

2022.01.02周报

纪新龙

本周计划任务

转入研究人脸特征追踪、定位、建模。

具体完成情况

重温之前读过的两篇硕士、博士毕业论文中相关部分。

下周计划任务

暂时未定

具体完成情况

毕设的框架

张磊老师在本周三下午的演示中问我，我的毕设框架是什么。我的毕设框架是如下一个三步走：

1. 学习图像的本征图分解，进而扩展到视频，实现一个高效的光照分离算法
2. 学习人脸的三维建模，及特征点对齐，实现一个换脸算法
3. 学习重光照理论，实现一套高真实感的视频换脸算法

我没有把周三下午的演示当作一个开题报告，所以就没有把这个框架写进我的ppt。

第一步的工作我已经做了很多了。目前我在第二步。我觉得第三步不会很难。

黄老师说，明确毕设的框架，是为了避免“这边做一点，那边做一点”，零零散散的做点事情然后拼到一起，而不是做一个系统性的工作。其实实现的三篇算法，每篇都花两周时间这样。对我来说意义主要是，在动手实现的过程中更深地体会了文章中的意思。光读论文容易读得粗，遗漏东西，动手实现的时候才倒逼回去更细地理解文章。第三篇针对视频的分解，最终效果一般，大部分先验是和图片分解一样的，只加了一个帧间一致性约束。然而帧间一致性约束的效果并没有显示出来，这么看是耽误了时间了。但是复现过程也有收获。视频分解的复现，主要让我收获了checkgrad函数的使用，以及倒逼我花时间去思考了几种优化方法，想通的东西以后将不会再有疑惑。

有些时候算法复现出来但是参数调不对，就会导致达不到想要的效果。就像第一个算法，当时效果也一般，后来我回去重看的时候调整L0系数，意外获得一个理想的效果（本来我以为L0项作用不大的）。

关于换脸算法

这周做的事情不多。周中时候身体不好，没兴趣看论文，就调了调之前复现过的三个算法的代码。以及本科期间最后一个综合课设到了结题时间了，三个组员刚刚考研结束，我作为组长帮他们一起给项目收尾，最终顺利通过验收了。后边除了毕设和读研就没有其他事情会打扰我了。

我的毕设题目是在视频换脸中用本征分解来改进效果。但是我在搜“换脸”相关算法的时候，发现这个词在学术上不止一种含义。只不过之前我的重心在本征图分解上面，就忽略了这一点。现在我做到这个地方了，我想整理一下这个地方的思路。

1. 基于3D模型和光流的方法

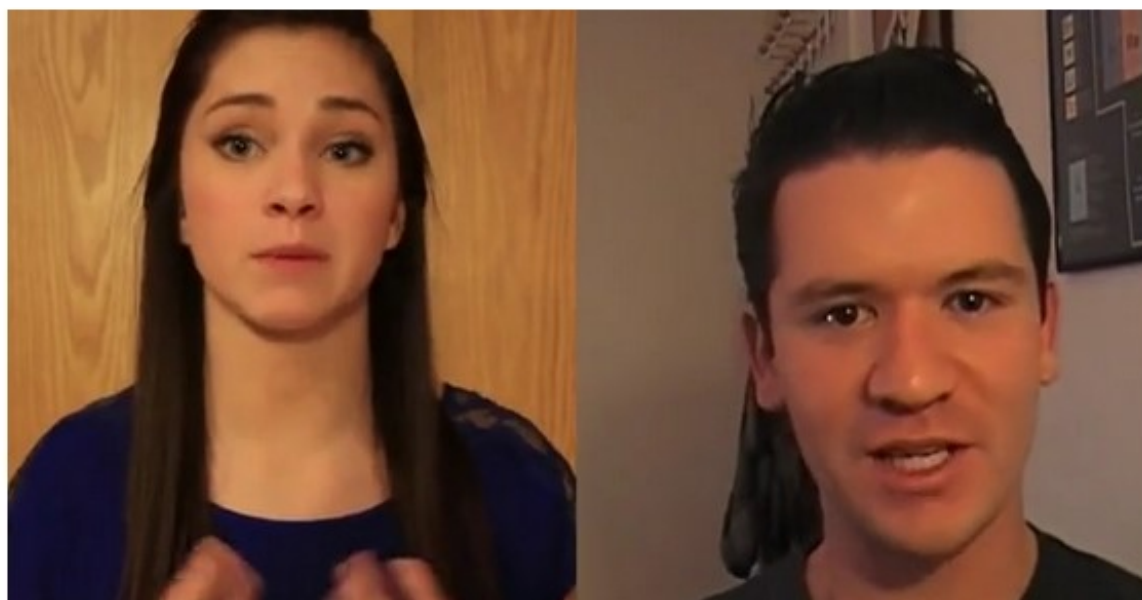
早期的“换脸”其实是面部表情迁移算法，被换脸的人脸部皮肤和形状没有被改变，但是他的表情被换掉了；

