

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 田旭泽

作者学号 22351187

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 2023 年 1月 9日

**1.引言**

三维可变形面部模型（3DMM）是一种高效的统计模型，用于呈现多样真实面部。该模型建立在从一系列样本中学习面部形状和纹理的先验分布技术上，这些样本具有密集的对应关系。

在3DMM中，最核心的挑战在于产生潜在的可变形表示方式。过去二十年来，随着数据规模、多样性和质量的提升，这一领域取得了显著进展。最初基于线性模型，随后扩展到基于多线性模型，以单独编码不同模式。然而，由于线性模型在处理复杂变化（如夸张表情）时表示能力相对有限，这些方法的效果并不理想。

在深度学习背景下，通过使用卷积神经网络或图神经网络，在2D图像或3D网格输入下研究了许多非线性模型。这些方法带来了一定的绩效提升，但由于离散表示策略在输入数据上的限制，面部先验无法充分捕捉，导致形状细节的丢失。

本文提出了一种新的3DMM，名为ImFace，通过学习内隐神经表征（INR）对传统3DMM进行了重大改进。ImFace构建了独立的INR子网络，以明确将形状变形分解为身份和表情两个变形场，并引入了改进的嵌入学习策略，以扩展表情的潜在空间，允许更多样化的细节。这种方式可以精确建模个体差异和细粒度变形，同时考虑了相关任务的灵活性。

此外，受线性混合蒙皮的启发，本文提出了神经混合场，用于将整个面部变形或几何体分解为语义有意义的局部隐式函数编码的区域，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。

另外，本文设计了一种新的预处理管道，绕过了对不透水面部数据的需求，适用于各种面部表面，包括硬件采集或人工合成。

**2.实现方法介绍**

ImFace利用内隐神经表征（INR）学习非线性的3D可变形人脸模型。它明确定义面部形状变形，将其分解为身份和表情两个独立的变形场，并采用深度Signed Distance Function（SDF，有符号距离函数）来表达模板形状。所有这些属性都通过一系列局部隐式函数的混合来表示，以获得更为详细的表征。

**2.1分离的INRs网络**

INRs的核心思想是训练神经网络以逼近连续函数，该函数通过水平集隐式表示曲面。这种函数可以以多种形式定义，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度SDF来全面呈现面部特征。它输出查询点的带符号距离s：

其中，表示查询点在3D空间中的坐标，和分别表示表情和身份嵌入。 我们的目标是学习一个神经网络的参数化，使其符合真实的面部形状先验。ImFace网络由三个Mini Nets块组成，分别是Expression Mini-Nets (ExpNet)、Identity Mini-Nets（IDNet）、Template Mini-Nets (TempNet)。这一设计清晰地展示了面部形状变形的学习过程，保证能够准确地建模个体间差异和微小变化。前两个网络分别学习与表情和身份变化相关联的独立变形场，而模板Mini-Nets则学习模板面部形状的带符号距离场。详见图1（a）。

