# 利用三维感知合成数据估计姿态对人脸识别的影响，并将其应用于质量评估

Marcel Grimmer、Christian Rathgeb、Christoph Busch

**摘要**——评估人脸图像的质量对于确保人脸识别系统具备足够的准确性至关重要。人脸质量标准化领域的最新进展（*ISO/IEC CD3 29794-5*）建议采用组件质量度量，将人脸质量分解为各个独立因素，从而为操作人员提供有价值的反馈，以便重新采集低质量图像。鉴于三维感知生成对抗网络的最新进展，我们提出了一个新颖的数据集 *Syn-YawPitch*，包含1，000个身份，涵盖不同的偏航角与俯仰角组合。利用该数据集，我们证明，俯仰角超过 30 度会显著影响当前人脸识别系统的生物特征识别性能。此外，我们提出了一种轻量且可解释的姿态质量预测器，该预测器符合 ISO/IEC CD3 29794-5 国际标准草案的要求，并将其与最先进的人脸图像质量评估算法进行基准比较。

**索引词**—人脸识别、误差分析、合成数据、3D感知人脸图像合成

arXiv:2303.00491v2 [cs.CV] 6 Dec 2023

## 1 引言

ACROSS 在全球范围内，人脸识别（FR）系统在日常生活中日益普及。在欧盟，生物特征识别被用于敏感应用，例如出入境系统（*EES*） [1] ，该系统计划建成一个用于登记第三国国民的自动化信息技术系统。此外，在法证调查中，成员国之间也会处理和交换人脸数据，以识别嫌疑犯或寻找失踪人员——这项工作通过申根信息系统（*SIS*） [2]实施。尽管对安全、便捷的人脸识别系统的需求十分广泛，但其性能仍深受待比对图像质量的影响——这是一个众所周知的挑战，长期以来一直受到研究人员、标准制定者和从业者的广泛研究。

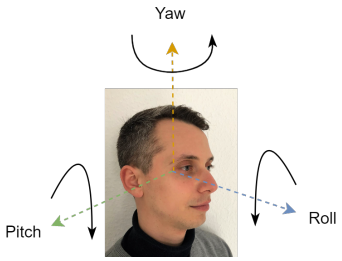


图 1：根据 ISO/IEC 39794-5 [10] 的规定，每种面部姿态均由偏航角、俯仰角和横滚角表示。

### 1.1 人脸图像质量

为避免在配合式图像采集场景中产生不准确的结果，质量评估算法会在样本交由识别系统处理之前，先剔除低质量样本。在此背景下，*ISO/IECCD3 29794-5* [3] 将以 $[0, 100]$，范围内的整数值表示的统一质量度量与组件质量度量区分开来，其中数值100 表示对于接收该样本的人脸识别系统而言具有完美效用（完美效用 [4]）。更准确地说，统一质量分数旨在预测给定样本会对识别准确率产生负面还是正面贡献，从而汇总给定人脸图像中的所有变化因素。虽然统一质量分数适用于减少 [5]生物特征识别系统的工作负载并提升其性能 [6] ，但其聚合特性使得无法确定导致质量低下的因素。此外，目前用于估计统一质量分数的最先进的方法 [7] [8] [9] 依赖于深度神经网络，而众所周知，这些网络会以牺牲可解释性为代价，换取复杂模型架构所带来的准确率提升。

相比之下，组件质量度量通过将各个变化因素的质量分解为单个分量标量来提高可解释性。因此，人脸识别系统的操作人员可以提供可采取行动的反馈，要求受试者纠正其行为，例如调整头部正面化或保持表情中性。在实践中，开发分量质量评估算法 的主要挑战之一，是将所有面部属性分离开来，以便逐一进行分析。为克服这一难题，我们在本研究中利用合成生成人脸图像的精细可控性，构建了一种轻量级姿态质量评估器。

• *M. Grimmer* 就职于挪威生物识别实验室（*NBL*）、挪威科技大学（*NTNU*），位于挪威约维克。电子邮箱：*marceg@ntnu.no*

• *C.Rathgeb* 和 *C. Busch* 任职于达姆施塔特应用科学大学（*h-da*）达姆施塔特校区的生物识别与互联网安全研究组（*da/sec*）。德国。电子邮箱：*christian.rathgeb@h-da.de,christoph.busch@h-da.de*



图 2：EG3D [16], 生成的3D感知人脸图像及其对应的3D网格，展示了StyleGAN2 [12] 的高视觉质量以及一致的人脸几何结构。

### 1.2 三维感知人脸图像生成

传统上，利用*2D*感知生成对抗网络 （生成对抗网络） [11] [12] [13]生成合成人脸图像，可以实现极高的照片级真实感，以至于人类感知 [14] [15]无法将其与真实数据区分开来。生成合成人脸图像后，必须编辑其面部属性，以分析此类变化如何影响相似度分数。

尽管二维感知生成对抗网络能够生成高度逼真的图像，但在操纵偏航角和俯仰角等属性时，它们通常难以保留面部几何结构（见图 1），从而导致关键身份信息的丢失。因此，一个新的研究方向是开发*3D*感知生成对抗网络，以便在保留身份信息的同时精确编辑面部属性，尤其是保持面部几何结构的一致性 [16]。与之前的工作相比 [17] [18] [19],, EG3D通过采用显式-隐式网络架构（见第 2 节），在几何一致性与内存利用率之间实现了权衡。在本工作中，如图 2 所示，我们利用 EG3D 的三维感知能力，基于合成数据研究偏航角和俯仰角对不同人脸识别系统的影响。具体而言，所有人脸图像均由在姿态均衡版本的 FFHQ 上预训练的 EG3D生成器生成 [11], [16] 数据集，分辨率为 $512^2$。

### 1.3 贡献与论文结构

我们首先在第3节介绍合成数据集的生成过程，并通过现成的运动恢复结构 [20] 和多视图立体视觉 [21] 流程进行三维一致性检查。进行三维一致性检查是必要的，因为使用生成对抗网络编辑俯仰角和偏航角时，通常会改变其他面部特征，这是由这些特征在潜在空间中的纠缠所致 [22]。接下来，第4节介绍我们提出的*SYP-Lasso* 姿态质量估计器，以及当前对*ISO/IEC CD329794-5* [3]所定义的姿态分量质量的实现。在第5节，我们分析了不同偏航角-俯仰角组合对五种人脸识别系统生物特征识别性能的影响：ArcFace [23], MagFace [7],Curricular-Face [24], AdaFace [25], 和 Cognitec [1] [2]。最后，第6节展示了我们的 SYP-Lasso 回归模型在*BIWI Kinect*人脸数据集 [26]上的有效性；在姿态依赖型质量预测任务中，该模型优于最先进的分量质量评估算法和统一质量评估算法。

总体而言，本工作的贡献可总结如下：

- 提出三维感知合成数据集（*Syn-YawPitch*），包含 1,000 个身份，共计 144,000 张图像，涵盖各种偏航角与俯仰角组合。
- 基于最先进的开源人脸识别系统和商用现成的人脸识别系统，估计不同偏航角与俯仰角组合的影响。
- 提出一种轻量级、符合 ISO/IEC 标准的 [3]姿态质量估计算法 [27]（*SYP-Lasso*），可生成可解释的质量度量。

## 2 相关工作

本节首先总结二维感知人脸图像合成领域的最新进展，随后介绍聚焦于向3D感知人脸图像生成过渡的相关研究。最后，第2.3节将说明过去如何使用合成数据来评估人脸识别系统。

