**Disentangled Representation Learning for 3D Face Shape**

本文，我们提出一种新的3D人脸分离表示。具体地说，给定的3D人脸形状被分解为身份部分和表情部分，它们皆用非线性方式编码。为了解决这个问题，我们提出了一种3D人脸网络的属性分解框架。为了更好地表示不同旋转角度的人脸，人脸形状通过基于顶点的形变特征而不是欧式空间坐标来表示。实验结果描述了我们的方法比我们基于身份和表情部分效果更好。而且，我们方法比现存方法可以实现更自然表示迁移结果。

**1 Introduction**

3D人脸模型是由几部分构成的如身份、表情、外观、姿势等，3D人脸形状是由身份和表情属性决定的。将3D人脸形状解耦为两部分是计算机视觉领域一个重要难题，但是它可从其他应用中得到启发如面部组件转移、人脸动画、化身动画等。这篇文章的目的为3D人脸形状开展一个属性分解模型使给定的人脸形状可以较好的通过身份和表情来表示。

在本文中，我们为3D人脸形状构建一个强有力表达的非耦合参数空间。一些传统线性方法已经分解标识和表情两部分，当时他们被线性模型的表达能力所限制。尽管深度学习被认为潜在的提升方法，如何设计深度学习方法并不明确，如神经网络结构、3D人脸形状特征表示。此外，另一个挑战是如何利用现有人脸数据集进行身份和表情的学习。

为了重新描述这个问题，假定身份和表情被独立编码为zid 和 zexp向量，S¯+AidZid +AexpZexp

S¯是平均形状。基于顶点的三维形状变形表示方法，捕捉局部变形梯度同时基于顶点定义。本文的主要贡献包含下面几个方面：

1）我们提出通过非耦合方式学习3D人脸形状，从语义层面对身份和表情进行编辑。

2）我们提供一种新颖的非耦合框架来定义3D人脸网络。基于顶点的形变表示可应用于我们的框架，同时它比欧几里得坐标有更好地表现。

3）实验结果表示我们的方法能够更好地解耦合身份和表情。因此，基于我们方法的表达传递可以得到更好地结果。

**2 相关工作**

线性3D人脸模型 3DMM的相同工作，线性参数模型被广泛用来表示3D人脸。Vlasic等人提出多线性模型来解耦合到不同模型，Cao等人采用双线性模型通过身份和表情参数来表示3D人脸。最近，其他方法也被提出来提升准确率，如通过更大规模数据集来提升3DMM的能力，或者使用立体表情来更好地捕捉面中部。

非线性3D人脸模型 最近，一些研究提出了基于深度学习的非线性参数模型嵌入三维人脸形状。Liu等提出多层感知残差模型表示3D人脸。Tran等对3D人脸提出编码-解码结构，其为非线性3DMM的一部分。Bagautdinov提出一种组合多变量自动编码结构来表示不同层的几何细节。Tewari等通过自监督方法生成3D人脸。Anurag等提出基于图卷子自动编码的3D人脸表示。这些工作采用深度神经网络来学习隐层参数空间来表示3D人脸，但是上述皆没有考虑人脸属性解耦合的问题。

深度学习进行3D人脸分析 深度学习近些年来吸引了越来越多的注意力。

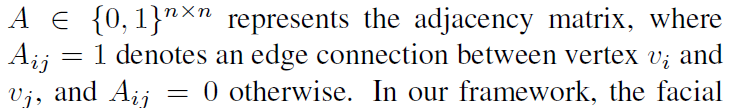
**3 解耦合3D人脸表示**

**3.1 概况**

给定3D人脸集合，我们的目标是得到一个身份和表情的紧凑表达方式。在表情分析中常用观察方式是人类表情依赖高维曲面，如图1所示，每个表情流被渲染为黄色。由于不同个体表情流为相同的，个人表情流可以用相同表情的均值脸表示。另一方面，每个个体都有他的中性表情，其可作为每个流的原始点，被用来表示他/她的身份属性。同样地，在均值人脸上的相同表情表示他/她的表情属性。这两个网络被分别定义为身份网络和表情网络。

基于这个观察，我们的非耦合3D人脸表示包括两部分：分解和融合网络。分解网络解耦合属性通过将输入人脸网络分解为身份网络和表情网络两部分。融合网络从身份网络和表情网络中恢复原始人脸。

我们用图结构定义人脸网络，图的定点为V， 边为 M = (V;A)，|V|= n .



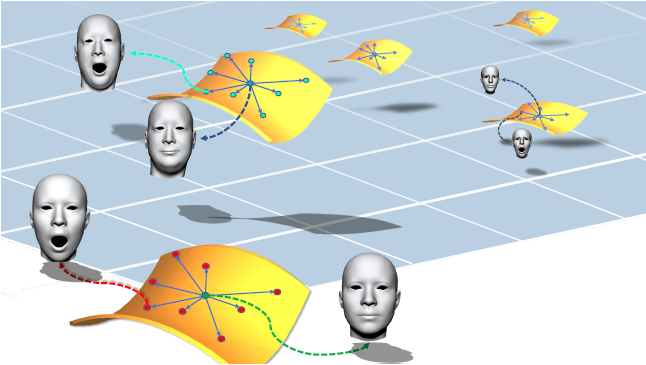
在我们的框架中，训练集中的人脸网络包含相同连接，同时每个顶点与一个特征向量Rd相连。

C://Users/DELL/AppData/Local/YNote/data/sina2610774271/96d803b6433246eea724919184e35f96/clipboard.png

在我们提出的方法中，3D人脸网络由身份表示网络和表情网络两部分组成，

C://Users/DELL/AppData/Local/YNote/data/sina2610774271/0f7c5610e6ce4b8e820d98172a76d632/clipboard.png

