

指导教师： 杨涛

提交时间： 2015.3.28

The task of
Digital Image Processing

数字图像处理

School of Computer Science

No: 1

姓名： 丁宇

学号： 2012302421

班号： 10011202



贝叶斯视图合成和基于图像的绘制原则

摘要

在本文中，我们致力于解决关于从一组输入图像中如何综合出新颖视图的方法问题。一些最先进的方法，例如非结构化的卢米图，已经使用了启发式信息来结合原有的图像，它们经常利用了场景的几何形状的显式或隐式的近似值。同时已经提出了的启发式方法已经在很大程度上被探索和证明以至于让工作更加有效率，最近推出的贝叶斯猜想，让一些先前提出的启发式算法更加规范化，指出了其中的物理现象可能恰恰藏在这些图像之后。然而，一些重要的启发性问题还没有被带入考虑并且缺乏合适的规范化。我们提供了一个全新的基于物理的生产模式和与其相适应的最大后验概率的估计，提供以启发式为基础的方法和贝叶斯公式之间所需的统一。关键的一点是要系统地考虑由几何代理的不确定性所引起的错误。我们提供了一个广泛的讨论，分析所得到的方程式怎样来解释启发式在以前的方法中的发展。此外，我们表明我们的新型的贝叶斯模型显著提高了新的视图的质量，特别是当场景几何估计是不准确的时候。

1.简介

我们解决新的视图合成的问题是基于图像呈现的领域（简称 IBR），目的是利用一系列任意搭配输入图像，从不同的观点中综合出新的观点。大部分最先进的方法都是运用启发式的状态的方法来定义能量或目标功能最小化，实现了优异的成绩。在 IBR 一个重大突破是 Buehler 等人的鼓舞人心的工作。他们定义的“理想性能”任何 IBR 算法

应该有。这些指令仍然在当前最先进的方法中占据优势。

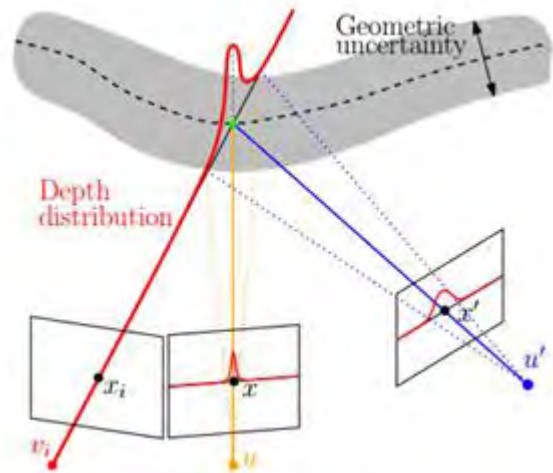


图 1: 沿相机 V_i 的射线光的深度分布将不同地传播，具体取决于呈现相机 u 或 u' 的视角，该角度越大，预计的不确定性就越大。

然而，最近，对贝叶斯形式主义的使用在 IBR 的技术已被引入，就在 Wanner 和 Goldluecke 提出的工作里。他们为新颖视图的合成提供了第一贝叶斯框架，利用一个基于物理学的生成模型描述图像形成过程，并取得其最大后验（MAP）估计。此外，它们的变分方法不仅解决新视图合成的问题。它直接解决新的超分辨率图像的合成，并提供给其他相关的问题，即图像去噪或图像去模糊一个坚实的框架。

有趣的是，虽然[4]和[28]已经解决了同样的问题，他们的理论结果不收敛成一个统一的框架。一方面，由 Buehler 等人在[4]中提出的口授的准则已经证明是非常有效的，但缺乏一个正规推理来支撑它们。此外，目前还不清楚应如何处理一些理想的性质之间的平衡。一个说明性的例子是极点一致性及分辨率灵敏度之间的折衷。前者指出，“当期