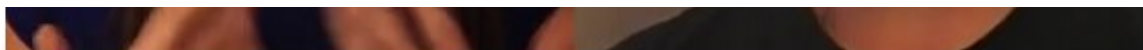


2. 基于形状拟合的算法

基于形状拟合的方法是最直观的换脸算法，即检测到的关键点然后计算两个人脸形状之间的变形，再添加图像融合等后处理技术，目前在天天P图等应用中的换脸算法如此。

Before...





After...



3. 基于风格迁移的算法

换脸算法保留了被换脸的表情，姿态，形状，同时使用了新脸的纹理，要实现的就是纹理的迁移，这可以使用已经较为成熟的风格迁移算法。



(a)

(b)

(c)

第一种“换脸”只换了表情，第二种既换了表情又换了皮肤，第三种换脸不换表情只换皮肤。 还有许许多多其他方法，但最终效果都可以归入这三种。毕设的问题描述中说“现有的人脸图像的应用，如人脸识别、换脸等，往往对光照影响考虑不充分，或者仅做简单的处理，导致效果不理想”。我认为只有换掉被换者的大块皮肤的时候，光照才会对结果产生很大影响。也就是说，上述我想要改进的换脸应用是上述总结中的第二、三种“换脸”。在这个地方我想明确一下换脸指的是哪种换脸，如果我理解错了请指出。

人脸三维重建技术

目前人脸三维重建技术主要包含两大类。第一类是基于三维形变模型(3D

morphable model)的重建技术。这类算法从一个基于统计原理得到的人脸几何结构表达出发, 根据目标人脸特征点的位置以及可能存在的表情对已知的平均人脸几何模型G进行变形, 从而得到一个符合目标人脸姿态、表情的三维模型。这类算法旨在得到大致的人脸几何结构, 并不专注于对人脸皮肤细节的恢复。

第二类算法主要依赖于从明暗恢复形状算法(Shape from shading)及其变形来恢复人脸的三维结构。除了单张图片的应用, 也有学者将目光聚集到了个人图片集或者视频流上。相较第一种算法, 基于从明暗恢复形状的算法可以更加准确地恢复人脸皮肤上诸如皱纹、粉刺等细小特征的几何结构。

