

Gram matrix是一种表示图像风格的矩阵。其计算方法为：

若，

采取均方误差。

两种损失采取加权和的形式，构成Identity Loss。

**全损失**

所有先前损失的线性组合。

**实验**

由参数z=7预训练得到。

Generator的learning rate选取0.0002。

Discriminator的learning rate选取0.0001。

300 epochs。

256256的图像。数据增强，每个图像进行三种可能的翻转（flips）。

使用Convolutional Pose Machine(CPM)产生2D-pose。CPM失败的图像被删除。

在剩下的图片图片中，选择24145训练，5000测试。

数据集：DeepFashion

使用pose conditional adversarial

Cycle-GANs

The loss function used in image style transfer using convolutional neural networks.

**疑问**

1. 为什么在Image Adversarial Loss中还放入的误差？
2. Gaussian peak是什么？
3. 中的分母是什么意思？