望的光线通过源照相机的投影中心时可以细微地修复“，而后者指出，”在现实中，图像像素不是单个射线真正的单位，而是对着一个小的立体角的一组光线的集合。这个角度范围应由理想的呈现算法来解释。“该极点一致性与角度偏差期有关联，而分辨率灵敏度是由有关的雅各比平面单列矩阵导出。这两种启发式方法似乎都是合理的，但谁应该占主导地位呢？属性之间的权重选择是由用户调谐，同时在他们的实验中，参数必须由不同场景做出不同调整。

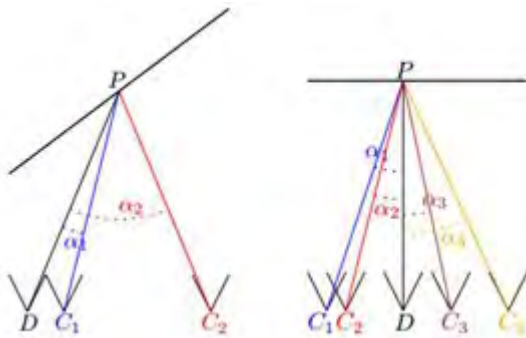


图 2：视线 D 是从摄像头 C_i 通过使用[28]产生的。左：

由于透视效果的作用，相机 C_2 将比相机 C_1 更受到青睐。然而，D 和 C_1 之间观看的光线角距离是比 D 和 C_2 之间小得多。右：配置了一个水平面。所有摄像机将具有相同贡献，尽管有不同的视角。

另一方面，现有的贝叶斯模型[28]是能够解释一些启发式方法的，但仍然违反了那些，似乎明显并且已被证明有效的方法。例如，我们在能源中的场景几何发现了一个分析推演的透视效果的影响。调查结果在[4]中证实了由 Buehler 等人所提出的启发式：正是雅可比变换的推导与视野有关。然而，当仔细分析在[28]中的最后的方程，一个在[4]中提出的重要的期望

特性，至今下落不明：视线的最小角度偏差并不执行此特性，甚至在某些情况下违反，如图 2 所示。

最新技术的生产模型和被普遍接受的启发式方法提出的能量之间的区别是什么促使了现在的工作。我们的目标是保留从贝叶斯

形式主义所产生的固有的无参数能的优点，同时推动了[28]的图像形成模型的界限，并提供一个新的模型，其能够解释在 IBR 中大部分目前公认的在本领域中最先进的直觉。

作用：这个所提出的方法的核心理论贡献就是通过几何代理估计误差的郎伯图像合成工序里的错误的系统模型化。我们把这个误差叫做深度不确定，由输入图像的深度估计得到。除了这个错误以外，我们也按照普遍的高斯建模思考这个图像传感干扰。我们广泛地分析获得的能量的理论含义，同时讨论我们建模中最先进的启发式方法的形式推理。这项工作明确地为第一贝叶斯构想导出了[4]的启发法。

从具体实际的角度来说，我们通过和采用贝叶斯体系的最优存在法进行对比，在数字方面评估了我们方法的性能。实验数据表明，关于大众数据的客观度量，我们完成了最新的数据。此外，我们也能处理超分辨率，利用[28]中建立的体系结构。这个新的模型并不是不需要任何代价，因为他的最优法并不直观。然而，现有的方法能让我们去客服困难。公开的源代码可在此网站获得：<http://sf.net/projects/cocolib/>。

<http://sf.net/projects/cocolib/>。

2. 相关工作

自从 McMillan 和 Bishop 在全景模型[14]中做出了早期工作后，许多 IBR 的技术都因为某些目的得到了发展，例如，自由视点呈现[24]，图像变形[30]或图像视图插补[21]。由 Shum 等人[19]做的分类表明许多 IBR 方法使用了一个几何代理，并且他们做了一个 IBR 连续性的分类，分类方法取决于他们用了多少几何。在一个连续的一端，我们有一些方法，并不用什么东西，只需收集大量的输入图像，例如光场呈现[13]，同心交织图案[20]。在相反的另一端，我们将呈现技术依赖于直观的几何，使用精确的几何模型但