### 2.1 二维感知人脸图像合成

在生物识别应用的背景下分析合成数据，需要图像合成模型生成质量与真实数据难以区分的人脸图像。也就是说，Karras等人 [28] 是最早提出能够生成高质量图像（1024x1024）的GAN架构的研究者之一；该架构通过逐步提高生成器和判别器处理的图像分辨率来实现这一目标。借助 StyleGAN [11], , 研究者引入了基于风格的GAN架构，以避免传统生成对抗网络潜在空间中观察到的特征纠缠。中间潜在空间这一新概念，以及混合多个身份风格的可能性，为现代无条件GAN提供了更多控制其输出的手段。

尽管 StyleGAN 在无条件图像合成和潜在空间解耦方面不断取得新突破，但由于集成自适应实例归一化 [29]操作以及网络架构设计存在重大缺陷 [12]，其应用仍受到明显图像伪影的限制。随着 StyleGAN2 [12]和 StyleGAN3 [13], 的后续改进，StyleGAN 常见的液滴状伪影和纹理粘连效应得以消除，生成的人脸图像达到人眼难以与真实图像区分的程度 [15] [14]。

[1] https://www.cognitec.com/ (FaceVACS，版本 9.6.0)[2] **免责声明：**请注意，本研究中使用 CognitecSDK 进行的评估完全由作者执行。所呈现的结果未必能全面反映该技术的全部能力。

### 2.2 3D感知人脸图像合成

尽管传统生成对抗网络能够生成极为逼真的人脸图像，但在学习操控面部属性（例如姿态或表情）所需的基础三维结构方面仍然存在困难 [17]。为克服这一障碍，另一条发展路线是采用几何感知生成对抗网络 [17] [18] [19] [16]，，从二维训练数据中学习三维表示。通常，三维物体表示既可以通过显式方式学习，例如使用深层体素网格 [30]，；也可以通过潜在表示 [31]隐式学习。通过三维卷积显式处理体素网格的计算开销极大，因此限制了其在大规模高分辨率图像中的实际应用。相比之下，隐式架构虽然更加紧凑，但由于需要大型全连接神经网络来存储所有信息，运行速度往往过慢，通常只能用于单图像优化场景。

为提高计算效率并减少内存占用，Chan等人 [16] 提出了一种显式-隐式方法，结合了两者的优势。他们基于StyleGAN2 [12]生成器骨干网络，通过神经辐射场显式地表示空间信息 [31]。随后，这些特征被传递至三平面解码器，由其隐式学习底层人脸的更紧凑表示。最后，可微神经体渲染器将给定特征转换为二维人脸图像。为将头部姿态角与其他面部属性解耦，StyleGAN2映射网络以相机参数为条件 [32]，并在训练期间以50%的概率随机交换这些参数。在推理过程中生成人脸图像时，这种条件机制已被证明能够显著减少头部姿态与训练数据集中观察到的其他变化因素之间的纠缠，例如正面肖像中的微笑表情。在本研究中，我们利用3D感知图像合成领域的这些进展，生成了一个大规模数据库（“Syn-YawPitch”），其中包含一致的偏航角和俯仰角变化，以便按照*ISO/IEC CD3 29794-5* [3]的要求推导姿态质量估计器。

### 2.3 生物识别中的合成数据

计算机视觉和计算机图形学近期在生成逼真且可控的人脸图像方面取得了进展，由此推动了各种有益于生物识别系统的应用。正如 Joshi等人 [33]，所示，合成数据集广泛用于人体分析领域的各种应用，包括评估生物识别性能、评估生物识别系统的可扩展性，以及通过扩大真实数据集的规模和多样性来减少偏差。使用合成数据替代或扩充真实数据，依赖于合成样本所具备的高度逼真性和可控性；我们预计这一趋势未来将持续发展。在使用合成数据进行人脸识别的研究中，Kortylewski等人 [34]的开创性工作之一使用三维形变模型人脸模型，通过在留出姿态区间的合成数据集上进行训练，分析人脸识别系统在推理过程中的泛化能力。

Colbois等人 [35] 创建了包含与 Multi-PIE 数据集 [36]相似变化度的合成人脸图像，以比较六种人脸识别系统的生物特征识别性能与原始 Multi-PIE 数据集上报告的性能。通过观察到性能差异很小，作 [37]者得出结论：其合成数据集可替代真实数据，用于评估人脸识别系统的性能，同时满足隐私保护要求。Zhang等人 [37]也得出了类似结论，他们使用三种质量评估算法，定量分析了合成人脸图像与真实人脸图像之间的差异。在后续研究中，Grimmer等人 [38] 使用两种人脸年龄增长技术 [39] [22] 生成了年龄差异最多达60 年的匹配样本，以研究人脸身份随时间变化的程度。

此前，Grimmer等人[40] 通过对 StyleGAN 的潜在空间应用主成分分析，探索新的方向，从而生成非确定性的匹配样本。以非确定性方式创建匹配样本，旨在检测人脸识别系统中的漏洞；如果只关注简单的面部属性（例如微笑），这些漏洞可能会被忽略。

## 3 SYN-YAWPITCH

姿态质量估计器（PQE）的推导需要分别分析偏航角和俯仰角的组合，以评估它们对人脸识别准确率的影响。因此，必须精确控制每个协变量，同时保持其他人脸特征和环境因素固定。在这项工作中，我们借助 Chan 等人 [16](*EG3D*)的工作，利用 3D 感知图像生成领域的最新进展。具体而言，我们生成了一个数据库，其中包含偏航角和俯仰角组合各不相同的合成匹配样本（*Syn-YawPitch*）。

尽管滚转角会影响生物特征识别性能，但应用人脸对齐算法可以减少这类变化。因此，所有 EG3D 训练样本都经过了预处理，使生成器无法在推理过程中操纵滚转角。不过，我们提请注意 Lu 等人 [41]，的研究发现，即滚转角处于 $\pm 30°$范围内时，生物识别性能稳定。我们还将在第 6.1 节证明，评估数据库中的大多数人脸图像的滚转角均处于 $\pm 20°$ 范围内。

我们生成了1，000 个身份（身份标识）的人脸图像，以评估不同偏航角-俯仰角组合下的生物特征识别性能。具体而言，我们通过笛卡尔积 $\Phi_y = \{90°, 80°, 60°, 40°, 20°, 0°\}$ 和$\Phi_p = \{90°, 80°, 60°, 40°, 20°, 0°\}$定义的偏航角（$\Phi_y$）与俯仰角（$\Phi_p$）构建网格状配对。按照这一方法，Syn-YawPitch 数据集中的图像总数增至144，000，每个身份包含144 幅图像。表1详细比较了 Syn-YawPitch 与现有包含头部姿态变化的数据集。

### 3.1 头部姿态重新标注

尽管 EG3D 能够在保持人脸几何结构一致的情况下编辑偏航角和俯仰角，但我们发现指定角度与实际角度之间存在差异。指定角度偏离正面视角越多，这种影响就越明显，并且很可能源于 EG3D 训练数据集中的头部姿态失衡。然而，准确估计实际头部姿态角是一

