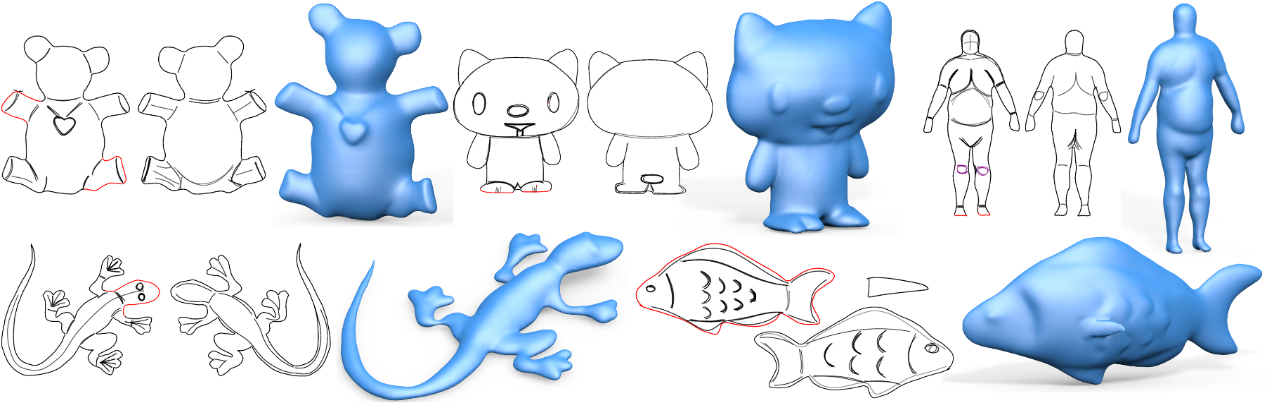
下图是用我们的方法计算的草图和相应的3D形状。红色轮廓曲线部分被重新绘制以提供更清晰的边界位置数据（参考第4节）。比如抬起泰迪四肢和壁虎头以及弯曲鱼体。人体草图的紫色曲线用目标曲率指定（参考第3.1节）。

对于每个草图，我们的方法是使用CNN来预测3D表面补丁。两个或多个表面补丁融合成完整的形状。后(底)视图重复使用前(顶)视图的轮廓曲线，但是多了内部笔画。 鱼的那个小的三角形草图是为了对其侧鳍建模。 请注意，较暗和过度勾画的地方通常对应于较强的曲率，这看起来自然而直观。



草图提供了一个直观的用户界面，用于显示各种形状。 虽然人类可以很容易地通过想象交流的形状，但在算法上复制这个过程需要解决许多歧义。（2D转3D 需要解决丢失的一维信息）现有的基于草图的建模方法通过费力的用户注释或将建模形状限制为特定的类别来解决这些歧义。 我们提出了一种由稀疏2D自由形式草图构建3D表面的通用方法，该方法通过将卷积神经网络 (CNN) 合并到草图处理流程中来克服这两个限制。

给定 3D 表面的 2D 草图，我们使用 CNN 来推断表示表面的深度和法线图。为了消除歧义，我们引入了一个中间 CNN 层，该层对表面的密集曲率方向或流场进行建模，并产生一个额外的输出置信度图以及深度和法线。流场指导我们随后的表面重建以提高规律性；无监督训练的置信度图测量歧义，并为数据拟合提供稳健的估计量。为了减少输入草图中的歧义，用户可以通过在稀疏点处提供可选的深度值和笔画的曲率提示来优化他们的输入。我们的 CNN 在一个大型数据集上进行训练，该数据集是通过使用模仿人类对自由形状草图的非真实感线渲染 (NPR) 方法渲染各种 3D 形状的草图而生成的。我们使用 CNN 模型来处理单视图和多视图草图。使用我们的多视图框架用户通过在不同视图中绘制草图逐步完成形状，生成完整的闭合形状。对于每个新视图，建模由早期视图中积累的表面部分草图和深度提示辅助，使用预测的置信度作为权重，将部分表面融合成一个完整的形状。

# 1 INTRODUCTION

2D草图描绘了3D形状的丰富的几何特征，如轮廓、遮挡及隐藏轮廓、山脊和山谷以及阴影线，从而为3D形状提供了一种简洁直观的表达形式。为此提供了一组非真实感的渲染(NPR) 方法 。已成功用于从 3D 形状中自动提取富有表现力的草图。从稀疏的 2D 草图建模 3D 形状的逆过程也展示了它在对传统 CAD 系统难以实现的自由曲面建模方面的巨大潜力。然而，由于草图的简单性和 2D 到 3D 维度的高程导致固有的模糊性，这个逆问题更具挑战性。以前的许多成果都使用人工注释的几何先验知识来消除歧义；例子包括（见原文）。虽然这些方法具有强大的建模能力，但它们通常需要用户详细注释草图，以便可以应用它们各自的几何先验。最近的新作品，使用数据驱动的方法并训练了一个包罗万象的机器学习模型，例如用卷积神经网络 (CNN)，将输入草图直接映射到 3D 形状 。这些方法通过为每个对象类别（例如椅子、飞机、字符）构建特定的机器学习模型来解决歧义问题，但是这会将建模的形状局限到仅具有用于训练的大型 3D 数据集的类别。

本论文提出了一种基于草图的方法，该方法能够对常见的自由曲面和完整的 3D 形状进行建模，并且具有稀疏但富有表现力的用户交互效果。用户在 2D 平面上绘制线条草图，对深度和法线表示的曲面片进行建模。在多个观测面上逐步绘制草图会形成完整形状的表面补丁。我们方法的核心是一个 CNN，它将输入的 2D 草图映射到其目标捕获的常规自由曲面。

我们在一个大型生成的数据集上训练 CNN，该数据集具有不同的 3D 形状和由 NPR 方法渲染的平面草图，NPR方法模仿人类如何描绘模型。由于与相同稀疏输入草图相对应的不同 3D 形状具有固有模糊性（就是2D图一样，3d图不一样），所以为此类任务训练的简单 CNN 结构表现不佳。因为没有针对于普通自由曲面建模的特定类别先验信息，我们通过使用基本的几何和统计学原理来解决歧义。首先，我们不是直接从草图映射到几何图形，而是引入了密集曲率方向（流）场的中间层，CNN 首先对其进行回归，然后与草图结合以推断曲面形状。流场为几何推断提供了密集的局部指导，并且用具有响应草图线的增强变换产生了更规则的表面（不懂）。其次，我们使用在不确定区域上较低而在其他情况下较高的置信度图明确地对草图的感知模糊性进行建模。虽然对人类观察者来说很明显，但模棱两可或

置信图很难用手工规则量化。我们让 CNN 预测置信度图，并通过使用稳定的统计数据以无监督的方式对其进行训练。最后，我们允许修改预测的表面，使 CNN 能够识别用户可选择添加的规范稀疏点深度​​提示（就是第一张图人为添加的红线 边缘轮廓 紫圈的注释能让计算机看得懂），或沿笔划的尖锐特征和表面曲率提示。

除了基于单视图表面建模，我们还开发了一个多视图系统，通过在不同视图中绘制来逐步创建完整的 3D 形状。在先前视图中绘制曲面片后，用户通过在新视图中绘制来旋转和更新部分 3D 形状，其中渲染的部分草图和从现有曲面采样的深度提示可以帮助建模。最后，通过融合由相应置信度图加权的表面，生成完整的 3D 形状。