上述三元组被用来训练我们的网络。



**光谱图卷积** 如对常规2D图像做的卷积操作，我们采用图卷积操作，光谱图卷积，用来提取网络上有用的顶点特征。我们首先提供一些关于此卷积的背景，更多细节可以在下述文章中发现。

正如我们定义网络 M=（v, A）,正则化拉普拉斯矩阵可以定义为 L = I - D-1/2AD-1/2,其中D是度矩阵，特别的，一个对角矩阵，I为单位矩阵，光谱图卷积定义在图的傅里叶变化域，在其中拉普拉斯矩阵L，特征向量为U

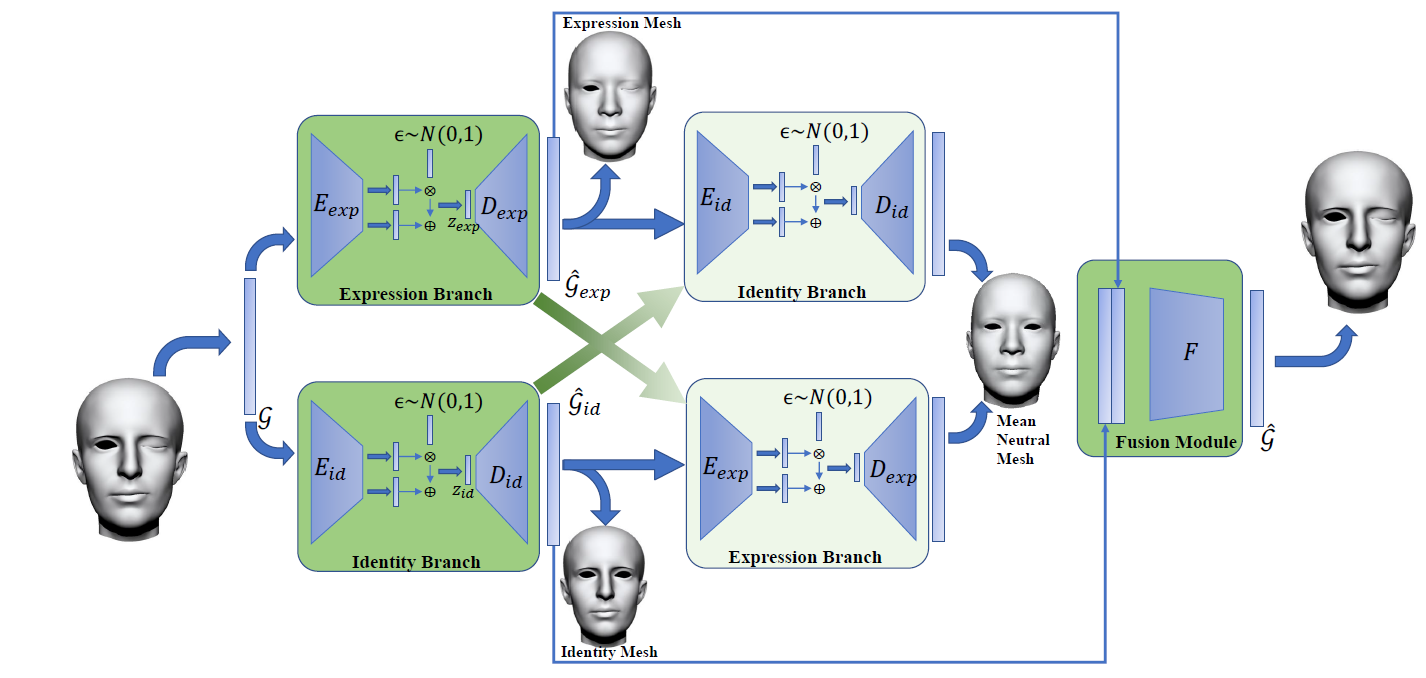
在傅里叶空间的卷积被定义为

**变形表示法**在现有3D人脸表示工作中，在三维空间中的欧几里得坐标是最常用的顶点特征。基于光谱图卷积，我们可以使用定义在顶点上的其他特征。正如在[24]中指出的，光谱图卷积是拉普拉斯平滑的一种特殊形式。

**3.2 分解网络**

输入分解网络是3D人脸网络可变表示特征，我们的目标是将它解耦合为身份和表情属性。将网络M匹配到其他两个三元组(Mid,Mexp)。

分解部分包含两个并行相同结构的网络，一个来提取Mexp的表情网络，另一个来提取身份网络Mid。讨论身份分支作为一个例子，



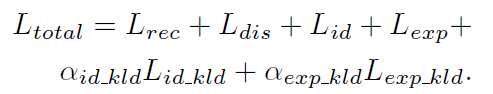
**3.3 融合网络**

正如描述，从分解和融合属性重建原始输入是必要的。因此，我们自然提出一个融合模型来融合身份和表情网络对来重建。这个模型保证我们的分解本质上是无损失的。由于网络结构组是同构的，我们可以通过聚焦于从身份和表情网络获取的顶点特征来得到一个新的图。新图与原始输入有相同的边集和顶点集，除了在每个顶点上的2d维度特征。

**3.4 训练过程**

我们首先序列化的训练分解网络和融合网络。C://Users/DELL/AppData/Local/YNote/data/sina2610774271/6ad00244731441fd862ce93a345aea7a/clipboard.png

解耦合保证身份部分不包含表情信息，同时表情部分不包含任何身份信息。总结，完整损失函数如下所定义：



**数据扩增**

我们开发了一个新的数据扩增方法来克服过拟合问题。给定训练集中，m个身份样本