

硕 士 研 究 生 读 书 报 告



题目《ImFace: A Nonlinear 3D Morphable Face Model with Implicit Neural Representations 》 读书报告

作者姓名 周强

作者学号 22251038

指导教师 李启雷

学科专业 软件工程

所在学院 软件学院

提交日期 2022 年 月 日

摘要

本文将会介绍《ImFace: A Nonlinear 3D Morphable Face Model with Implicit Neural Representations》一文的内容，主要包括一下几个方面：1、文章所要解决的问题；2、解决问题的方法；3、论文的创新点。

# 1.引言

3D可变形面部模型（3DMM）是一种效果良好的统计模型，可以用于渲染各种各样的真实面部。该模型是通过从具有密集对应关系的一组样本中学习面部形状和纹理的先验分布的技术来建立。

在3DMM中，最根本的问题在于生成潜在的可变形表示的方式，在过去二十年中，随着数据在规模、多样性和质量方面的改进，已经取得了显著的进展。这些方法最初是基于线性模型的，并进一步扩展到基于多线性模型的，其中不同的模式被单独编码。不幸的是，由于线性模型的表示能力相对有限，这些方法在处理具有复杂变化（例如夸张表达式）的情况时效果不佳。在深度学习的背景下，通过使用卷积神经网络或图神经网络，在2D图像或3D网格的输入下研究了许多非线性模型。它们确实带来了绩效收益；然而，受限于输入数据上离散表示策略的分辨率，面部先验不能被充分捕获，从而导致形状细节的丢失。

本文提出了一种新的3DMM，即ImFace，该模型通过学习INR（内隐神经表征）对传统的3DMM进行了大幅升级。为了捕捉非线性面部几何变化，ImFace构建了单独的INR子网络，以显式地将形状变形分解为分别用于身份和表情的两个变形场，并引入了改进的嵌入学习策略来扩展表情的潜在空间，以允许更多样化的细节。通过这种方式，可以对个体间差异和细粒度变形进行精确建模，这同时考虑了应用于相关任务时的灵活性。此外，受线性混合蒙皮的启发，本文提出了一种神经混合场，以将整个面部变形或几何体分解为由一组局部隐式函数编码的语义有意义的区域，并通过轻量级模块自适应地混合它们，以减少参数实现更复杂的表示。本文还设计了一种新的预处理管道，该管道绕过了现有基于SDF的INR模型中对不透水面部数据的需求，并适用于各种面部表面，即硬件采集或人工合成。

# 2.实现方法介绍

利用INR学习非线性3D可变形人脸模型。ImFace明确地将面部形状变形分解为分别与身份和表情相关的两个单独的变形场，并学习深度SDF（Signed

Distance Function有符号距离函数）来表示模板形状。

所有字段都与一系列局部隐式函数混合，以获得更详细的表示。

## 2.1分离的INRs网络

INRs的基本思想是训练神经网络以拟合连续函数，该函数通过水平集隐式表示曲面。该函数可以多种格式定义，例如占用率、SDF或UDF。我们利用基于表情和身份的潜在嵌入的深层SDF进行全面的面部表示。它输出与查询点的带符号距离s：

其中是查询点在3D空间中的坐标，和分别表示表达式和身份嵌入。

我们的目标是学习一个神经网络来参数化，使其满足真实的面部形状先验。Imface的网络由三个Mini Nets块组成，分别为Expression Mini-Nets (ExpNet)、Identity Mini-Nets（IDNet）、Template Mini-Nets (TempNet)，它明确地解开了面部形状变形的学习过程，确保可以准确地建模个体间差异和细粒度变形。前两个网络分别学习与表情和身份变化相关联的单独变形场，模板迷你网络学习模板面部形状的带符号距离场。如图1中（a）所示。