我们通过大量示例验证我们的方法，并将其与以前的方法、替代结构和其他网络进行比较，以对特定类别的对象进行建模。结果表明，我们的网络生成具有丰富变化和微妙细节的规则形状，同时需要较少的用户输入。新手用户的评估还表明，我们的工具使自由形状的建模对普通用户有效且易于使用。

# 2 RELATED WORK

在几十年的研究中，基于草图的自由形状建模积累了丰富的文献。然而，我们非常粗略地将之前的工作分为两组，一组是更经典的几何推理方法，它依赖于几何先验来重建 3D 形状，另一组是构建机器学习模型以从特定 3D图形 学习先验的数据驱动方法。我们的方法实际上借鉴了两组的想法，因为我们首先建立在几何原理上，包括投影曲率方向和 3D 形状之间的映射以及法线/深度一致性，以解决通用的自由形式建模问题，但也利用了强大的拟合CNN 执行解析草图和生成密集曲率方向场以及从投影几何数据映射到空间深度和法线的多因素非线性任务的能力。下面我们简要讨论一下之前最相关的作品

几何推理法。大量先前的工作都集中在开发有效的几何原理上，在此基础上，可以从用户草图中准确推断出 3D 形状而不会产生歧义。用户勾勒出轮廓与膜功能相结合 [Igarashi 等人。 1999；乔希和卡尔 2008；尼伦等人。 2007；叶等人。 2016;张等人。 2001] 或其他平滑插值函数 [Bernhardt 等人。 2008；奥尔森等人。 2011；施密特等人。 2005] 可以快速确定平滑的低频 3D 形状。草图轮廓曲线变形并混合用户注释的图元（例如广义圆柱体、金字塔等）以形成复杂的形状 [Chen 等人。 2013；金戈尔德等人。 2009;什托夫等人。 2013]。基于艺术家如何描绘具有规则组合的轮廓、特征和代表性曲面曲率线的自由形状，一系列作品 [Bae et al. 2008； Iarussi 等人。 2015；李等人。 2017；施密特等人。 2009;邵等人。 2012；徐等人。 2014] 解决 2D 到 3D 的歧义，并将平面草图转换为 3D 数据。

我们的方法是根据数据驱动的方法来学习基于同一组原则的几何推断，即形状规则和轮廓、特征和代表性曲率线的形状。另一方面，虽然这些先前的方法依赖于详细的用户注释将草图解析为不同功能的曲线，并且经常使用昂贵的非线性数值优化来解决 2D 到 3D 的转换，但我们利用 CNN 模型来解析草图并推断几何图形提高了效率并减少了用户草图和注释。见Sec 5.2 比较。

数据驱动的方法。对于许多常见的物体和场景，通常我们人类通过首先识别它们是什么然后将记忆中相同类别的先前形状与观察结果匹配来设想它们的 3D 形状。这个想法是基于草图建模的一系列数据驱动方法的基础，因为它们通常将建模任务分为两个步骤：首先通过数据库针对输入草图搜索匹配的形状，然后根据需要调整检索到的形状以适应输入草图。（看到一个熊，首先想起来它有四个爪子，尖牙，棕色。然后再看图片这只胖瘦，结合对熊的固有印象及图片熊的特征构建图片熊的3D模型）示例包括纯基于草图的检索 [Eitz et al. 2012；苏等人。 2015；王等人。 2015c]，以及随后的改编和组合检索，如 Sketch2Scene [Xu et al. 2013]用于场景建模和[Guo et al。 2016;李和芬克豪斯 2008；谢等人。 2013] 用于对象建模。虽然这些方法为用户尝试建模的特定类别对象提供了丰富的先验知识来显着减轻用户负担，但是为一个类别构建的工具并不能推广到其他类别。相比之下，我们的机器学习模型学习的是更通用的几何重建过程，而不是特定类别的 3D 对象的知识，这使得我们的方法对特定类别的对象进行建模的效率可能较低（精准度低），但更通用，提供了更精细的形状控制水平。

该领域的后续工作并未明确分离模型搜索和调整步骤，而是依靠强大的深度神经网络直接从草图映射到 3D 数据，示例包括 [Delanoy et al. 2017；伦等人。 2017；苏等人。 2018]。 [苏等人。 2018] 通过编码器-解码器网络从特定类别的 2D 草图预测法线补丁，最大限度地减少法线拟合损失和对抗损失，并将用户指定的法线样本作为可选输入。 [德拉诺伊等人。 2017] 使用 CNN 将草图映射到表示 3D 形状的体积占用网格，并允许用户在新视图中绘制草图时通过更新器 CNN 对形状进行增量更新。然而，结果表明，为每个对象类别训练的 CNN 并没有推广到其他类别。此外，体积表示限制了建模形状的分辨率。 Lun 等人的工作。 [2017] 将来自规范视点（正面、侧面、顶部）的特定类别平面草图输入到具有编码器和 13 个解码器的 CNN，每个解码器输出 13 个预定义视点之一的深度和法线图，然后将它们融合在一起成 3D 网格。

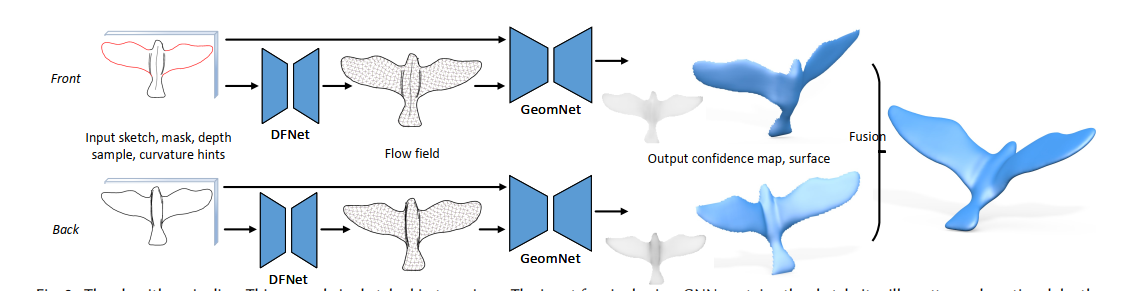
不同于 [Delanoy 等人。 2017] 和 [Lun 等人。 2017]解决了训练类别的完整 3D 形状的生成，我们的工作重点是对表示为深度图的自由曲面进行建模，同时还提供了一种多视图融合方法来将表面组合成完整的 3D 模型。通过使用一般几何规则和从草图中学习形状的先验来一次建模一个表面，我们的方法对形状类别是不可知的。然而，我们注意到将一个完整的 3D 形状分解为多个表面块以按顺序建模并不总是那么简单。为用户构想，我们认为这是为无类别优势付出的应有代价。为了帮助用户建模，我们的多视图交互过程允许用户逐步在任意视图中为形状的不同部分绘制草图，在其他视图中建模的表面有助于在新视图中绘制草图（第 4 节）。

程序和参数模型提供了另一种先验知识，有效地将建模任务简化为从草图到模型参数的映射。许多作品从数据中学习映射，用于建模城市建筑 [Nishida et al. 2016]，面对 [Han et al. 2017] 等 [Huang et al. 2016]。这些方法是为给定的参数模型量身定制的，不能推广到其他自由形状。

最近的工作直接从 2D 图像重建深度和法线图表示的 3D 形状和场景 [Eigen and Fergus 2015;塔塔尔琴科等人。 2016;王等人。 2015a]，点云 [Fan et al. 2017] 或体积网格 [Choy et al. 2016;塔塔尔琴科等人。 2017；吴等人。 2016]，利用数据驱动和 CNN 模型。在本文中，我们专注于从包含比图像更稀疏的信息的草图中重建高质量的 3D 形状，并为用户提供方便的 3D 建模控制。