很少的图像，例如深层图像[17, 6]和依靠视觉的手感绘图[9]。在这其中，我们通过使用一个几何的内在描述找到了一些方法，例如依赖光电流的视图插补技术[7, 27]，使用极几何沿视线建立连接的传送方法[12]，还有卢米图[11]，使用了大致明确的几何学和相对大量的图像。

当Buehler等人引进非结构化卢米图呈现时[4]，他们建立的七个“理想属性”，所有的IBR方法都应该符合：几何代理的使用，非结构化的输入，极点的一致性，最小角度偏差，连续性，分辨率灵敏度，相当于光一致性和实时性。这项工作一直在社会上非常重要，大多数IBR方法也是遵循这些原则。虽然贝叶斯形式主义是一种常见的方式去处理空间的超分辨率在多视点以及轻现场设置[3,10]，他们最近才因为华纳和Goldluecke[28]的工作而被运用到IBR。虽然他们的工作为分辨率灵敏度特性提供了一个物理解释，最小角度偏差也可以在最后的方程中使用。最有趣的是，Vangorp等人[26]试验检验这些IBR方法的属性很容易创造视觉假象，并他们的主要成果之一是把识别角度偏差作为重点属性来加以考虑，以避免视觉假影。

即使通过现有最先进的3D重建技术在估算几何代理方面也是惊人的，考虑到他们是完美的，似乎是一个强大的假设：即使是最好的方法在最后的估计中有不确定性。当然，新的观点合成容易因为一个差的（隐性或显性）重建而在某些区域产生视觉假象。其中一个解决这个问题的方法就是提高采集设置，由Zitnick等人完成[31]。他们实现一个足够好的重建，促进了令人印象深刻的新观点的合成。然而，他们的设置在很大程度上被制约。

在[22]中，当从2个图像中呈现一个新的视角时，Takahashi在几何代理呈现中研究误差的理论影响。我们通过解决照相机的配置加强了[22]，并提供一种有效的方法来找到解决方案，都明确地留作今后的工作。在[23]中，Takahashi和Naemura使用深度的不确定性的信息来影响正则项（前）。但是，这

个想法确实还没有考虑到最小角偏差，因为每个摄像头不允许有不同的贡献量。我们在这项工作中解决这个问题。

3.新型视图综合生成模型

我们的目标是从一个新的视点 c 合成一个（可能是超分辨率）图像 $U: \Gamma \rightarrow \mathbb{R}$ ，使用一组从任一点 c_i 捕获的图片 $V_i: \Omega_i \rightarrow \mathbb{R}$ 。我们假设我们有一个几何代理的估计，其中足以建立视图之间的对应关系。更正式地说，这个几何代理包含一个向后的变形标记 $\tau_i: \Omega_i \rightarrow \Gamma$ ，从每个输入图像到新的视点，以及二进制闭塞掩码 $m_i: \Omega_i \rightarrow \{0,1\}$ ，它选择最有价值的点当且仅当 Ω_i 中的点在 Γ 中可视。如果我们限制 τ_i 到设定的可见点 $V_i \subset \Omega_i$ ，它是单射并且它的左逆 $\beta_i: \tau_i(V_i) \rightarrow \Omega_i$ 被很好地定义，见图3。

