

# 第二周工作汇报

Shilong Zhang

8.27至 9.2

1.因为最后张老师决定做动作迁移，读了几篇与此相关的文章。2.，因为要用GAN，看了一些台大李弘毅的GAN相关的课程。3.学习了一些机器学习理论。

## 1 Paper Reading

### 1.1 Everybody Dance Now

#### 1.1.1 Abstreact

Transfer motion between human subjects in dierent videos.

#### 1.1.2 Method overview

pose detection  $\rightarrow$  global pose normalization  $\rightarrow$  mapping from normalized pose stick fgures to the target subject.

**Pose detection** Using a pretrained pose detector to get the representation of resuilting pose stick figure.

**Global pose normalization** calculate the scale and translation in the y direction (具体细节? )

## Adversarial Training of image to image Translation

### 1. pix2pixHD framework

D由三个不同规模的描述符构成，其应该趋向于真实的x与y打高分，x与G(x)打低分，即最大化 $L_{GAN}$

G则希望以假乱真，其希望D给G(x)打高分，同时最小化两个损失 $L_{FM}$ ,D给出的一个feature—matching loss,以及一个VGGNET给出的重建误差。

### 2. Temporal smoothing

改进G，其input变为这一次的pose stick figure( $x_{t-1}$ ) + 上次的 $G(x_{t-1})$ 来生成下一帧，而D则两帧一起评分。

### 3.Face GAN

To add more detail and realism to the face region.

有一点类似于残差网络的思想，加一个恒等映射，那么经过与D的对抗，生成的至少不会比映射前的差。

### 1.1.3 Experiments

质量评价中是让生成的图像用Detector检测出关键点，算这些关键点与原始pose的距离，为了避免有些关键点没有检测到的问题，只算全检测到的。

## 1.2 Dense Pose Transfer

### Abstract

Surface-based pose estimation + deep generative models,a two-stream architecture

warping module(基于UV贴图，包含较多纹理细节) + predictive module (Data drive)，具有互补的优点。

两个并行部分最开始都使用一个叫做Densopose的网络，他可以先把pixel分类到24个预先设定的身体表面模块，之后回归精确坐标 •

### 1. Predictive stream

a conditional generative model ,

### 1.3 机器学习理论

#### **PAC ((Probably Approximately Correct) Learning**

The learner: the strategy to select hypothesis from the hypothesis set

ML Pipeline: 如何根据Data Sampling 在 Hypothesis Set里面选择一个去逼近target function, 这三个因素也是决定机器学习效果的最主要的三个因素。注意越多的采样点可以让我们向target function有一个高概率的逼近。

**PAC learnability** H 是否可以在多项式时间内, 找到一个高概率的target function 的高精确度估计。

**Sample complexity** 保证 PAC Solution 的最少训练样本。

**Consistent Learner** Learner 选择的假设完美的符合训练数据。

**PAC Bound** Version space 里面, 至少出现一个在真实分布上的错误率小于 $\epsilon$ 的概率。