# 3 SINGLE VIEW MODELING

在单个视图中，我们从稀疏的平面草图中恢复深度和法线数据。 这个过程有两个主要挑战。 首先，草图中的稀疏笔划具有不同的含义，每一种对相应 3D 形状的邻近区域的影响不同； 我们需要一致地解析笔画并在整个平面区域内插值它们的数据以推断 3D 表面。 为了解决这个问题，我们依靠 CNN 模型以最少的用户规范自动解析不同的输入行，从而节省了大量的用户工作量。 如同图2的示例，从未标记的输入笔划中自动地将脊和谷区分开。



其次，2D 草图对于它们所代表的 3D 形状具有固有的模糊性，这可能会导致任何强大的机器学习模型在尝试对重建进行无脑回归的方式失败。 以前的方法通常通过限制常见类的形状来解决歧义； 因此，这样的模型适用于它所训练的特定类别，但不能推广到其他类别 [Delanoy 等人。 2017； 伦等人。 2017]。 相反，我们努力实现更通用的自由形状建模，并专注于使用几何原理和可选的用户输入来消除歧义。

总而言之，我们单视图建模的核心是一个两阶段的CNN 回归模型（图 2）：

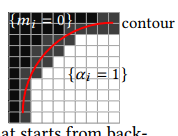
1. 给定输入草图，第一阶段子网络 (DFNet) 对进行回归流场，流场是描述表面曲率方向并指导其重建的密集信号（第 3.2 节）。
2. 第二阶段子网络（GeomNet）采用草图和流场的信息，并预测深度/法线图，以及显示输入草图每个点的模糊度的置信度图（第 3.3 节）。

此外，用户可以通过在笔划上提供曲率提示或在稀疏样本点上提供深度值（就是修改笔画，画的弯一点或者深一点）来进一步修改表面并解决歧义； 我们的 CNN 模型经过训练可以利用这些可选输入。 接下来我们详细讨论单视图建模。

## 3.1 Input and output

### Input

单视图建模的输入主要是由单通道的灰度图像表示的2D草图，即为描绘3D对象在当前视图上的平行投影。我们要求将轮廓和遮挡轮廓与草图中的其他笔划区分开来，方法是强制（遮挡）轮廓为纯黑色，而其他笔划具有非零灰度值。这种区别对于明确指定深度和法线与（遮挡）轮廓中断的位置至关重要，而其余笔划，包括脊/谷、暗示轮廓、曲率线等，主要描述平滑的形状变化。事实上，这样的线条样式符合 NPR 文献 [Rusinkiewicz 和 DeCarlo 2013]，这表明人们使用粗线条来强调边界、不连续性和大曲率。我们的训练数据生成和建模工具遵循这种风格。图 1,2,15 显示了笔触样式如何自然导致不同表面变化的示例。请注意，对于真实的用户草图，由折线表示的输入笔画在被渲染并馈送到网络之前会被略微平滑以去除噪声伪影。



二进制轮廓蒙版由输入轮廓草图构建而成，用于网络输入和训练损失计算（就是说二进制蒙版的输入是由草图生成的，描述草图的前景背景，计算训练损失）。 草图轮廓遮罩 {mi } 从 mi = 0（插图，黑色）的背景中描绘出 mi = 1 的前景对象，并通过从背景像素开始在任意草图像素处停止的泛洪过程计算； 对于有孔的草图，用户必须在每个孔内选取一个背景点，以启用正确的掩码计算。 {mi } 是网络的输入。 训练样本的形状轮廓掩码 {mi′} 是渲染几何的前景像素掩码； 它通常与草图轮廓蒙版 {mi } 的不同之处在于缩小了一个小边距，因为轮廓线确实有几个像素的宽度。 信号过滤掩码 {αi }（插图，白色）去除不连续（遮挡）轮廓像素处的无效信号、锐利特征（如果有）和背景。 {mi′} 和 {αi } 用于训练损失评估

深度样本是可选输入，用于在某些点提供深度提示。 它们由单个通道图编码，样本点的像素具有指定的深度值，其他像素为零值（3D 模型已定位以确保正深度值）。 样本点可以沿着轮廓放置并由单个像素编码，也可以在放置在轮廓内并由 3x3 像素补丁编码。

曲率提示是可选的用户输入，用于指定表面法线经历突然变化的尖锐特征的存在，以及表面曲率，或表面法线如何变化，跨越山脊、山谷或尖锐特征。 我们使用具有三个通道的地图来对提示进行编码。 一个通道是二进制掩码，对于锐利特征的像素为 1，否则为 0。 第二个通道也是二进制掩码，用于指示存在目标曲率，其值在第三个通道中给出。 对于笔划上的像素目标曲率值为 ∈ [−2, 2]，其中 t 是训练 3D 模型上相应空间曲线的单位切向量，⊥ 是将向量旋转 90° in 切平面，Δn 是固定步长下表面法线向量沿 t⊥ 的差值（我们使用对象边界框对角线长度的 3%）。 注意曲率值的符号表示整个笔划表面的凸度。

总共，我们用于表面预测的 CNN 的输入图有 6渠道。 （灰度图1个，轮廓蒙版1个，深度1个，曲率3个）

### Output

输出图由 1 个通道的深度、3 个通道的法线向量和 1 个通道的置信度值组成。 输出图与输入具有相同的分辨率。虽然图像大小可以是任意的，因为 CNN 是完全卷积的，但我们使用 256 × 256 来平衡 3D 数据的质量和训练成本。

## 3.2 Flow field regression

草图中的许多线提供有关表面弯曲（或曲率）方向的信息，这对于恢复 3D 形状至关重要。 我们没有像以前那样通过启发式算法和非线性优化来解析笔画和解决流场，而是通过 CNN 预测来自动化整个过程。

值得一提的是，在 [Groueix et al. 2018] 使用图集参数化来调节解码器网络预测的表面。 类似地，在我们的框架中，流场为第二阶段几何重建提供了在 2D 域的每个点处的局部平滑引导，以实现数据的规范传播。 它会导致更规则的形状和增强的变化，如 Sec 5.4. 所示。.

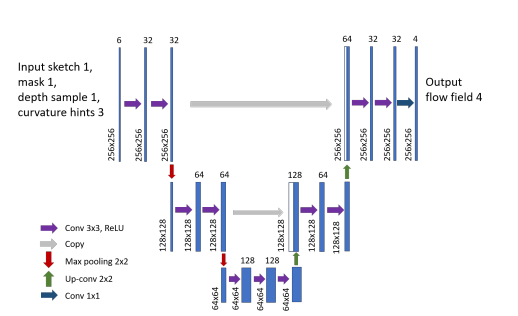
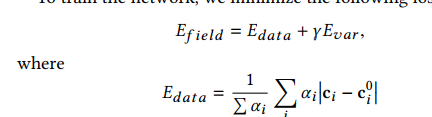


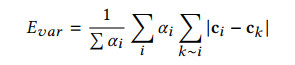
图 3. DFNet 的结构。 它是一个编码器-解码器网络，具有三个域分辨率。 输入是一个 6 通道的图像，其中包含草图、其轮廓蒙版以及可选的深度样本和曲率提示。 输出为 4 通道流场。 每个特征图旁边的数字显示其空间大小和通道数。

(‘’’以下文字暂时不懂‘’‘)