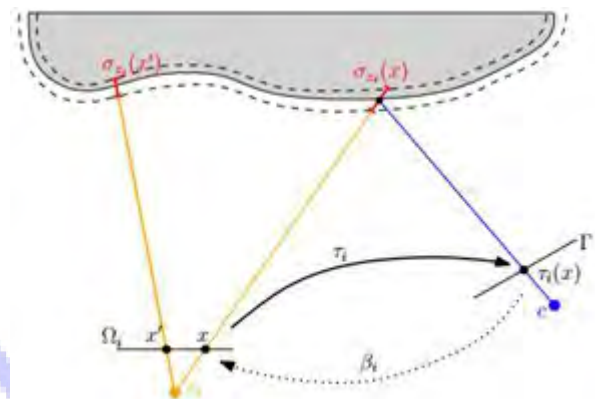


图3 从图像平面 Ω_i 转移 τ_i 到目标图像平面 Γ 。深度不确定性 σ_z 可能在像素点间不同。

理想的成像模式。为了考虑从超分辨率新视图输入新视图的分辨率损失，我们通过在VI的图像形成过程中应用模糊内核中的 b 模拟二次抽样过程。它对应于摄像机 i 的点扩散函数（PSF）。 v_i 的各像素来自场景的光线的集合存储积分强度，并且这个新的点 u 将总被认为比输入的图像有更高的分辨率。

我们暂时抛开可视度的效果，假设所有点都可见。还假设我们有一个完美的向后的伸缩

映射 τ_i^* , 从 Ω_i to Γ , 也有一个完善的输入影像 V_i^* 。假设朗伯成像模型, 新视图和输入视图之间的理想化确切的关系是

$$v_i^* = b * (u \circ \tau_i^*), \quad (1)$$

存在关系 \circ 的函数组合操作。然而, 该观察到的图像 v_i 和几何 τ_i 是不完美的, 而我们需要考虑这些因素的成像模型。

传感器误差和图像错误。首先, 我们考虑用方差 σ_s^2 在所有摄像机上用高斯传感器噪声, 而传感器噪声方差 σ_s^2 和二次取样内核 B 可能是图像之间不同, 为符号的简单起见, 我们将假定它们是相同的。

$$\sigma_{g_i} = \sigma_{z_i} \left| b * \frac{\partial(u \circ \tau_i)}{\partial z_i} \right| = \sigma_{z_i} \left| b * \left((\nabla u \circ \tau_i) \cdot \frac{\partial \tau_i}{\partial z_i} \right) \right|.$$

其次, 我们认为错误在几何估计, 这意味着相应的向后伸缩 τ_i 是与理想的映射 τ^* 不同。这在图像形成过程中包含一个强度错误 e_{g_i} ,

$$\epsilon_{g_i} = b * (u \circ \tau_i^*) - b * (u \circ \tau_i) \quad (2)$$

有关强度的错误的 uncertainty e_{g_i} 表示为 σ_{g_i} : $\Omega_i \rightarrow \mathbb{R}$ 。需要注意的是都具有强度单位。

考虑到上述错误, 图像形成模型变为:

$$v_i = b * (u \circ \tau_i) + \epsilon_{g_i} + \epsilon_s \quad (3)$$

虽然我们做出了共同的假设, 即 ϵ_s 遵循高斯分布, e_{g_i} 的分布对于我们是未知的。我们知道是 e_{g_i} 是与几何误差密切相关的。在下一节中, 我们研究他们分布之间的关系。

图像误差对几何误差的依赖。该几何代理为每个在 Ω_i 的像素 x 产生一个深度检验 z_i , 其与不确定性 σ_{z_i} 关联, 给了我们一个沿着视线的深度分布, 如图 3。我们现在考虑几何代理的估计的错误 e_{z_i} , 用世界单位表示。以前的图像错误 e_{g_i} 是依赖于底层的几何误差。注意, 图像误差具有强度单位不得与有几何单位的 e_{z_i} 混淆。与此相反的模糊核和传感器噪声, 我们允许这些不同错误的存在, 是对于每个视图, 并为每个视图的不同

像素, 如让符号更加详细。

我们假设, 对于深度估算的误差分布是正常的,

$\epsilon_{z_i} \sim \mathcal{N}(0, \sigma_{z_i}^2)$ 。现在的目标是

导出这个分布如何生成在图像形成处理的颜色误差分布。传播具有任意功能的分布并不简单, 即使如果在我们的情况下, 该深度误差分布被假定为高斯, 并且将仅沿着该视图之间的几何形状来传播。相比于计算沿着视线的全色分布, 我们线性化考虑 v_i 到 ZI 的一阶泰勒展开式的。这意味着, 所得到的颜色分布也是高斯, 平均 $u \circ \tau_i$ 和标准偏差

$$\sigma_{g_i} = \sigma_{z_i} \left| \frac{\partial v_i}{\partial z_i} \right|. \quad (4)$$