表1：具有头部姿态变化的公开可用数据集概览。值得注意的是，在受约束数据集中，Syn-YawPitch 的每个受试者的图像数量最多，并且包含类似网格的偏航角-俯仰角组合分布，以便于系统评估协议。

| 数据集 | 受试者 | 图像 | 合成? | 受约束? | 姿态变化 |
|---|---|---|---|---|---|
| Syn-YawPitch（我们的方法） | 1,000 | 144,000 | Yes | Yes | 偏航、俯仰 |
| CMU Multi-PIE [36] | 337 | ~750,000 | No | Yes | Yaw |
| SynHead [42] | 10 | ~510,000 | Yes | Yes | 偏航、俯仰、横滚 |
| SASE [43] | 50 | 30,000 | No | Yes | 偏航、俯仰、横滚 |
| 指向'04 [44] | 15 | 2,940 | No | Yes | 偏航、俯仰 |
| AFLW-2000 [45] | 2,000 | 2,000 | No | No | 偏航、俯仰、横滚 |
| BIWI [26] | 20 | ~15,000 | No | 部分 | 偏航、俯仰、横滚 |
| CAS-PEAL-R1 [46] | 1,040 | ~30,000 | No | Yes | 偏航、俯仰 |
| FEI [47] | 200 | 2,800 | No | Yes | 偏航、俯仰 |
| M2FPA [48] | 229 | ~400,000 若 | | Yes | 偏航、俯仰 |
| CPLFW [49] | 5,749 | 13,233 | No | No | 野外环境 |

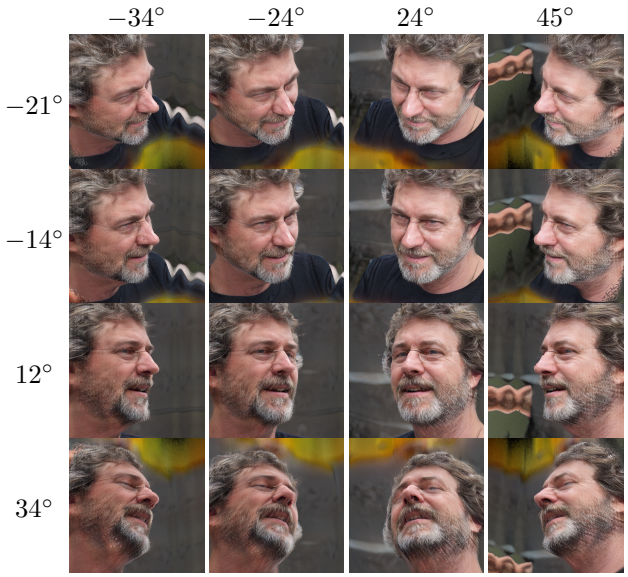


图 3：Syn-YawPitch 样本示例，展示了多个偏航角与俯仰角组合下的视觉质量。这些示例突出了 EG3D [16] 仅改变摄像机角度的能力，同时不会引入正面视图中观察到的偏差，例如微笑的面部表情。

这是评估其对人脸识别效用影响的关键前提。因此，我们使用现成的姿态估计器 [32]为每个样本重新标注。具体而言，我们定义了一个函数，将原始 EG3D 姿态标签 $\Phi_y$ 和 $\Phi_p$ 映射到对 100 张随机生成图像应用姿态估计器后所得的中位数 $\{x_1, x_2, \ldots, x_{100}\}$:

$$\hat{\Phi}_y = \text{Median}(\mathcal{P}_{\Phi_y}(x_1), \mathcal{P}_{\Phi_y}(x_2), \ldots, \mathcal{P}_{\Phi_y}(x_{100})) \quad (1)$$

and

$$\hat{\Phi}_p = \text{Median}(\mathcal{P}_{\Phi_p}(x_1), \mathcal{P}_{\Phi_p}(x_2), \ldots, \mathcal{P}_{\Phi_p}(x_{100})) \quad (2)$$

ˆ ˆ其中 $\Phi_y$,$\Phi_p$ 表示偏航角和俯仰角的中位数值，$\mathcal{P}_{\Phi_y}(x_i)$ $\mathcal{P}_{\Phi_p}(x_i)$ 表示第 i 张随机生成的人脸图像 $x_i$估计出的偏航角和俯仰角。计算出所有中位数值后，所得查找表便可根据外部头部姿态估计器 [32]，将 EG3D 的姿态标签转换为更准确的标签。重新标注样本的图像示例见图 3。请注意，通过重新标注我们的 Syn-YawPitch 数据库，我们发现 EG3D 的生成能力存在不平衡——

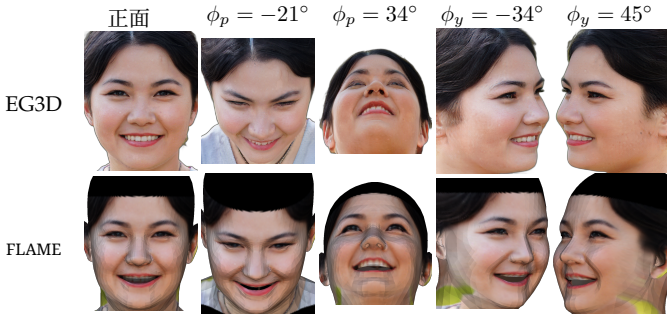


图 4：使用三维感知生成对抗网络（EG3D [16]）和三维可变形人脸模型（FLAME [50]）生成的人脸图像的视觉对比。

例如，指定 $\Phi_y = 90°$ 对应于调整后的数值 ˆ $\Phi_y = 45°$，而将符号翻转为 $\Phi_y = -90°$ ˆ则对应于 $\Phi_y = -34°$。我们认为，这些姿态不平衡是在 EG3D 训练阶段形成的，原因是训练数据集（*FFHQ* [11]）中某些姿态区间的样本较少。尽管 EG3D 能够生成超出训练阶段所见范围的姿态角，但由于伪影数量会明显增加，我们不建议选择高于 90° 的角度。

在图 4 中，我们展示了一个示例，说明 EG3D 和最先进的三维可变形人脸模型 (3DMM) [50], [51] 从不同姿态角度生成图像的视觉质量。虽然 3DMM 的参数化能够精确控制各种面部属性，但渲染出的人脸图像缺乏照片级真实感，也无法呈现面部毛发、牙齿或眼球等重要细节。

### 3.2 实现细节

由于 EG3D 的架构采用了 StyleGAN2 [12], 的生成器，我们遵循 Colbois 等人 [35] 的建议，限制每个潜代码与此前抽取的样本保持最小距离。如[35], 所示，这种采样策略可避免潜向量从潜在空间中彼此接近的区域抽取而导致相似者比率异常偏高。

生成 Syn-YawPitch 时，我们选择0.5 的截断因子，以生成图像质量稳定的人脸图像。截断因子可视为一种参数，它通过限制随机抽取潜向量的潜在空间区域，来控制感知图像质量与多样性之间的权衡 [37] 。我们还利用合成数据的高度灵活性，构建了一个独立数据集，并使用0.75的截断因子计算具有代表性的异人比对分数。这一选择受到 Zhang 等人 [37], 的启发，他们已证明，截断因子高于0.75 时，所得异人比对分数分布与真实数据集中的分布相似。

### 3.3 几何一致性

为确保 EG3D [16]在调整观看角度时保留面部几何结构，我们采用一种独立的运动恢复结构（SfM） [20]技术

该技术用于从二维图像序列中重建三维结构。因此，给定来自多视角的一组二维人脸图像，SfM 旨在估计其底层三维结构。本节将从面部几何结构的角度，考察基于三维感知生成对抗网络（EG3D [16]）和二维感知生成对抗网络（StyleGAN [11]）的人脸图像示例的身份保持效果。

我们发现，尽管 StyleGAN 能够生成高质量图像，但其生成器并未针对在改变相机视角时保留面部几何结构进行优化。图 5 展示了这一问题：其中给出了从同一合成身份在不同观看角度下生成的 [20] 320 张人脸图像重建得到的三维点云。给定一个初始潜代码，InterFaceGAN[22]分别通过在潜在空间中以极小步长操控偏航角和俯仰角，生成经过操控的样本。点云表明，正如图 5 所示，StyleGAN 默认并不适合在编辑偏航角或俯仰角时保留面部几何结构。