DFNet 子网络（图 3）是基于常用的 U-Net [Ronneberger et al. 2015]，具有三个级别的图像分辨率。 给定之前讨论过的所有输入，网络为第 i 个像素输出一个具有四个通道 ci = (ci0,ci1,ci2,ci3) 的密集映射，它通过编码非正交 4 方向场 u, v ∈ C ，。 这种表示通过避免匹配各个方向的需要来促进相邻像素处的方向比较。 为了训练网络，我们最小化以下损失函数



测量回归场与 3D 形状的投影曲率方向场 {ci0} 的绝对差，以及



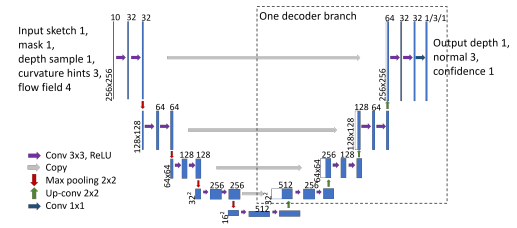
是回归场的总变化，第 k 个像素是第 i 个像素的右邻或上邻。 场总变化的惩罚是必要的，因为表示表面弯曲方向的稀疏草图线不能完全捕捉地面实况投影方向场的方差，我们鼓励从线条中导出规则方向场。 我们使用了 γ = 0.1

信号过滤掩码 {αi } 重置由（遮挡）轮廓线、锐利特征（如果有）和背景覆盖的像素处的损失。 使用遮罩的原因是，首先，（遮挡）轮廓上的点的法线向量背向侧面，并且 3D 曲率方向的投影在那里退化，导致这些像素上的训练数据不可靠，其次，表面在 （遮挡）轮廓像素，并且不期望它们之间的方向场平滑度。 当存在尖锐特征时，它们的像素具有不可靠的法向量，并且也被排除在损失计算之外。 当输入到 GeomNet 子网络时，输出字段也被 {αi} 屏蔽。

(‘’’以上文字暂时不懂‘’‘)

## 3.3 Robust flow-guided surface regression

输入草图、剪影蒙版和可选的深度样本和曲率提示与回归流场叠加在一起，并输入 GeomNet 子网络以预测正常和表示 3D 表面的深度图。 GeomNet（图 4）是一个具有五级图像分辨率的编码器-解码器结构和三个解码器分支共享同一个编码器，输出分别是法线图、深度图和置信图。

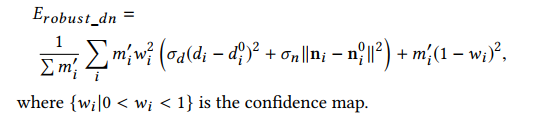


(‘’’以下文字暂时不懂‘’‘)

训练网络，我们最小化以下损失函数：



稳健的数据拟合。 第一项测量回归深度 di 和法线向量 ni 与从 3D 形状中获取的真

值d0i 和 ni0 的差异：

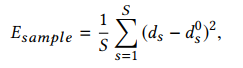
解码器分支通过最后一层 sigmoid 激活输出置信度图。 置信度图的函数本质上是一个鲁棒的估计量 [Black and Rangarajan 1996]：它趋向于 1，按照 Erobust\_dn 中第二项的要求，强制学习真实值数据，但如果回归的深度和法线不能 匹配基本事实。 直观地说，置信图估计输入草图的每个点有多少模糊度，因为模糊度是训练有素的几何回归模型不匹配的最终原因。 参见图 9 示例置信度图，显示与预期相符的歧义量。

尽管对人类来说模糊性是显而易见的，但它很难量化。 这里置信图的训练是无监督的，避免了生成真实值数据的需要。 置信度图的好处主要是通过放宽高度模糊区域的深度和正常拟合损失来使学习更加平易近人。 这在第二节的消融测试中显示。 5.4. 此外，在融合过程中，来自多个视图的表面被调整为完整的 3D 形状，这有助于对点进行加权，以便我们可以保持可靠并更改其余部分（第 4 节）。 最后，它可以被视为对用户的反馈，即网络预测在哪里不确定，并且可以使用更多的输入数据来减少歧义。

σd , σn 归一化深度和正态回归的误差。 为了估计它们，我们在没有置信度图的情况下训练深度和法线的直接回归（第 5.4 节），并在测试数据集上收集预测深度和法线的平均误差，据此我们得到 σd = 920 ≈ 1/( 2×0.02322)，σn = 430 ≈ 1/(2×0.03392)。

稳健的数据拟合与贝叶斯深度学习框架 [Kendall and Gal 2017] 有相似之处，其中由于 CNN 模型和数据噪声导致的不确定性由与我们使用的特定置信度图非常相似的变量建模。 然而，由于在我们的任务中，训练数据是从地面实况 3D 形状生成的，因此置信度图主要模拟 CNN 回归模型的不确定性，这直观地源于将草图映射到 3D 形状的模糊性。

深度样本约束。 如果提供了深度样本，则样本点处的回归深度图应与指定的深度值 ds0 匹配，我们鼓励通过以下方式：



其中 S 是样本点覆盖的像素数。 深度和正常的一致性。 最后，对于有效的 3D 形状，回归的法线向量和深度值在几何上应该是一致的，通过以下方式测量：



其中 tix = 1, 0, (dj − di)/k , tiy = (0, 1, (dl − di)/k) 是像素 i 处的切向量，通过与其右侧和上相邻像素 j 的有限差分计算， l，k = 0.00784 是映射到训练数据规范尺度的像素宽度。 {αi } 用于重置不连续处的损失

我们使用 β = 5 来强调与给定样本点的匹配。 因为最初在训练过程中回归的法线和深度远不是正确的，所以我们放宽了一致性要求，逐渐增加； 具体来说，λ 从 5×10−4 变为 2，在训练期间每 5k 次迭代乘以 3.985，总共大约 42k 次迭代。 在实际用例中，真实数据中没有形状掩码 {mi′}； 输出贴图用轮廓掩码 {mi } 过滤

(‘’’以上文字暂时不懂‘’‘)

## 3.4 Data generation and network training

**数据生成**。 训练样本是成对的输入，即草图、掩码、可选的深度样本和曲率提示，以及相应的输出，即来自 3D 形状的投影曲率方向场和深度/法线图。 数据生成的详细信息在补充材料中； 下面我们给出一个概述

用于数据生成的自由平滑形状是计算机图形学中的常见模型。 它们涵盖了广泛的类别，包括动物、雕像和人物，其细节可以通过整洁的 2D 草图合理描绘。 为了使用户能够指定清晰的特征，我们还通过迭代应用滚动引导法线过滤来生成具有尖锐折痕的模型 [Wang et al. 2015b] 平滑形状。 通过使用这组多样化的对象，我们可以从独立于特定形状类别的草图中进行几何推断，从而针对通用的自由形状。

训练样本是通过模仿人类如何素描来描绘 3D 形状而生成的。草图由 NPR 使用程序 [Rusinkiewicz 和 DeCarlo 2013] 渲染生成。我们选择 NPR 的参数，以便人类观察者很容易从渲染的 2D 草图中相对较好地推断出 3D 表面。用于渲染的视点和缩放因子在有效范围内均匀采样，但经过过滤，以便观察到的表面块足够大并且大部分是正面的，这符合用户通常对对象的姿势。可选的深度样本沿部分轮廓曲线放置以锚定表面；在建模过程中，对于许多形状，轮廓可以简单地初始化到绘图平面，然后通过重新绘制来调整（第 4 节）。很少（少于 5 个）额外的深度样本在轮廓内生成，并放置在表面的极端深度值上。为检测到的 3D 表面的尖锐特征和脊/谷曲线随机生成可选的曲率提示。