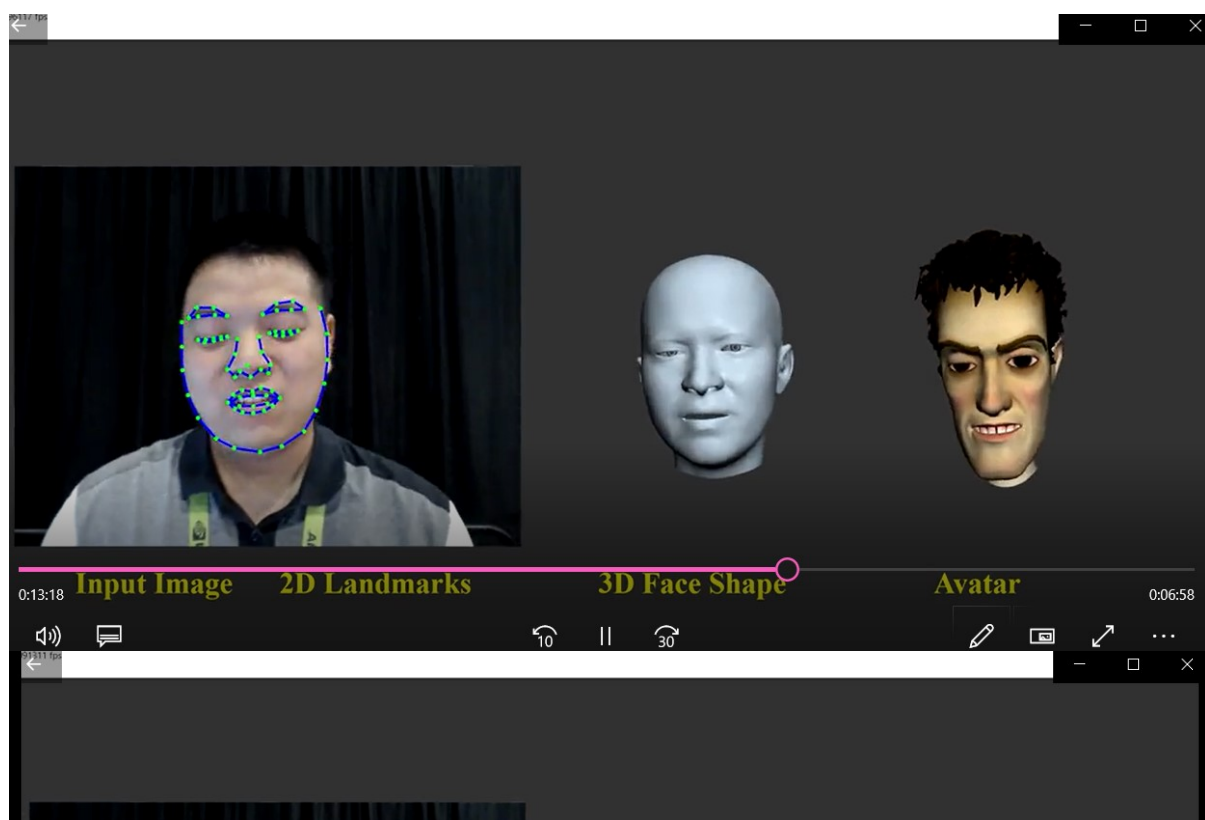
浙大的DDE

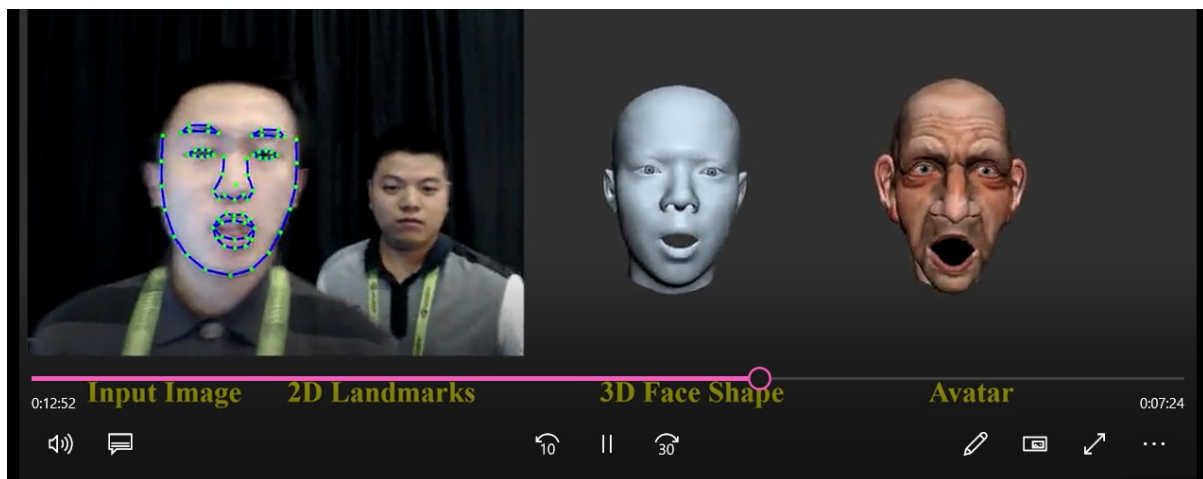
之前的我看到浙大的文章

Displaced Dynamic Expression Regression for Real-time Facial Tracking and Animation

之后, 对他们算法的效果印象很深。且他们的算法涉及人脸特征点的追踪, 在演示视频里也有一个三维模型和动画三维模型, 我就以为他们的这个方法可以为我所用。但是这周有了上述思考之后, 我又回去看那篇论文, 我发现他们提出的DDE回归方法的亮点是(免摄像头校准地)从实时视频中**追踪人脸的表情**, 并在这个过程中校准摄像头的参数。也就是说他们论文中对人脸三维模型的重建其实是不存在的, 他们只是**从视频中定位人脸的特征点并以此追踪到人的表情, 再把表情显示到一个通用的人脸3D模型上**。