运用式子 (1)。连锁规则, 我们发现

$$\sigma_{g_i} = \sigma_{z_i} \left| b * \frac{\partial(u \circ \tau_i)}{\partial z_i} \right| = \sigma_{z_i} \left| b * \left((\nabla u \circ \tau_i) \cdot \frac{\partial \tau_i}{\partial z_i} \right) \right|. \quad (5)$$

MAP 估计和能量。在贝叶斯法则中, 新视图的 MAP 估计可以发现, 因为图像 \hat{u} 使能量最小化

$$E(u) = E_{\text{data}}(u) + \lambda E_{\text{prior}}(u), \quad (6)$$

其中数据术语 $E_{\text{DATA}}(U)$ 可从生成模型推导出, 并且 $E_{\text{prior}}(u)$ 作为一个平滑项, 随后详细介绍。我们方法里 $\lambda > 0$ 是唯一的参数, 并且它控制该结果的平滑度。

让我们考虑两个误差源为独立的, 外加的和高斯的。然后它们的和也是正态分布具有零

均值和方差 $\sigma_s^2 + \sigma_{g_i}^2$ 。这个数据从式 (3)

的生成模型计算 由下式给出:

$$(u) = \sum_{i=1}^2 \frac{1}{2} \int_{\Omega_i} \omega_i(u) m_i (b * (u \circ \tau_i) - v_i)^2 dx, \quad (7)$$

和

$$\omega_i(u) = (\sigma_s^2 + \sigma_{g_i}^2)^{-1} \quad (8)$$

该数据项是一个类似于在先前[28]的模型中发现的那个, 除了因子 $\omega_i(U)$, 它可以看作是取决于对深度的不确定性和潜在图像 U 的量来计算。如果没有深度的不确定性,

这个项就会降低至 σ_s^2 , 这恰好给出在[28]

中找到的能量。

从公式 (5) 中, 我们可以观察到如果向量

$\partial \tau_i / \partial z_i$ 的长度减小, 在 ω_i 中的术语 $\sigma_{g_i}^2$ 也在变小。 $\partial \tau_i / \partial z_i$ 表示当其深度 z_i (x_i) 变化时, 从原来的图 V_i 中的点 x_i 在新图 u 的投影变化量。这个向量点朝向 v_i 上的点 x_i 发出的极线的方向, 其幅度减小与来自新视图 u 和从原始视图 v_i 发出的光射线之间的角度有关。如图 1 表示, 术语 $\sigma_{g_i}^2$ 因而解释了 [4] 中极小的角度偏差“理想的属性”, 它未解释 [28] 中的原因。

让我们来更精确地分析在何种情况下的权重 $\omega_i(u)$ 达到其最大值 $1/\sigma_s^2$, 这是在以前的模型中找到的值。发生这种情况有三种可能。第一个是, 如果 $\partial \tau_i / \partial z_i = 0$, 即在 v_i 中某点的深度对它再投影到 \tilde{u} 没有影响。这只能发生在两个光射线是相同的情况下, 其对应于 [4] 中

的极点一致性特征。第二个是如果 $\nabla u = 0$, 即呈现图像在所考虑的点不具有梯度或纹理: 在此情况下, 在深度估计上的错误对呈现视图没有影响。最后的情况是, 如果 ∇u 在呈现点正交于从摄像机 I 穿过极线的方向呈现点: 摄像机 I 的深度估计中的一个小错误不会对呈现视图有影响, 因为这种错误影响的方向相切一个轮廓。