总共有 260k 和 58k 样本分别用于训练和测试。

**网络训练**。 网络在两个阶段的过程中进行训练。 首先，我们训练 DFNet 子网络 10 个 epoch，最小化损失函数 Efield。 接下来，我们使用损失函数 Etotal 训练 GeomNet 子网络 10 个 epoch，同时修复 DFNet。 每个网络都在 Tensorflow 中实现，并由 Adam 求解器 [Kingma and Ba 2014] 在具有 4 个 Nvidia GeForce 1080Ti GPU 的机器上以固定学习率 10−3 进行优化。 DFNet 训练需要 8 小时，GeomNet 需要 16 小时。

# 4 MULTIPLE VIEW MODELING

要对完整的 3D 形状进行建模，必须使用来自不同视图的多个表面。 在多个视图中绘制草图可以对不同的部分进行建模，以融合成最终的完整形状。 在多视图过程中，现有表面可以通过在重叠区域提供部分草图和深度数据来辅助新零件的建模。 在本节中，我们将讨论如何将我们的单视图建模网络与多视图草图相结合，并启用方便的用户交互。

在旋转视图中重新绘制轮廓。 一个视点的等高线深度在通过旋转视图看时变成了剖面曲线，这使得通过在后一个视图中重新绘制剖面曲线来修改前一个视图的轮廓深度非常方便。 一个例子如图5所示：在原始视图中，我们要修改的部分轮廓曲线标记为（a）； 在另一个视图中给定目标重新绘制的轮廓曲线 (b)，用作 CNN 输入的附加深度样本提示，用于更新表面 (c)(d)

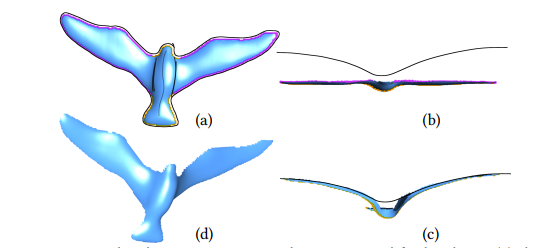
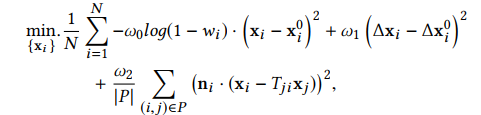


图 5. 在侧视图中重新绘制轮廓以修改形状。 (a) 与草图重叠的当前表面。 要重新绘制的轮廓线部分标记为紫色。 (b) 在侧视图中看到的表面，以及与标记轮廓匹配的新轮廓曲线。 (c) (d) 用新的深度提示更新表面

在帮助下以新视图绘制草图。在新视图中进行草绘时，先前视图的曲面会投影到当前视图，这有助于在当前视图中进行草绘和建模。 一方面，为现有表面生成显着的 NPR 曲线，用户可以简单地重复使用这些曲线。 另一方面，可见的现有表面为当前表面预测提供了深度线索。 例如，在图 2 中，将前视图表面的轮廓曲线及其深度数据直接复制到后视图，用户只需绘制内部笔划即可对后表面进行建模。

多视图融合。 在最后一步中，将所有表面放置到适当的位置并融合成一个完整的 3D 形状。 具有重叠区域的邻近表面可能不一致，我们通过以下二次优化解决：



其中 N 是来自所有视图的点数，xi 是第 i 个点在其局部视图的坐标系中的 3D 位置，wi 是预测置信度值，xi0 是预测位置，ni 是预测法向量，Δ· the Tutte Laplacian 算子 [Gotsman et al. 2003] 用于建立在像素的 3×3 邻域上的三个坐标（我们使用简单的 Tutte 权重，因为域是常规像素网格），以及 P 匹配点对的集合，其中第 j 个点来自其对应视图 通过帧变换 Tji 重新投影到第 i 个点的视图中，使其具有与 i 相同的 x,y 坐标

为了避免由于观看遮挡和不完美的部分重建带来的错误匹配，我们过滤掉一对匹配点，如果它们的距离大于包含 i、j 的两个视图的较小轮廓边界框对角线长度的 10%，或者如果 它们预测的法向量形成一个大于 70° 的角度就会被过滤掉。 置信度值映射到对数尺度，以强制保留更可靠的表面，同时调整置信度较低的部分。 拉普拉斯偏差的惩罚是尽可能地保持预测的形状。 我们使用 ω0 = 0.1,ω1 = 100,ω2 = 0.1 来实现平滑的适应过程，从而保留临界表面变化。

优化问题求解4次迭代； 在每次迭代中，更新点对应关系并使用迭代共轭梯度求解器求解线性最小二乘问题 100 次迭代。 图 6 显示了紧密融合补丁的示例。 请注意，如果我们使用包含大量冗余的密集像素的适当子集作为变量，则可以大大加快此融合过程。 这可以通过下采样来实现，我们将其留给未来的工作。

最后，通过 Screened Poisson Reconstruction [Kazhdan and Hoppe 2013] 从融合点集中提取表面网格，我们将适应点及其基于新位置更新的法向量传递给该点。

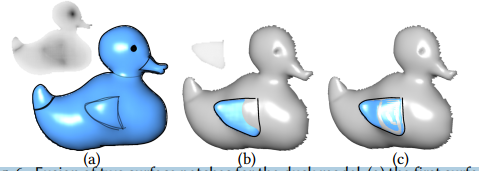


图 6. 鸭模型的两个表面补丁的融合。 (a) 绘制第一个表面，其置信度图显示在左上角。 由于自遮挡，鸭翼在此视图中不容易建模，置信图在此显示较低的值。 在稍微旋转的视图 (b) 中，鸭翼的草图绘制得当，左上角显示了高置信度。 (c) 两个表面在重叠区域紧密融合，偏向具有较高置信度值的翼片。

# 5 RESULTS AND DISCUSSION

## 5.1 Results

为了对无类别且可能是抽象的目标自由形状进行建模，交互式草图绘制过程通常非常直观：用户逐步探索和绘制以在新视图中添加更多细节甚至表面补丁。 图 7 显示了如何通过在基本平面上添加凸块和孔来绘制一个简单的抽象凹凸平面。 图 8 显示了通过一系列基本用户输入和编辑对各种恒星的建模，其中预测的表面相应地响应笔画、曲率提示、锐利特征和深度样本。 更多结果可见图。 1、2、13、15。随附的视频显示了使用我们的工具进行的实时建模。

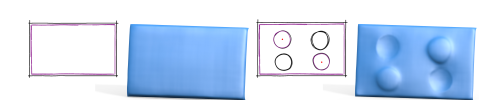


图 7. 一个简单的抽象形状在绘制过程中。 要创建这个凹凸平面，用户可以简单地首先使用轮廓和尖锐特征绘制一个平面，然后通过在内部绘制圆圈并使用深度样本（红点）将孔推到平面下方来添加凹凸

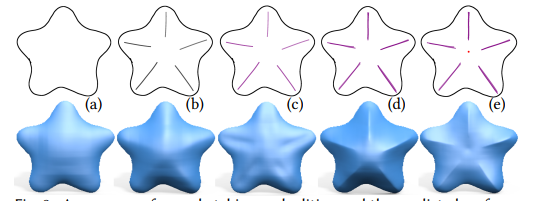


图 8. 一系列用户草图和编辑以及预测的表面。 (a) 仅轮廓及其表面。 (b) 画了五个内笔画。 (c) 内部笔划被赋予带有负号的曲率提示。 (d) 内部笔画被尖锐的特征所取代。 (e) 将内部深度样本（红点）放置在中心并将区域拉平。