相比之下，EG3D 的三维感知网络架构包含一个以图像的相机外参和内参为条件的神经体渲染模块。在训练过程中采用这一设置后，EG3D 可用于推理，并能一致地控制偏航角和俯仰角。图 5 展示了示例图像，呈现了同一身份在多个观看角度下模拟生成的人脸图像。与 InterFaceGAN 相比，EG3D 对几何人脸结构的保留效果明显更好，且使用了数量相同的、具有不同偏航角和俯仰角的 320 张输入图像。在满足面部几何结构一致这一前提后，我们为下一节展示的实验结果奠定了坚实基础。

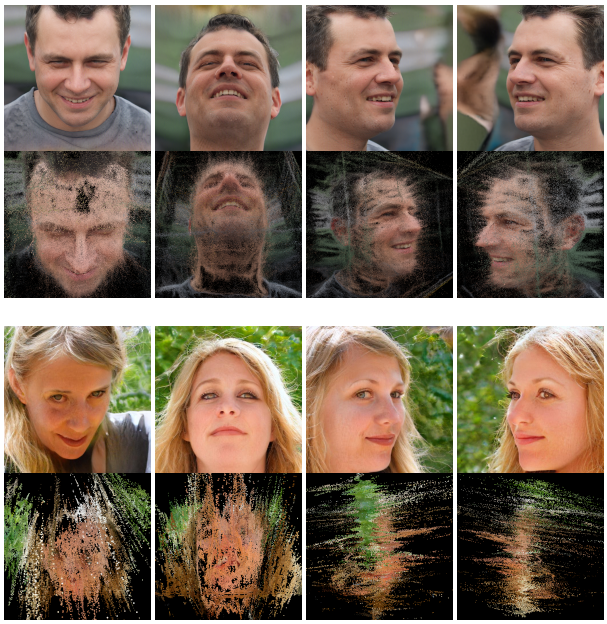

图 5：使用 EG3D 生成的人脸图像示例 [16]（上）和 InterFaceGAN 生成的人脸图像示例 [22]（下）。3D 点云由开源运动恢复结构 [20]软件工具生成。[3]。给定同一受试者在多个视角下的一组 EG3D 人脸图像，生成的稠密点云重建呈现出连贯的面部结构。相比之下，具备 2D 感知能力的 InterFaceGAN 方法 [22] 无法生成几何一致的图像序列，这一点可由 3D 点云中不存在的图案看出。

表 2：使用 5 个不同的人脸识别系统，在 Syn-YawPi 上训练的 SYP-套索回归 PQE 系数概览。每个系数都可以解释为相应姿态角对与正面人脸图像进行比较时预期相似度的贡献。请注意，在此表中，我们使用多项式次数 $n = 1$，以便于解释各个系数。

| 人脸识别系统 | $\phi_{p[-]}$ | $\phi_{p[+]}$ | $\phi_{y[-]}$ | $\phi_{y[+]}$ |
|---|---|---|---|---|
| ArcFace [23] | 1。12e−02 | −1.13e−02 | 8。73e−03 | −6.27e−03 |
| MagFace [7] | 1.14e−02 | −1.17e−02 | 8.76e−03 | −6.17e−03 |
| CurricularFace [24] | 1。18e−02 | −1.23e−02 | 9.20e−03 | −6.68e−03 |
| AdaFace [25] | 1。23e−02 | −1.26e−02 | 1.02e−02 | −7.06e−03 |
| Cognitec[1] | 8.83e−03 | −1.07e−02 | 6.67e−03 | −4.70e−03 |

## 4 姿态质量估计器

本节介绍基于我们的 Syn-YawPitch 数据集构建的姿态质量估计器（PQE），旨在根据人脸图像的偏航角和俯仰角预测其人脸识别效用。此外，我们还介绍了 ISO/IEC CD3 29794-5 [3] 的当前实现，并在实验中将其用作基线姿态质量估计器。请注意，我们提出的姿态质量估计器，即 *SYP-Lasso*，以及基于 ISO/IEC 的姿态质量估计器，均按照标准化要求实现，用于量化给定人脸图像的效用；其质量度量被限制为 $[0, 100]$，范围内的整数标量，其中 0 表示效用下限，100 表示效用上限。

### 4.1 基线姿态质量估计器

为建立公平的基准，我们采用 ISO/IEC CD329794-5 [3] 定义的组件质量度量，分别根据偏航角和俯仰角估计效用：

$$\text{ISO-PQE}(\Phi_y) = 100 \cdot cos^2\Phi_y \tag{3}$$

and

$$\text{ISO-PQE}(\Phi_p) = 100 \cdot cos^2\Phi_p \tag{4}$$

[3] https://colmap.github.io/

与本研究不同，ISO/IEC CD3 29794-5 [3] 中定义的效用估计分别针对偏航角和俯仰角计算，因此忽略了这两个因素之间潜在的相互相关性。为提高与我们提出的 SYP-Lasso 模型的可比性，我们选取 $\Phi_y$ 和 $\Phi_p$ 中的较小值，以得出融合的姿态分量质量度量。

### 4.2 SYP-Lasso PQE

给定正面人脸图像与同一身份但偏航角和俯仰角经过调-整的图像进行比对后得到的相似度分数，我们可以通过数学模型捕捉头部姿态与相似度分数之间的关系。因此，为推导出我们提出的 SYP-Lasso PQE，我们作出以下假设：1）当将正面参考图像与头部旋转的探测

样本进行比对时，我们假设相似度分数较低是由偏航角和俯仰角的变化直接导致的。这一说法基于以下事实：只有摄像机视角发生了变化，从而导致偏航角和俯仰角变化，而其他面部属性保持不变。2）由于相似度分数仅取决于单一变化因素，我们进一步假设，相似度可以被视为一种效用度量，并符合 ISO/IEC CD3 29794-5 [3] 中规定的要求。

给定我们提出的 Syn-YawPitch 数据集（其偏航角-俯仰角组合呈网格状分布），我们比较正面人脸图像与其在不同偏航角-俯仰角配置下的匹配对应样本，并使用套索回归模型 [27] 对相似度分数（即姿态质量度量）与头部姿态角之间的关系进行建模，该模型将以下项最小化：

$$\min_{\beta \in \mathbb{R}^T} \left\{ \frac{1}{N} ||s - \mathcal{P}\beta||_2^2 + \lambda ||\beta||_1 \right\} \tag{5}$$

在公式 5 中，$s = [s_1, s_2, \ldots, s_N]$ 表示姿态旋转探测样本与正面参考图像进行匹配比较所得的相似度分数，其中 $N = 144,000$ 表示 Syn-YawPitch 数据集的大小。此外，$\beta = [\beta_1, \beta_2, \ldots, \beta_T]$ 表示通过最小化公式 5[4] 并采用经验选定的正则化项 $\lambda = 1e{-}06$ 计算得到的套索回归系数，以减少过拟合，并通过自动移除预测价值较低的协变量来进行特征选择。最后，$\mathcal{P} \in \mathbb{R}^{N \times T}$ 中的每一行都包含如下所述的 $T = |\mathcal{C}|$回归协变量：

$$\begin{aligned} \mathcal{C} = & \bigcup_{i=1}^{n} \left\{ \phi_{p[+]}^i, \phi_{p[-]}^i, \phi_{y[+]}^i, \phi_{y[-]}^i \right\} \\ & \bigcup \left\{ a \cdot b \mid (a,b) \in \Phi_p^{i-1} \times \Phi_y^{i-1} \right\} \end{aligned} \tag{6}$$