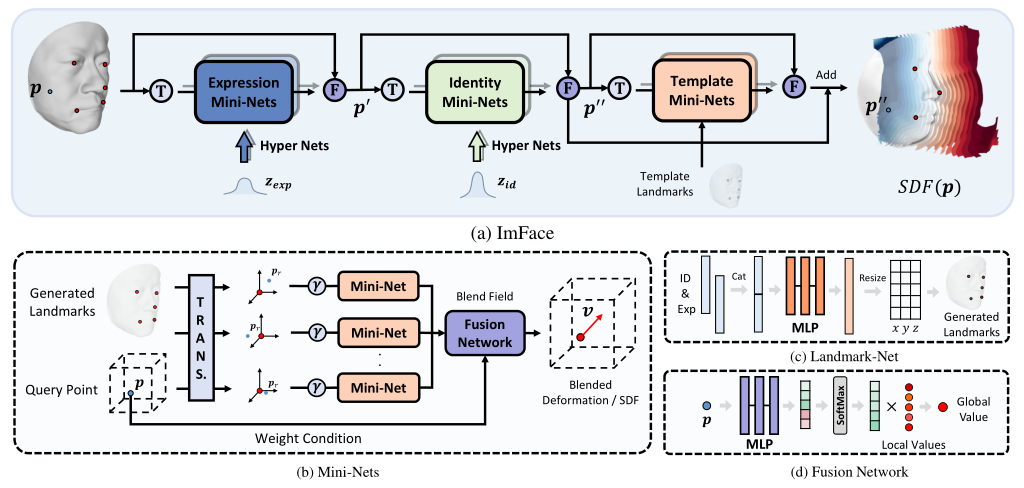


图1. ImFace概述

**2.2神经混合场**

Mini Nets块由三个子网络组成，共享一个通用的架构。它学习连续场函数：，用于生成全面的面部表示的神经融合域。特别是，为了克服单个网络的表达能力有限的问题，本文将人脸空间分解为一组具有语义意义的局部区域，并分别在混合之前学习每个区域的变形或有符号距离值。这种设计灵感源自最近INRs对人体的研究，该研究引入了线性混合蒙皮算法，使网络可以从身体部位的独立变换中学习。为了更好地表示面部表面的细节，我们替代了原始线性混合蒙皮算法中的常量变换项，并定义了神经混合场为：

其中，描述了第个局部区域的参数，表示第个混合权重，是对应的局部场。通过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，该方法提供了更强大的表示能力，特别在处理复杂的局部特征时更为有效。

**2.3改进的表达式嵌入学习**

目前，INR普遍使用一种自动解码器框架，以同时学习嵌入和网络参数。在之前的尝试中，每个表情类型都使用一个嵌入来编码，以便区分属性。然而，这种嵌入只能表示每种表情类型的平均形状变化，导致学习到的潜在空间无法捕捉个体之间更丰富的变化细节。为了克服这个问题，我们改进了学习方法，将每个非中性人脸扫描视为独特的表情，并为其生成特定的嵌入。这一改进大幅扩展了潜在空间，使其能够表示更为精细的细节。但是，我们也面临着一种潜在的问题，即身份属性再次陷入表情空间，导致身份映射的模糊性。为了解决这个挑战，当当前训练样本是中性面时，我们抑制了一些因素，即：

这意味着我们在学习过程中专注于共同学习中性面的形状表示，并且主要关注表情变化。此外，在训练过程中，我们只需要中性标签，避免了对密集表情标签的过度依赖。

**2.4数据预处理**

当前，对隐式函数的研究通常需要对输入进行严格处理，因为神经网络擅长拟合处处可微的函数。尽管像UDF这样的功能不需要严格的水密性，但它们在穿过表面时是不可区分的，并且在处理细节时效果欠佳。

为了解决这个问题，本文提出了一种有效的预处理管道，用于生成伪水密面，并在其上定义通用的SDF，从而使得可以像在水密对象上一样精确地学习几何图形和对应关系。

伪水密面的生成过程如下：首先，使用标记将面部严格对齐到正面，并将每个网格标准化为10cm单位。然后，将坐标原点设置在鼻尖后4cm的点，并将半径为10cm的球体定义为裁剪掉外部网格三角形的采样区域。接着，使用射线三角相交算法去除隐藏表面，例如鼻腔和口腔，并在定向和伪水密网格的x-y坐标上执行Delaunay三角化算法。

随后，针对面部曲面，进行SDF计算。首先生成伪水密面，然后通过对其进行距离变换来计算SDF值。样本的符号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。最后，对面部表面上的250,000个点和球体中的15,000个点进行统一采样，并计算它们的带符号距离和梯度向量。这些采样数据最终被形成为ImFace训练所需的三元组（查询点、梯度向量、有符号距离值）。

**3.本文创新点以及贡献**  
 本文介绍了一种基于INR的全新3DMM，通过两个独立分离的变形场来编码复杂的面部形状变化，以一种细致而有意义的语义方式学习强大的表示。同时，提出了一种有效的预处理管道，为非水密的3D面定义通用SDF，使得首次可以使用INR进行相关处理。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，并且在3D人脸重建方面胜过目前同行业先进技术。