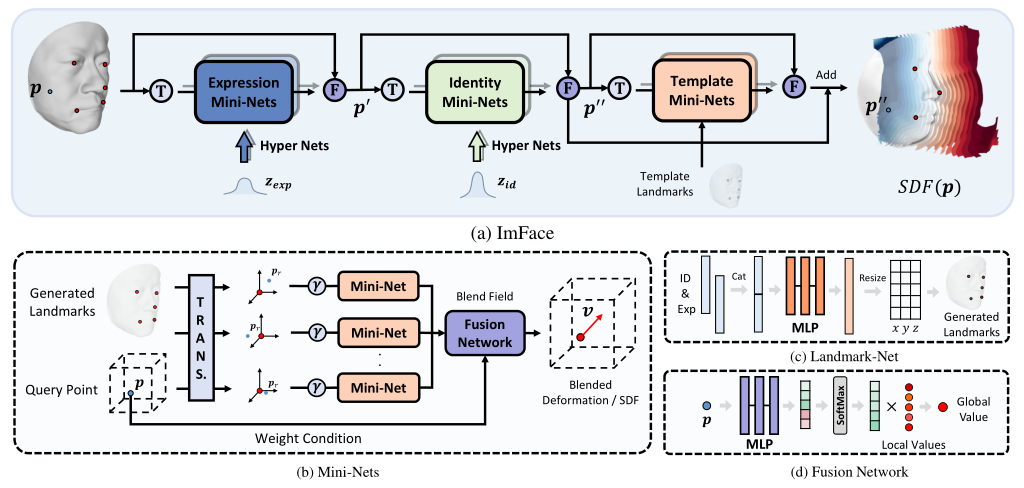


图1. ImFace概述

## 2.2神经混合场

Mini Nets块是三个子网络、和共享的通用架构。它学习连续场函数：， 以生成用于全面面部表示的神经融合域。特别是，为了克服单个网络的有限表达性，本文将人脸空间分解为一组语义上有意义的局部区域，并在混合之前分别学习（例如变形或有符号距离值）。这种设计的灵感来自最近INRs对人体的研究，该研究引入了线性混合蒙皮算法，以使网络从身体部位的单独变换中学习。为了更好地表示详细的面部表面，我们用替换原始线性混合蒙皮算法中的常量变换项，并将神经混合场定义为：

其中是描述第个局部区域的参数，是第个混合权重，是相应的局部场。以这种方式，在一系列局部场上执行混合，而不是计算一些固定位置的输出值的加权平均值，从而在处理复杂的局部特征方面产生更强大的表示能力。

## 2.3改进的表达式嵌入学习

INR目前广泛采用一种自动解码器框架，以联合学习嵌入和网络工作参数。在前面的尝试中，每个表达式类型都通过一个嵌入进行编码，以使属性解开。不幸的是，这样的嵌入仅仅能够表示一种表达类型的平均形状变形，使得学习到的潜在空间无法捕捉个体之间更多样的变形细节。为了避免上述困境，我们改进了学习策略，将每个非中性人脸扫描视为唯一的表达式，并为其生成特定的嵌入。通过这种方式，潜在空间显著扩展，这使能够表示更细粒度的细节。另一方面，存在一种潜在的失败模式，即身份属性再次陷入表达式空间，崩溃为身份映射。为了应对这一挑战，当当前训练样本是中性面时，我们抑制，写为：

其中表示来自中性面的点。通过应用这种学习策略，和共同学习中性面上的形状表示，而只关注表情变形。此外，在训练过程中只需要中性标签，绕过密集的表达标签

## 2.4数据预处理

由于神经网络擅长拟合处处可微的函数，当前对隐式函数的研究通常需要严密的输入。虽然像UDF这样的功能不需要水密性，但它们在穿过表面时是不可区分的，并且在处理细节方面也不太称职。

本文提出了一种有效的预处理管道，该管道生成伪水密面并在其上定义通用SDF，以便可以像在水密对象上一样精细地学习几何图形和对应关系。

伪防水面生成。使用界标将面部与正面严格对齐，并将每个网格标准化为10cm的单位。坐标原点设置在鼻尖后4cm的点，然后将半径为10cm的球体定义为裁剪掉外部网格三角形的采样区域。射线三角相交算法用于去除隐藏表面，如鼻腔和口腔，然后在定向和伪水密网格的x-y坐标上执行Delaunay三角化算法。

面部曲面上的SDF计算。生成伪防水面后，可以通过对其进行距离变换来计算SDF值。样本的符号仅由其到最近曲面的距离向量与z轴正方向之间的角度确定。面部曲面后面的坐标值定义为负值。统一采样每个面部表面上的250000个点和球体中的15000个点，并计算它们的带符号距离和梯度向量。采样数据最终被公式化为用于ImFace训练的三元组（查询点、梯度向量、有符号距离值）。

# 3.本文创新点以及贡献

本文提出了一种新的基于INR的3DMM，它通过两个明确分离的变形场来编码复杂的面部形状变化，以细粒度和语义有意义的方式学习强大的表示。

同时本文提出了一个有效的预处理管道，它为非水密3D面定义了通用SDF，使INR能够首次使用它们。

文章通过实验证明，ImFace在合成具有可信细节的高质量3D人脸方面具有优势，在3D人脸重建方面优于最先进的同行。