多项式次数为 $n$，俯仰角 $\Phi_p = \{\phi_{p[+]}, \phi_{p[-]}\}$ 和偏航角 $\Phi_y = \{\phi_{y[+]}, \phi_{y[-]}\}$ 则分别拆分为正值和负值，以便 SYP-Lasso 学习与姿态相关的性能不平衡，这是推导姿态效用度量的关键。

如图6所示，根据经调整的 $R^2$ 指标，超过 $n = 2$ 后继续提高多项式次数只能带来微小的性能提升。该指标表示协变量能够解释目标变量（即相似度分数）方差的百分比。鉴于选择 $n = 1$时系数具有可解释性，表 2 首先展示了各协变量的原始影响，随后将在第5节中进行更全面的讨论。最终，我们提出的 SYP-Lasso 模型采用 $n = 2$，在模型复杂度与效用估计性能之间实现了最佳权衡；如图6所示，该模型在人脸识别开源系统和 FaceVACS（Cognitec）上的经调整 $R^2$ 分别最高达到 83% 和 61% 。

最终，在推理过程中，SYP-Lasso 将流程简化为简单的线性回归。此外，我们将估计的姿态效用缩放并裁剪到 $[0, 100]$ 的范围，以满足*ISO/IEC CD3 29794-5*的要求。

[4] https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso.html

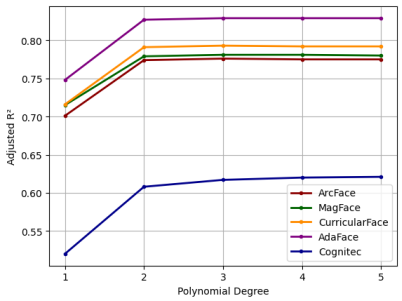


图 6：曲线表明，对于多项式次数为 $n > 2$的情况，调整后的$R^2$ 仅有小幅提升。请注意，在第 6.1 节中，我们选择多项式次数为 $n = 2$，因为在姿态质量估计中，它在模型复杂度与性能之间实现了最佳权衡。

## 5 姿态影响估计

Syn-YawPitch 的生成使我们能够分析不同偏航角与俯仰角组合对生物特征识别性能的影响。我们利用合成数据集的受控结构，计算等错误率（EER）、错误不匹配率（FNMR）和错误匹配率（FMR），以评估实验中纳入的人脸识别系统的性能。具体而言，我们报告固定为 1% 的 FMR 下的 FNMR（即 FNMR@1%FMR），以及固定为 1% 的 FNMR 下的 FMR（即 FMR@1%FNMR），以评估生物特征识别性能。

图 7 展示了三维性能曲面网格，其中网格上的每个点代表基于 1,000 个同人比对分数计算出的性能。这些分数是通过将 1,000 个身份标识的正面人脸与其匹配对应样本进行比对，并采用特定的偏航角-俯仰角组合获得的。该可视化有助于分析偏航与俯仰的相互作用，以及它对各种指标下识别性能的影响。如第 3.2 节所述，我们选择 0.75 作为截断因子，随机生成 1,000 对非同人样本，且每对样本均为正面视图。在所有实验中，所得的非同人比对分数分布保持不变，用于计算所有生物特征性能指标。

在 $[-20°, 20°]$的区间内，识别性能保持稳定。然而，当偏航角和俯仰角同时设定为较大值时，我们观察到所有错误指标和人脸识别系统整体上都有所上升。尤其是，超过 $30°$ 的正俯仰角似乎会导致性能大幅下降，错误不匹配率和错误匹配率均超过 80%。尽管我们最初预计，某些人脸识别系统可能擅长处理俯仰角变化，而另一些系统可能更擅长处理偏航角变化，但研究结果表明，所有人脸识别系统在处理偏航角变化时的表现始终更好。这一观察使我们认为，除非所有（通用型）人脸识别系统都天生偏向于处理偏航角变化，否则它们难以识别俯仰角同样较高的人脸图像，因为此时可见的人脸区域有限。

为进一步分析这种不均衡，图 8 展示了固定俯仰角或偏航角为 $0°$，同时遍历另一角度时得到的交线。作为

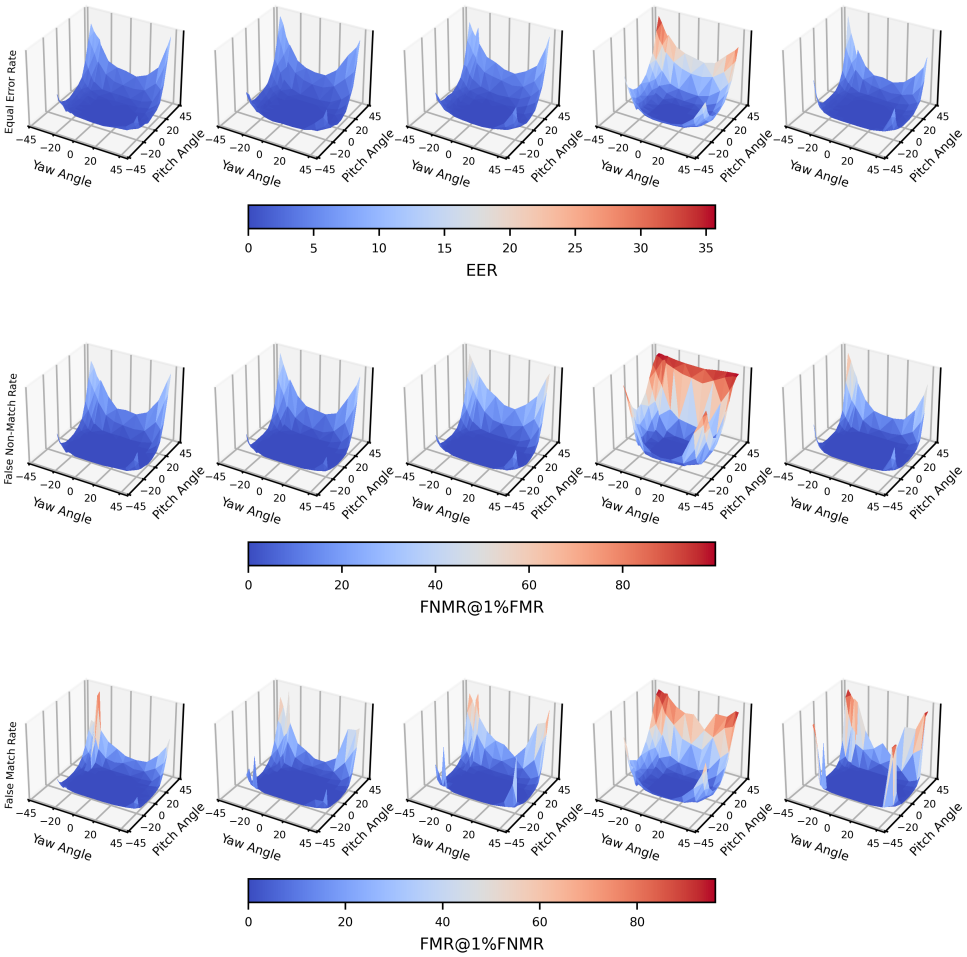


图 7：展示不同偏航角-俯仰角组合下生物特征识别性能的三维表面网格，每一列对应一个人脸识别系统：ArcFace（左）、MagFace（中左）、CurricularFace（中）、AdaFace（中右）、Cognitec（右）。