通过对像素网格进行三角剖分来构建表面块，并将其提升到预测的深度值并使用预测的法向量进行渲染。 对于单曲面建模，除非另有说明，否则在绘图平面中默认为轮廓曲线提供深度样本。 重新绘制的轮廓曲线用红色标记； 内部深度样本（如果有）也显示为红色点。指定为锐利特征或具有目标曲率值的笔划以紫色标记。

为了更好地理解稳健数据拟合损失的行为，我们用或多或少的模糊草图可视化了代表性示例的置信图（图 9）。 正如我们所料，简单的草图或使用深度提示和曲率提示会导致低歧义并提高预测的置信度值。 另一方面，相对杂乱的过度草绘和复杂笔划，以及存在未确定表面法线跳跃的尖锐特征，可能会更加模糊并降低置信度值。

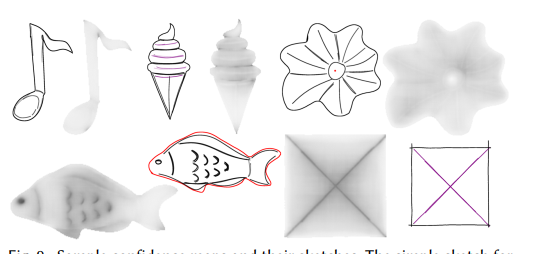


图 9. 样本置信度图及其草图。 音符的简单草图，冰淇淋的带有曲率提示的笔画（紫色）以及弯曲盘子的深度样本（红点）都可以减少歧义和更明亮的置信度图。 相比之下，像鱼中的过度绘制和杂乱的局部细节以及金字塔的未知法线跳跃的尖锐特征（紫色）具有更多的模糊性和更暗的置信度图。 结果 3D 形状如图 1 和图 3 所示。 1、11、15

运行。 CNN 前向传递的单视图表面预测是瞬时的（大约 42 毫秒），这允许交互式反馈和平滑的草图探索。 由于所有有效像素的大量变量，使用多视图融合时相对较慢，但即使对于多达五个表面补丁，也只需不到 10 秒。 泊松重建通常需要不到 3 秒的时间从点集中提取表面网格。 该时间是在配备 Intel Xeon 2.0GHz CPU 和 Nvidia GeForce GTX 980 GPU 的台式 PC 上报告的

## 5.2 Comparison

基于非学习的方法。对于许多经典方法，例如 [Igarashi et al. 1999；尼伦等人。 2007]，建模过程是一个交互式和增量会话，其中每个单笔画定义一个新的 3D 边界或特征曲线，通过求解预定义的几何平滑度先验（如双调和方程）来修改当前的基本形状。相比之下，对于每个视图，我们的方法都将 2D 草图作为输入，并在一次通过中生成一个详细的 3D 形状，该形状遵循学习的几何先验。一方面，建模范式的这种差异使我们的方法成为对先前方法的补充，因为它可以通过在单个视图中进行草图绘制来进行有效的表面补丁建模。另一方面，学习的几何先验可能比手工规则更合理。在图 10 中，我们看到要对瓶子形状进行建模，使用我们的方法将 2D 草图直接转换为 3D 表面 (d)，而对于具有边界位置和法线约束的双调和表面，结果形状失去了我们的关键特征想要 (b)，并且即使在提供额外的内部曲线约束 (c) 时也过度圆整。

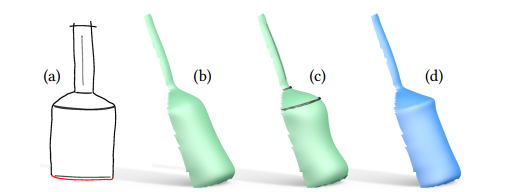


图 10. 与双谐波表面的比较。 (a) 瓶子形状的输入草图。 轮廓的红色部分被重新绘制成半圆。 (d) 我们的预测结果。 (b) 使用取自 (d) 的边界位置和法线约束，计算出的双谐波表面不能保留瓶子形状的关键特征。 (c) 即使我们用从 (d) 中提取的提升空间笔划（黑色曲线）进一步约束双调和表面，它也过于圆润并失去了所需的特性。

BendSketch [李等人。 2017] 从 2D 草图中重建具有复杂曲率变化模式的自由曲面，但每个笔划的类型和含义必须由用户指定作为输入，以便可以正确应用预定义的几何规则。 相比之下，由于我们的数据驱动方法，我们的方法用更简洁的笔触和很少的注释来解析用户的手绘草图。 示例如图 11 所示，其中两种方法都使用非常不同的草图创建了相似的表面补丁。

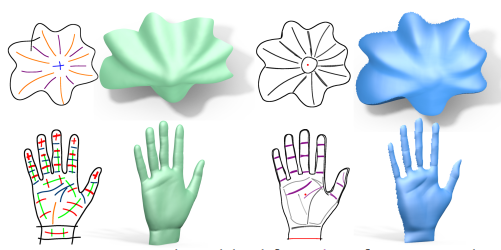


图 11. 与 BendSketch 的比较 [Li et al. 2017]。 我们重新创建了与他们的论文（左）相似的形状（右）。 虽然他们的方法需要关于更多笔划类型的详细注释（脊、谷、曲率线、尖锐特征等，通过颜色显示）和调整笔划曲率大小（通过线宽显示），但我们的方法使用更简洁和 自然灰度草图，几乎没有其他规格，例如深度样本（红点）或清晰的特征（以紫色标记）。 对于我们的手部模型，红色轮廓被重新绘制为稍微偏离平面。

基于学习的方法。 如第二节所述。 2，大多数现有的基于学习的方法对特定类别的形状进行建模，而我们的方法针对的是通用的自由形状。 尽管如此，我们还是在这里与多视图解码器方法进行比较 [Lun et al. 2017]，它输出要融合到 3D 形状中的预定义视图的中间深度和法线图，对它们的形状类别之一进行建模，字符，以展示不同网络结构对结果质量的影响。

为了进行有意义的比较，在 [Lun et al. 2017]，我们使用前视图中的草图作为唯一的输入。对于他们的多视图解码器网络，输入是前视图草图图像，输出是相同数量的视图的 13 个深度和法线图，包括一个正面视图和对象周围的 12 个定期采样的球面视图。他们的网络经过训练以最小化数据拟合损失和捕获特定类别属性的对抗性损失。为了统一，我们从他们的网络中删除了掩码回归目标。对于我们的网络，输入的前视图草图首先通过第一阶段流场回归网络，然后草图和流场图一起通过第二阶段几何网络输出前视图的深度和法线图。注意到数据集中的许多 3D 模型是缺乏平滑度的低多边形形状，我们将规则权重放宽到 λ = 0.01（第 3.3 节）以训练我们的网络；由于训练集中没有笔画类型数据，我们使用保守的信号过滤掩码 {αi}，对于所有笔画像素和背景，该掩码为零。使用训练数据中的深度图生成地面实况曲率方向场。我们用他们网络中使用的较小的编码器-解码器结构替换了我们的基本 UNet 组件，该结构适合具有 10k 个样本的小型数据集；结果，我们的网络比他们的网络小两倍多。所有网络都经过 10 个 epoch 的训练以收敛