下图是他们的两个不同脸型的演示者, 可以看到中间的那个人脸3D都是一样的, 再次说明他们不在乎脸型这种更细节的形状特征, 他们的DDE算法在乎的只是人的表情。

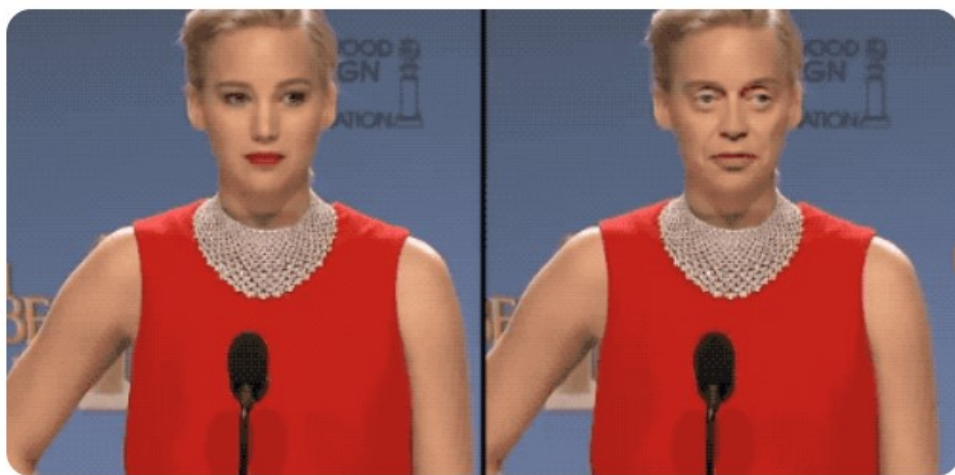




这种对人脸三维模型的重建，不是完全的重建，而只是表情这种局部特征的抽象。这和我在上述“关于换脸算法”中列举的第一种只迁移表情不迁移皮肤的算法在所谓的人脸3D重建方面是一个思路。如果我跟踪他们的DDE这个算法，确实能学到针对表情的人脸重建和人脸特征点追踪方面的东西。但是在“关于换脸算法”那一节，我认为表情迁移不会太受光照影响，受光照影响比较大的换脸应该是那种换掉皮肤的换脸。

从“光照不自然”角度对换脸的认识

我搜到一篇名为“基于光照方向一致性的换脸视频检测”的文章，说的是针对大部分Deepfake 换脸视频与真实视频在外部成像环境上的差异，提出一种基于光照方向一致性的换脸视频检测方法。这似乎是与我们的课题“视频人脸的本征图分解及在换脸中的应用”针锋相对的鉴伪技术。我们希望改进光照对换脸的干扰，他们希望利用未处理好的光照揪出被换脸的视频。下图是Deepfake的效果图，可以看出Deepfake是一种保留被换脸者的表情，替换皮囊的换脸。左图在光照上确实有点不自然。



使用换脸技术，导演史蒂夫·布西密的脸被换到了詹妮弗·劳伦斯的身上，但是布西密的动作和表情都是劳伦斯的。

总结

本周工作量不多，主要精力在了解换脸算法的不同技术路线上，这是为了确定以后要做的事。

基于目前我对换脸、浙大DDE、《基于光照方向一致性的换脸视频检测》的进一步了解，基于我目前的认识水平，我认为有两条模糊的路线：

1. 传统的基于人脸建模，跟踪人脸的表情；利用特征点对齐，将别人的表情换到被换脸者的人脸上。保持被换脸者的皮肤、衣着、背景、光照等不变。
2. 从Deepfake这种深度学习入手。目前搜换脸，似乎各种神经网络是最多的，ai一直比较火。但是我大致看了Deepfake的内容，和其他神经网络一样，不太会借助人脸建模这种传统的建模方法。由于其不可解释性，人脸特征点这种信息也是糅合在网络里成为各种不可见的参数存在的。我对神经网络理解有限，我知道它需要大量数据来训练，效果可能也比较好，但是如果专注于训练网络可能就不会专门去走规划中的人脸三维建模、人脸特征点对齐等技术了。我听说ai的门槛比图形学低，我想学门槛高的东西。但是全世界都在做ai，这个势头也不好避免，而且目前我看到的“换皮肤”的换脸几乎都是用神经网络做的（我查的资料有限），个人看来换皮肤式的换脸更需要引入本征分解弱化光照。或许表情迁移式的换脸也会有光照的问题，我不确定。

关于两个路线，希望老师给点意见！有认识不对的地方，也请指正。