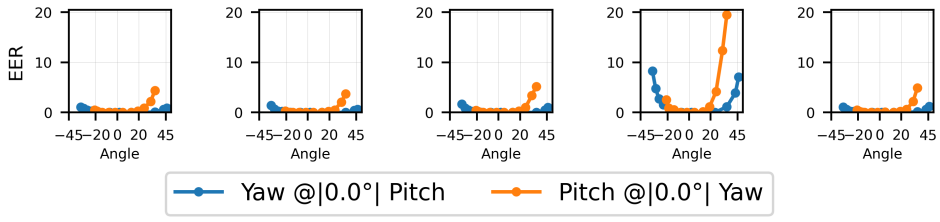


图 8：三维性能表面网格的交线，显示将俯仰角或偏航角固定在 $0°$ 时四种人脸识别系统的等错误率：ArcFace（左）、MagFace（中左）、CurricularFace（中）、AdaFace（中右）、Cognitec（右）。俯仰角超过 $30°$ 后，错误率显著上升。

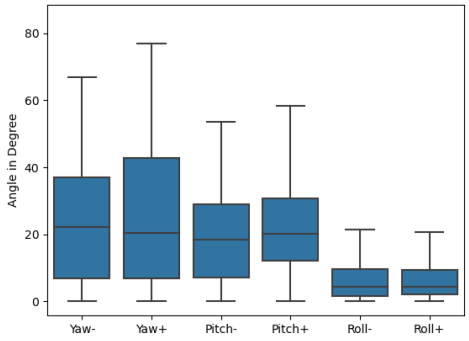


图 9：箱线图展示了 BIWI 数据库中偏航角、俯仰角和横滚角的分布，负角度以绝对值报告。

图 7 中的视图表明，超过 $30°$ 的正俯仰角对识别性能的影响最大，这一点从橙色曲线显著向外突出可以看出。尽管负俯仰角最多只能生成至 $-21°$，但我们预计这类变化度对性能的影响同样显著，这一点可从左侧尾部区域开始出现的轻微曲线偏转看出。正、负偏航角的影响已被证明相近，这符合我们的预期，因为通过简单的水平翻转即可消除偏向某一侧的影响。

表 2 显示了相同的不平衡现象，其中列出了在四个基础协变量 $\Phi_p$ 和 $\Phi_y$ 上训练 SYP-Lasso（见公式 5）所得的系数。该表表明，俯仰角变化的负面影响被过度加权，这一点从其系数的高绝对值可以看出。相比之下，正、负偏航角对相似度分数的影响要小得多。此外，高达 $83\%$ 的调整后 $R^2$ 值表明，在 Syn-YawPitch 数据集中，相似度分数与头部姿态旋转之间存在很强的线性关系。

根据本节的观察，人脸识别系统操作员应考虑俯仰角变化超出 $\pm30°$时对性能的敏感性。此外，我们想强调当前开源人脸识别系统能够稳健地处理 $\pm20°$范围内的偏航-俯仰交叉情况，这充分展示了深度神经网络 [52][53] 能够泛化至各种相机视角的能力。

## 6 姿态质量估计

如表 2 所示，头部姿态与相似度分数之间存在很强的线性关系，本节将分析我们的 SYP-Lasso 模型根据偏航角和俯仰角变化估计人脸识别效用的有效性。具体而言，我们针对每个经评估的人脸识别系统分别拟合一个 SYP-Lasso 回归模型，以学习相似度分数与偏航角和俯仰角组合之间的、特定于人脸识别系统的关系。我们将所提出的方法与以下方法进行基准比较：1）第 4.1 节介绍的与人脸识别无关的 ISO/IEC 相关姿态质量估计器；以及 2）四种成熟的统一质量评估算法：FaceQnet v1 [8], SER-FIQ [9], MagFace [7],和 CR-FIQA [54]。

我们基于*BIWI Kinect* 头部姿态（BIWI）数据库 [26] 开展实验，该数据库包含 $15k$ 幅来自24 名个体的图像，其偏航角范围为 $\pm75°$，俯仰角范围为 $\pm60°$。所有样本均使用三维传感器采集，并包含深度信息和精确的姿态标签。因此，可以排除由标签不准确引起的真实值误差传播。

图9展示了 BIWI 数据集中偏航角、俯仰角和横滚角的分布。由于大多数横滚角是在 $-20°$ 至 $20°$ 的区间内采集的，并且 EG3D 仅支持编辑偏航角和俯仰角，因此本文不对横滚角进行分析。这一决定基于先前研究 [41]的观察：在 $\pm30°$ 范围内，横滚角对人脸识别性能的影响较小。

我们再次强调，使用真实数据评估姿态质量估计性能至关重要，这样才能排除合成与真实域差异，避免 SYP-Lasso 学到有意义的模式。尽管必须纳入真实数据，但若要收集行为方式与合成样本同样受控的真实匹配样本，就难以避免成本过高。面部属性之间的纠缠可见于图10，其中展示了评估数据集中的部分样本。尽管所有样本的采集都确保了头部姿态旋转是主要变化因素，但不可控的微表情仍会对比对分数产生额外影响。



图 10：用于将我们提出的 SYP-Lasso PQE 与基线姿态质量估计器进行基准测试的部分图像示例。

### 6.1 SYP-Lasso基准测试

我们遵循*ISO/IEC FDIS 29794-1* [4] 和美国国家标准与技术研究院(NIST) [55]的规范，通过根据错误率-丢弃率特征曲线（见图 11）评估我们 SYP-Lasso 质量估计器的性能。每条曲线显示持续丢弃估计姿态质量最低的图像时的错误不匹配率。在此情境下，理想曲线对应于错误不匹配率随着低质量图像的丢弃比例增加而成比例下降。为便于比较各条曲线，报告了丢弃率最高为曲线下面积部分值（pAUC），丢弃率上限为$20\%$ [6] [56]。

从图 11 可以明显看出，在所有评估的人脸识别系统中，SYP-Lasso 的性能均显著优于基于 ISO/IEC 的基线方法，表现为部分曲线下面积更低。此外，SYP-Lasso 的结果可与最先进的统一质量评估算法相媲美，甚至在五个人脸识别系统中的三个系统上超过了这些算法的

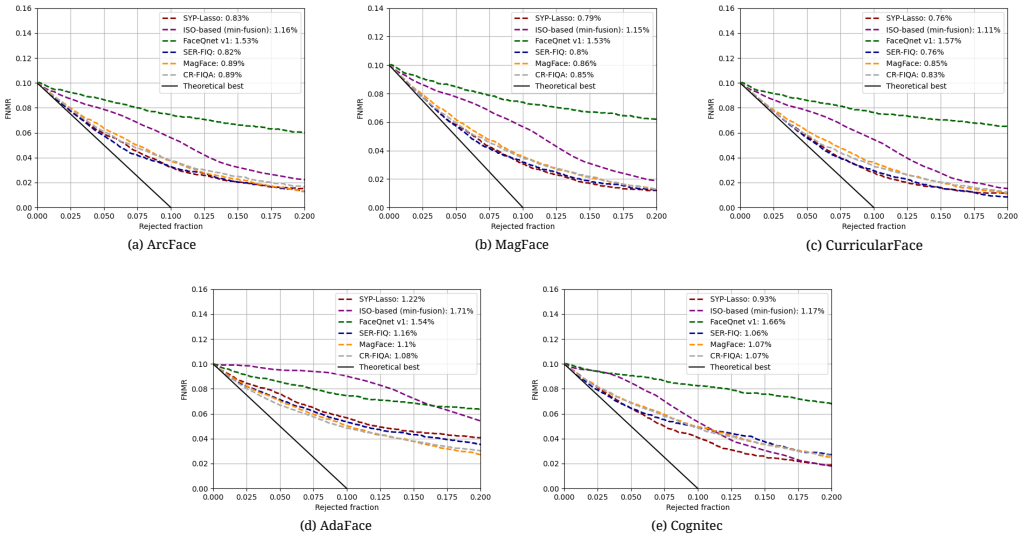


图 11：错误与弃置特征曲线表明，将基线和基于 ISO 的姿态质量估计替换为 SYP-Lasso 可显著提升性能。图例中报告的部分曲线下面积值可用于对曲线进行定量比较。我们着重展示了 SYP-Lasso 的优越性能，从而证明我们的 Syn-YawPitch 数据库能够代表真实世界中采集的数据。