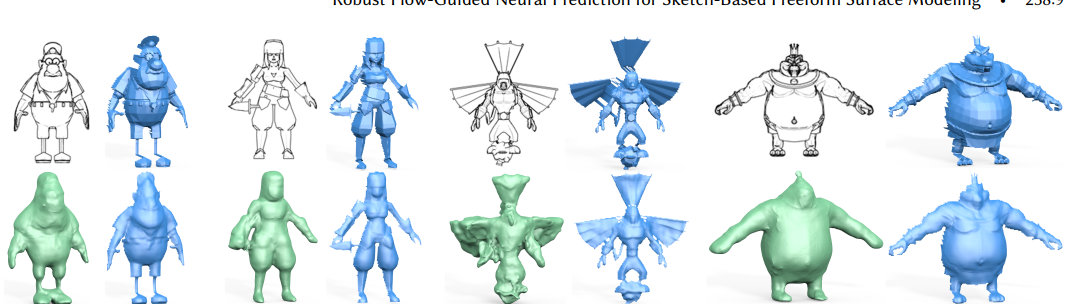


图 12. 与多视图解码器网络的比较 [Lun et al. 2017]，两个网络都在字符数据集上进行了训练和测试。 每个测试用例的输入草图显示在左上角，地面实况前视图深度图显示在右上角。 [Lun 等人的多视图融合了完整的 3D 形状。 2017]以绿色显示在左下角，而我们预测的通过后正则化改进的前视图表面在右下角以蓝色显示。 虽然我们的结果不是完整的 3D 模型，但它们捕捉到了草图所预期的更多关键细节。

他们结果的深度/法线误差为 0.0294/22.4°，我们的为 0.0332/18.6°。 视觉结果（图 12）显示我们的网络预测的形状可以更好地捕捉输入草图传达的重要细节。 它表明，我们的具有歧义估计的两阶段网络即使对于从草图进行特定类别的重建也是有益的。

## 5.3 User evaluation

我们邀请了五位对素描和 3D 建模知识或培训知之甚少的新手用户来评估我们的工具。 我们首先向参与者介绍线条草图如何描绘 3D 形状的背景，然后通过绘制简单示例向他们展示我们的工具是如何工作的，让他们熟悉用户界面。 培训时间为每人 20 分钟。 然后在评估测试中，要求用户完成三个建模任务，他们创建三个不同复杂度的目标形状。 目标形状，即鸟、泰迪熊和海豚，由参考图像或在另一个程序中渲染的 3D 模型给出

参与者创建的示例草图和 3D 形状如图 13 所示； 更多的用户创作和评价在补充材料中。 用户以不同的风格进行素描，这对他们来说很自然。 他们分别在 5 分钟、15 分钟和 20 分钟内为鸟类、泰迪熊和海豚创造了有趣的形状。 鸟类模型全部绘制在一个视图中，并通过前表面的对称反射完成。 泰迪模型在前视图和后视图中绘制，其中前表面轮廓被复制到后面以获得帮助。 海豚模型以俯视图、仰视图和侧视图绘制； 侧视图仅包含一个三角形草图，对顶部鳍片进行建模，与其他两个补丁融合以完成形状

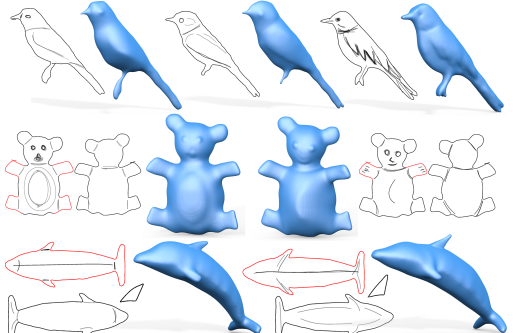


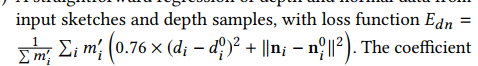
图 13 用户为评估测试中的目标形状绘制的各种样式的草图。 重新绘制红色轮廓以抬起泰迪四肢或弯曲海豚身体。 鸟的形状由一个正面草图制成，并由对称的背面完成。 泰迪熊由正面和背面草图的表面制成。 海豚在三个视图中绘制； 三角形草图为顶部鳍片建模。 平均而言，没有绘图或建模背景的新手用户分别在不到 5 分钟、15 分钟和 20 分钟内完成了鸟、泰迪和海豚形状的制作。

参与者的反馈大多是积极的。 经过 20 分钟的训练后，用户发现草图很容易理解，用于描绘形状。 在他们学会使用该工具后，他们喜欢花时间探索变化。 此外，用户可以在多个视图中找到草图，以使完整的形状易于使用。

## 5.4 Ablation study

没搞懂这部分干嘛的

为了进一步测试我们的网络结构的影响，我们对替代结构做了三个额外的实验

1. 来自输入草图和深度样本的深度和法线数据的直接回归，带有损失函数

深度拟合项的系数 0.76 来源于深度差 [0, 2.6] 的范围和法向量差的范数范围 [0, 2]

1. 修正回归，损失测量形状规则和深度样本约束，即 Edn + βEsample + λEr eä。
2. 具有置信度图和损失函数 Etotal 的进一步修改回归，但没有第一阶段网络或中间流场

这些网络具有与 GeomNet 相同的结构，即一个五级编码器-解码器，具有用于不同输出的单独解码器。 他们在与之前介绍的相同数据集上训练了 10 个 epoch 以收敛（第 3.4 节）

训练好的网络在测试数据集上进行评估。 代表性的视觉结果如图 14 所示。平均测试误差列于表中。 如图 1 所示，在我们对齐预测深度图和真实深度图以使它们具有共同的平均值之后，再次计算深度误差。 法线误差是从深度图和地面实况导出的法线向量之间测量的角度，以度为单位。 除了上述网络配置之外，我们还应用了拉普拉斯平滑 [Desbrun et al. 1999]（λ = 1, 5次迭代）对（1）的预测结果，检查这个简单的后处理是否可以提高结果质量

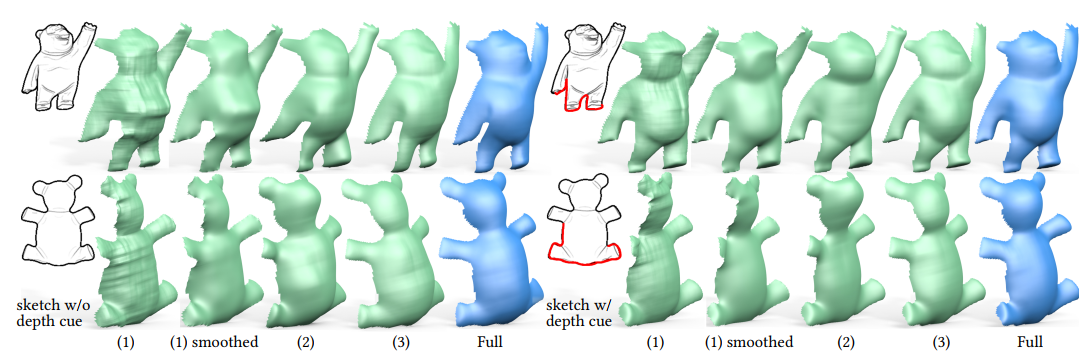


图 14. 用于消融测试的不同网络预测的表面。左列显示没有任何深度样本的输入草图的结果，右列显示沿部分轮廓曲线的深度样本（标记为红色）。草图是从现有的测试模型中呈现的。 (1) 是深度和法线的朴素回归。 (1) smoothed 将拉普拉斯平滑应用于网络预测。 (2) 在损失函数中增加 (1) 的规律性和深度样本约束。 (3) 进一步使用带有置信度图的鲁棒数据拟合。全网具有完整的两级结构。 (1) 产生非常嘈杂的结果，在拉普拉斯平滑后仍然失真。 (2) 降低噪音，但整体形状失真。 (3) 预测更规则的表面。但是完整的网络进一步提高了规律性，并且具有更丰富的表面变化，忠实地捕捉输入笔画的含义（例如，参见顶部熊的头部）。在给定深度线索时，所有网络都表现得更好，但完整的网络在缺乏深度样本的情况下最有弹性。