选择先验。 这个最优的方法在贝叶斯方法中引入来抑制对象图像潜在的配置。通常, 它被用来克服不稳定性问题: 在 Baker 和 Kanade [1] 的超分辨率分析中, 它们表明矩阵系统的零空间的维数随超分辨率因子的增加而增加。此外, 在新的视图合成中, 图像的某些部分可能不被任何贡献视图看出, 因此一个正规化允许以合理的的信息填补差距。因此, 现有的选择将对最终结果有显著的影响。

为了克服在超分辨率 [18] 的具体问题, 非常有趣的先验已经开始开发了。同时, 也有让从图像的集合 [16] 中学习通用图像先验的方法。然而, 在这项工作中, 我们重点在生成

模型, 我们使用基本的总变化作为一个正则化,

$$E_{\text{prior}}(u) = \int_{\Gamma} |Du|, \quad (9)$$

这是凸面的, 并在图像分析的问题 [5] 的文章中已被广泛地研究了。寻找最佳先验将是未来工作的一个主题。

最优化。 公式 (6) 中的结果很难得到最优因为 $\omega_i(u)$ 在公式 (7) 中的影响是在潜在图像 u 中的非线性结构。与 [8] 类似, 我们提出了重新加权迭代方法。我们使用 U 的估算, 在第一次迭代中设 $\tilde{u} = \frac{1}{n} \sum v_i \circ \beta_i$ 我们认为, $\omega_i(\tilde{u})$ 的每次迭代期间恒定, 使得简化能量得以突出。此外, 与 [28] 类似的争论, 我们可以证明功能衍生简化数据项是

$$dE_{\text{data}}^i(u) = \omega_i(\tilde{u}) |\det D\beta_i| (m_i \bar{b} * (b * (u \circ \tau_i) - v_i)) \circ \beta_i \quad (10)$$

在那里 $\bar{b}(x) = b(-x)$ 是共轭核心。它的功能导数是李普希茨连续的, 其允许经由快速迭代收缩和阈值算法 (FISTA) [2] 进行能量最小化。有了这个简单问题的解决方案, 我们更新 U , 从而获得新的权重, 和新的能量。我们再用 FISTA 解决它并且迭代。虽然在每次迭代中解决最小化问题是中凸的, 一般情况我们也不能希望找到式 (6) 的全局最小值。

4. IBR 法则的联系

正如我们在公式 (10) 中看到的, 对于每个视图加权因子都来自于两个方面。术语 $|\det D\beta_i|$ 和 [28] 中的是相同的, 并且对应于图像的变形的量度: 它是一个像素的表面从 u 投射至 v_i 。我们可以制定出它背后的意义: 当视点改变时, 观测的场景改变了多少?

术语 $\omega_i(u)$ 对应于深度不确定性, 正如在上一节中进行了说明。它背后的含义就是: 当观测的深度改变时, 观测的场景改变了多

少?

现在让我们认真地建立的提出的能源与[4]中IBR的“理想的特性”之间的联系。

利用几何代理和非结构化的输入。几何代理经由向后伸缩映射并入 τ_i ，输入可以是非结构化（即通用位置的随机的一组图像）。

极点一致性。如先前的解释，只要该光线从 x_i 发出并且 X 是相同的，加权因子 $\omega_i(u)$ 为最大，因此，如果照相机在 x 具有极点，则这款相机在 x ，通过 $\omega_i(U)$ 的贡献就能较高。极点一致性因此满足。

最小角度偏差。从方程(5)的 σ_{gi} 提供此启发式：如果保持所有其他方面不变（分辨率，到现场的距离等），然后方程(5)载体 $\partial \tau_i / \partial z_i$ 的幅度是完全成比例的角度，这个角度是从这两款相机的可见光射线到同一场景点之间。

分辨率灵敏度。此启发式由项 $|\det D \beta_i|$ 得来，它测量一个从 \tilde{U} 投影到 v_i 的像素的表面。我的相机的分辨率越大，这个表面也就越大，使分辨率的灵敏度得到妥善处理。