性能。我们强调，统一质量度量会考虑影响识别性能的所有质量要素，包括底层人脸识别系统所带来的人口统计偏差和非人口统计偏差 [57]。相比之下，组件质量度量侧重于研究单个质量要素的影响。

总之，与受 ISO/IEC CD3 29794-5 [3] 启发的现有姿态质量估计算法相比，我们证明了所提出的 SYP-Lasso 模型具有优越性。此外，在姿态质量估计方面，我们取得了与成熟的统一质量评估算法相当的结果，同时：1）仅使用其模型参数的一小部分；2）通过应用简单的回归模型，得到可解释的质量度量。

## 7 结论

在本研究中，我们介绍了 Syn-YawPitch，这是一种利用三维感知生成对抗网络 [16] 生成的合成数据集，旨在编辑偏航角和俯仰角时保留面部几何结构。此外，我们分析了偏航角和俯仰角对四种开源人脸识别系统及一种商用现成（COTS）人脸识别系统的影响。分析结果表明，超过 30°的正向俯仰旋转会显著降低所有人脸识别系统的生物特征识别性能。我们认为，这一趋势是由面部区域对识别系统的可见度降低所致。

头部姿态与匹配比较分数之间的相互关系，被提炼为一个轻量级的 Lasso 回归模型。事实证明，该模型优于基于 ISO/IEC 标准的基线方法，例如若干最先进的统一质量评估算法。SYP-Lasso 已向研究界开放，我们设想它未来的用途将符合当前工作草案的预期：*ISO/IEC CD329794-5* [3]，并进一步迈向更具可解释性的人脸图像效用度量。未来的研究可以着重开发通用姿态质量估计，因为在各种偏航角-俯仰角组合下，所评估的人脸识别系统的表现始终一致。

## 致谢

本研究得到了德国联邦教育与研究部以及黑森州高等教育、研究、科学与艺术部的支持，属于两部门对应用网络安全国家研究中心 ATHENE 的联合资助项目。

## 参考文献

[1] 欧盟理事会，“欧盟第2226/2017号理事会条例：建立出入境系统（EES）”，2017年。https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32017R2226。[2] 欧盟理事会，“欧盟第1862/2018号理事会条例：申根信息系统（SIS）的建立、运行和使用”，2018年。https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32018R1862。[3] ISO/IEC JTC1 SC37 生物特征识别，*ISO/IEC CD3 29794-5* 信息技术——生物特征样本质量——第*5*部分：人脸图像数据。国际标准化组织，2023年。[4] ISO/IEC JTC1 SC37 生物特征识别，*ISO/IEC FDIS 29794-1* 信息技术——生物特征样本质量——第*1*部分：框架。国际标准化组织，2023年。[5] D. Osorio-Roig、T. Schlett、C. Rathgeb、J. Tapia 和 C. Busch，“探索用于生物特征识别中工作负载降低的质量评分”，载于国际生物识别与取证研讨会论文集，2022年4月。[6] T. Schlett、C. Rathgeb、O. Henniger、J. Galbally、J. Fierrez 和 C. Busch，“人脸图像质量评估：文献综述”，*ACM*计算调查，第54卷，第10s期，第1–49页，2022年。[7] Q. Meng、S. Zhao、Z. Huang 和 F. Zhou，“MagFace：一种用于人脸识别和质量评估的通用表征”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第14225–14234页，2021年。