所有配置的结果都符合我们的预期。朴素拟合网络 (1) 产生的结果深度误差小但法线误差大，这在视觉上转化为全局不规则和局部噪声表面（图 14）。拉普拉斯平滑可以减少噪音但不会改善整体形状。与深度/法线一致性的附加目标相比，（2）法线误差大大减少，但深度误差比（1）增加；从视觉上看，结果形状是局部平滑的，但整体是扭曲的。具有置信度图建模模糊性的回归 (3) 不仅大大提高了表面规律性，而且在深度和法线方面的误差都比 (2) 低。具有流场引导的完整网络仍然具有几乎等于朴素预测的深度误差，但正态误差要小得多。从视觉上看，如图 14 所示，全网络的结果表面更加规则，丰富的变化忠实地捕捉了草图的意图。此外，虽然所有网络在给定更多深度线索时都会产生更好的结果，但完整的网络最具弹性

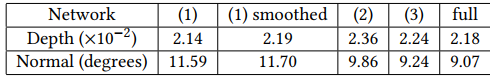


表 1. 比较网络的平均测试错误。 (1) 是深度和法线的朴素回归。 (1) smoothed 将拉普拉斯平滑应用于 (1) 的结果。 (2) 在损失函数中增加 (1) 的规则和深度约束。 (3) 使用与我们相同的损失函数和置信度图。 Full 是完整的两阶段网络。

由于缺乏深度线索，我们认为这是由于调节表面预测的密集流场。

# 6 CONCLUSION

在本文中，我们提出了一种基于草图的自由曲面建模方法，该方法使用 CNN 模型从稀疏的 2D 草图深度和表示表面的法线图推断。 与传统方法相比，新的基于学习的方法可以直接从 2D 草图中推断出更简洁的线条和更少的用户注释。 与现有的基于学习的方法仅通过建模类别特定的形状来解决 2D 到 3D 的模糊性不同，我们的方法针对通用的自由形状，并通过使用在我们的新型网络结构中实现的基本几何和统计规则来处理模糊性。

我们的网络有两个阶段：在第一阶段，网络从草图中回归密集流场，从而规范随后的几何重建； 在第二阶段，给定输入草图和流场，网络预测深度图和法线图，以及量化每个点的模糊度的置信度图，这是在无监督的情况下进行训练并使几何回归更加稳健。 我们的网络还允许用户通过提供深度采样点、锐利特征和曲率提示来修改表面。 为了训练网络，我们通过 NPR 渲染生成了一个数据集，并密切模仿用户实际绘制草图以描绘 3D 形状的方式。 最后，多视图草图提供了方便的工具，例如重新绘制轮廓，并通过融合对象不同部分的表面来建模完整的 3D 形状。

我们进行了验证、比较、用户评估和消融测试，表明所提出的方法是一种通过绘制直观且富有表现力的稀疏 2D 草图来建模各种自由形状的新方法。

**限制和未来的工作**。 我们的方法每次都对表面贴片进行建模，这是一种不适合高度结构化和对称形状（如 CAD 模型）的范例，通过绘制完整和规则的线框可以更方便地定义这些形状。

我们使用了固定分辨率 256×256 的绘图画布。 对于复杂的形状，很自然地以多种分辨率对不同的细节层次进行建模。 我们希望扩展我们的方法和用户界面，以实现这样一个多尺度框架，以灵活地对整体形状和精细细节进行建模。

在这项工作中，我们考虑了具有稀疏线条的草图作为输入。 然而，描绘阴影线索的密集笔触也被艺术家广泛用于增强 3D 感知。 将来，我们也希望为 3D 建模处理这些草图样式

解决歧义是数据驱动和机器学习任务中普遍存在的问题，我们相信鲁棒估计器的应用通常会有所帮助

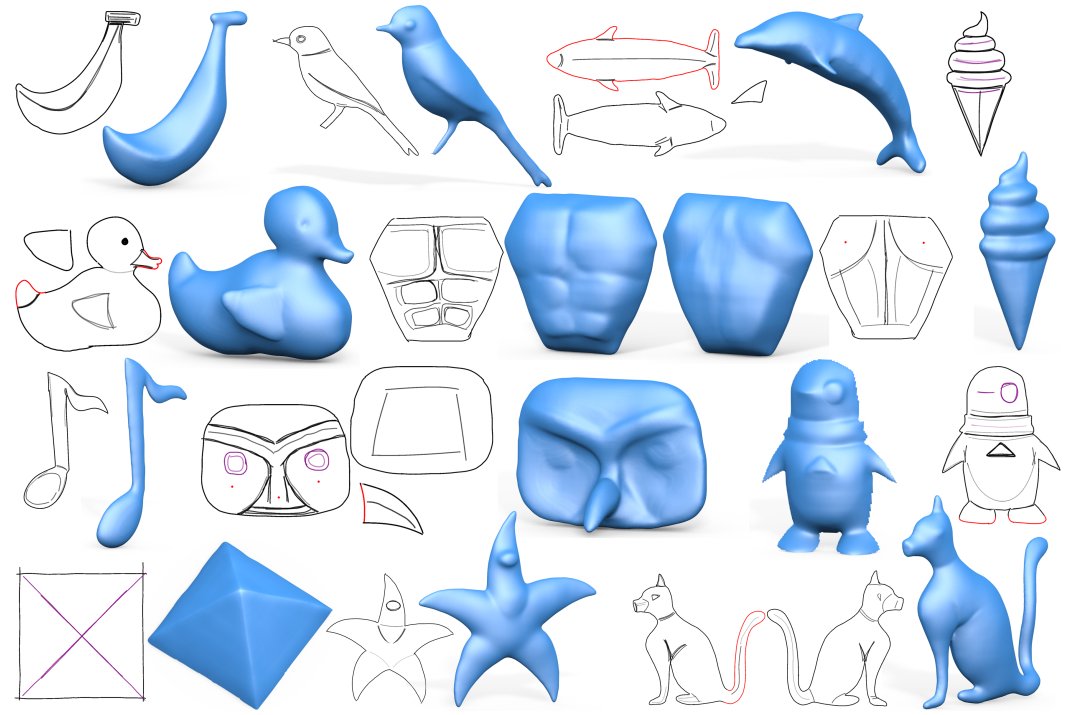


图 15. 更多草图和相应的形状。 重新绘制红色轮廓以提供额外的边界深度数据。 内部深度样本（红点）用于躯干背部和猫头鹰面部。 棱角分明的特征（紫色）用于金字塔、企鹅、猫头鹰的脸。 曲率提示用于冰淇淋。 海豚鳍、鸭翼和猫头鹰鼻子在各自的视图中由小三角形补丁建模，然后与其他部分融合。 企鹅和金字塔是单一的补丁； 其余的都是完整的形状。