[8] J. Hernandez-Ortega, J. Galbally, J. Fierrez, and L. Beslay, “生物特征质量：综述及其在使用FaceQnet进行人脸识别中的应用,” arXiv预印本 arXiv:2006.03298, 2020年。[9] P. Terhorst, J. N. Kolf, N. Damer, F. Kirchbuchner, and A. Kuijper, “SER-FIQ：基于随机嵌入鲁棒性的人脸图像质量无监督估计,” 载于IEEE/CVF计算机视觉与模式识别会议论文集，第5651–5660页，2020年。[10] ISO/IEC JTC1 SC37 生物特征识别，ISO/IEC 39794-5:2019 信息技术——可扩展生物特征数据交换格式——第5部分：人脸图像数据。国际标准化组织，2019年。[11] T. Karras, S. Laine, and T. Aila, “生成对抗网络的风格化生成器架构,” 载于 IEEE/CVF计算机视觉与模式识别会议论文集，第4401–4410页，2019年。[12] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, “分析并提升StyleGAN的图像质量,” 载于 IEEE/CVF 计算机视觉与模式识别会议论文集，第 8110–8119 页，2020年。[13] T. Karras, M. Aittala, S. Laine, E. H¨ark¨onen, J. Hellsten, J. Lehtinen, and T. Aila, “无混叠生成对抗网络,” 神经信息处理系统进展，第34卷，第852–863页，2021年。[14] S. J. Nightingale and H. Farid, “人工智能合成人脸与真实人脸难以区分，且更值得信任,” 美国国家科学院院刊，第 119 卷，第 8 期，第 e2120481119 页，2022年。[15] B. Shen, B. RichardWebster, A. O’Toole, K. Bowyer, and W. Scheirer, “关于人类对合成人脸的感知研究,” 载于第16届IEEE国际自动人脸与手势识别会议，第1–8页，IEEE，2021年。[16] E. Chan, C. Lin, M. Chan, K. Nagano, B. Pan, S. D. Mello, O. Gallo, L. Guibas, J. Tremblay, S.Khamis, 等，“高效的几何感知三维生成对抗网络,” 载于IEEE/CVF计算机视觉与模式识别会议论文集，第16123–16133页，2022年。[17] T. Nguyen-Phuoc, C. Li, L. Theis, C. Richardt, and Y. Yang, “Holo-GAN：从自然图像中无监督学习三维表示,” 载于IEEE/CVF国际计算机视觉会议论文集，第7588–7597页，2019年。[18] M. Niemeyer and A. Geiger, “GIRAFFE：将场景表示为组合式生成神经特征场,” 载于 IEEE/CVF计算机视觉与模式识别会议论文集，第11453–11464页，2021年。[19] S. Medin, B. Egger, A. Cherian, Y. Wang, J. Tenenbaum, X. Liu, and T. Marks, “MOST-GAN：用于解耦人脸图像操控的三维可变形StyleGAN,” 载于 AAAI人工智能会议论文集，第36卷，第1962–1971页，2022年。[20] J. L. Sch¨onberger and J.-M. Frahm, “重新审视运动恢复结构,” 载于计算机视觉与模式识别会议，2016年。[21] J. L. Sch¨onberger, E. Zheng, M. Pollefeys, and J.-M. Frahm, “用于非结构化多视图立体视觉的逐像素视图选择,” 载于欧洲计算机视觉会议，2016年。[22] Y. Shen, C. Yang, X. Tang, and B. Zhou, “InterFaceGAN：解读生成对抗网络学习到的解耦人脸表示,” IEEE模式分析与机器智能汇刊，2020年。[23] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, “ArcFace：用于深度人脸识别的加性角度间隔损失,” 载于IEEE/CVF计算机视觉与模式识别会议论文集，第4690–4699页，2019年。[24] Y. Huang, Y. Wang, Y. Tai, X. Liu, P. Shen, S. Li, J. Li, and F. Huang, “CurricularFace：用于深度人脸识别的自适应课程学习损失,” 载于 IEEE/CVF计算机视觉与模式识别会议论文集，第5901–5910页，2020年。[25] M. Kim, A. Jain, and X. Liu, “AdaFace：用于人脸识别的质量自适应间隔,” 载于 IEEE/CVF计算机视觉与模式识别会议论文集，第18750–18759页，2022年。[26] G. Fanelli, T. Weise, J. Gall, and L. V. Gool, “使用消费级深度相机进行实时头部姿态估计,” 载于联合模式识别研讨会，第101–110页，Springer，2011年。[27] R. Tibshirani, “通过套索回归进行回归收缩与变量选择,” 英国皇家统计学会杂志：B辑（方法论），第58卷，第1期，第267–288页，1996年。[28] T. Karras, T. Aila, S. Laine, and J. Lehtinen, “渐进式生成生成对抗网络，以提升质量、稳定性和变化度,” arXiv预印本 arXiv:1710.10196，2017年。[29] X. Huang and S. Belongie, “借助自适应实例归一化实现实时任意风格迁移,” 载于 IEEE国际计算机视觉会议论文集，第1501–1510页，2017年。[30] V. Sitzmann, J. Thies, F. Heide, M. Nießner, G. Wetzstein, and M. Zollhofer, “DeepVoxels：学习持久的三维特征嵌入……”，载于《IEEE/CVF计算机视觉与模式识别会议论文集》，第2437–2446页，2019年。[31] B. Mildenhall、P. Srinivasan、M. Tancik、J. Barron、R. Ramamoorthi和R. Ng，“Nerf：将场景表示为用于视图合成的神经辐射场”，《ACM通讯》，第65卷，第1期，第99–106页，2021年。[32] T. Hempel、A. A. Abdelrahman和A. Al-Hamadi，“用于非受限头部姿态估计的6D旋转表示”，arXiv预印本 arXiv:2202.12555，2022年。[33] I. Joshi、M. Grimmer、C. Rathgeb、C. Busch、F. Bremond和D. Dantcheva，“人体分析中的合成数据：综述”，arXiv预印本 arXiv:2208.09191，2022年。[34] A. Kortylewski、B. Egger、A. Schneider、T. Gerig、A. Morel-Forster和T. Vetter，“利用合成数据分析并减少数据集偏差对人脸识别造成的损害”，载于《IEEE/CVF计算机视觉与模式识别会议研讨会论文集》，第0–0页，2019年。[35] L. Colbois、T. de Freitas Pereira和S. Marcel，“关于使用自动生成的合成图像数据集进行人脸识别基准测试”，载于《IEEE国际联合生物识别会议》，第1–8页，IEEE，2021年。[36] R. Gross、I. Matthews、J. Cohn、T. Kanade和S. Baker，“Multi-pie”，《图像与视觉计算》，第28卷，第5期，第807–813页，2010年。[37] H. Zhang、M. Grimmer、R. Ramachandra、K. Raja和C. Busch，“合成数据在人脸识别中的适用性”，载于《国际生物识别与取证研讨会》，第1–6页，IEEE，2021年。[38] M. Grimmer、H. Zhang、R. Ramachandra、K. Raja和C. Busch，“时光飞逝：利用合成数据分析人脸老化对识别性能的影响”，载于《生物识别特别兴趣小组国际会议》，第1–6页，IEEE，2022年。[39] Y. Alaluf、O. Patashnik和D. Cohen-Or，“风格问题：使用基于风格的回归模型进行年龄变换”，《ACM图形学汇刊》，第40卷，第4期，第1–12页，2021年。[40] M. Grimmer、H. Zhang、R. Ramachandra、K. Raja和C. Busch，“由身份先验引导的非确定性合成人脸数据集生成”，载于《挪威信息安全会议论文集》，2021年。[41] B. Lu、J.-C. Chen、C. D. Castillo和R. Chellappa，“对非受限人脸验证中协变量效应的实验评估”，《IEEE生物识别、行为与身份科学汇刊》，第1卷，第1期，第42–55页，2019年。[42] J. Gu、X. Yang、S. De Mello和J. Kautz，“动态面部分析：从贝叶斯滤波到循环神经网络”，载于《IEEE计算机视觉与模式识别会议论文集》，第1548–1557页，2017年。[43] I. L¨usi、S. Escarela和G. Anbarjafari，“SASE：用于人体头部姿态估计的RGB-深度数据库”，载于《欧洲计算机视觉会议》，第325–336页，Springer，2016年。[44] N. Gourier、D. Hall和J. L. Crowley，“通过稳健检测显著面部结构估计人脸朝向”，载于《FG Net指示性手势视觉观测研讨会》，第6卷，第7页，Citeseer，2004年。[45] X. Zhu、Z. Lei、X. Liu、H. Shi和S. Z. Li，“大姿态下的人脸对齐：一种三维解决方案”，载于《IEEE计算机视觉与模式识别会议论文集》，第146–155页，2016年。[46] W. Gao、B. Cao、S. Shan、X. Chen、D. Zhou、X. Zhang和D. Zhao，“CAS-PEAL大规模中国人脸数据库及基线评估”，《IEEE系统、人类与控制论汇刊A辑：系统与人类》，第38卷，第1期，第149–161页，2008年。[47] C. E. Thomaz和G. A. Giraldi，“一种新的主成分分析排序方法及其在人脸图像分析中的应用”，《图像与视觉计算》，第28卷，第913页，2010年6月，第902页。[48] P. Li、X. Wu、Y. Hu、R. He和Z. Sun，“M2FPA：用于面部姿态分析的多偏航角、多俯仰角高质量数据集与基准”，载于《IEEE/CVF国际计算机视觉会议论文集》，第10043–10051页，2019年。[49] T. Zheng和W. Deng，“跨姿态LFW：用于研究非受限环境下跨姿态人脸识别的数据库”，技术报告18-01，北京邮电大学，2018年2月。[50] T. Li、T. Bolkart、M. J. Black、H. Li和J. Romero，“从四维扫描中学习面部形状与表情模型”，《ACM图形学汇刊》，第36卷，第6期，第194–1页，2017年。[51] R. Danˇeˇcek、M. J. Black和T. Bolkart，“EMOCA：情感驱动的单目人脸捕捉与动画”，载于《IEEE/CVF计算机视觉与模式识别会议论文集》，第20311–20322页，2022年。

[52] K. He、X. Zhang、S. Ren和J. Sun，“用于图像识别的深度残差学习”，载于*IEEE*计算机视觉与模式识别会议论文集，第770–778页，2016年。[53] J. Behrmann、W. Grathwohl、R. T. Chen、D. Duvenaud和J.-H. Jacobsen，“可逆残差网络”，载于国际机器学习会议，第573–582页，PMLR，2019年。[54] F. Boutros、M. Fang、M. Klemt、B. Fu和N. Damer，“CR-FIQA：通过学习样本相对可分类性进行人脸图像质量评估”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第5836–5845页，2023年。[55] P. Grother、A. Hom、M. Ngan和K. Hanaoka，“持续进行的人脸识别供应商测试（FRVT）第5部分：人脸图像质量评估（第4版草案）”，美国国家标准与技术研究院技术报告，第1卷，第5期，2021年。[56] T. Schlett、C. Rathgeb、J. Tapia和C. Busch，“关于生物特征质量评估算法评估的考量”，*arXiv*预印本 *arXiv:2303.13294*，2023年。[57] P. Terh¨orst、J. N. Kolf、N. Damer、F. Kirchbuchner和A. Kuijper，“人脸质量估计及其与人脸识别中人口统计偏差和非人口统计偏差的相关性”，载于国际联合生物识别会议，第1–11页，IEEE，2020年。