等价射线的一致性。“通过空间的任何空区域中，射线沿给定视线的应始终重建，不论视点位置（除非由其他目标规定...）”[4]。通过我们的框架，只是细微地感到满意，因为权重是根据相机参数连续地变化（通过向后伸缩映射 τ_i 的连续变化）。沿着光射线移动新视图摄像机（这是在[4]中用于描述该属性的情况）仅仅是一个特殊的情况。

连续性。IBR中的连续性原则要求最终呈现图像连续地具有不同于原来视图的相机参数。这意味着照相机之间不存在接缝处之间的可见性的边界，其可以发生在靠近视图中的场景中每个摄像机的视场的交叉点的边界，或在每个摄像机看到的深度不连续。典型的启发式去执行这种形式的连续性是在附近的可见性边界或其视场的边界[15, 4]降低一个摄像头的贡献率。

我们的方程式并不符合该属性，并且所获得的权重不降到零当接近可视边界的时候。这可能

轻松地被强制执行，通过使靠近深度的可视地图和可视不连续性平滑化，在不改变这些函数的零集的情况下。但是，由于我们宣称

拥有一个完全基于物理的贝叶斯公式，可视地图上的任何操作应有一个物理解释来确定，这方面我们仍然很迷失，而这是我们今后工作的一部分。

需要注意的是，在能量中的最优术语减少了这些问题，尤其在视觉伪像中，这是由于没有妥善处理连续性。然而，现有的新观点并不能完全解决连续性问题，这取决于场景和相机几何。

实时性。最后的“理想的属性”，是为了实现实时性。目前我们的方法还没有实时的，主要是由于MAP估计的计算复杂：从8个源图像中呈现出一768×768的图像，2到3秒是必要的。然而，无论是分辨率的算法和硬件体系结构都发展的很快，可以预计在未来的几年也会有更好的表现。

如果超分辨率并不重要，不是解决全图的问题，当使用实时正规化修补方法的形式来获得可接受的结果时，它似乎是合理的。

性能之间的平衡。关于[4]，我们方法的优点之一是不同特性之间的平衡不是由用户定义的参数来处理，而是从一个形式上推理暗示。想象一下，有两个摄像头的配置：一个配有低最小角距离但是高分辨率灵敏度变化，另一个配有高最小角距离但是低分辨率灵敏度变化。究竟哪一个会对最终的图像影响更大呢？在[4]中，角距离因分辨率灵敏度的比率 $1/0.05=20$

（走廊数据集）而优选。在我们的公式中，这些变化是完全基于物理的。图像之间的

$\Delta \alpha$ 的角度偏差按比例变化至 $\frac{1}{\sin^2 \Delta \alpha}$ ，

基于 σ_{gi}^2 的变化。一个透视效果或分辨率的差异引起的图像比例因子是按比例变化至

$\frac{1}{s^2}$ ，基于 $|\det D \beta_i|$ 的变化。这些因素之间的平衡被适当地处理，通过考虑传感器噪声

σ_s^2 。

一个例外是权重因子 λ ，在现有术语中使用。注意，这是基于贝叶斯原理的图像分析的所有工作的共同之处：因为目前还没有有

意义的方式来获得图像空间中的先验分布，需要用客观先验来进行正规化工作。当然，你也可以使用现有的方法[16]，允许直接从输入图像估计这个现有的东西，从而获得完全无参数的模型。

5.实验

简化的摄像头配置。虽然我们正在处理一般情况下的新视图的合成，为了简化最优化过程的执行，在实验中，我们假设我们的相机具有一个简化的配置。具体来说，所有的视点都在一个共同的平面，该平面平行于所有图像平面，也就是说，我们正在卢米图参数[11]中处理一个4D光场。新视图也合成在相同的图像平面上，这意味着， τ_i 是简单地给出，通过正比于归一化差异 d_i 的变换， $\tau_i(x) = x + d_i(x)(c - c_i)$ 。

(11)

标准化的差距体现在每个像素单元，并且连同与其相关联的不确定性，又与深度相关，通过：

$$d_i(x) = \frac{f_i}{z_i(x)} \text{ 和 } \sigma_{d_i}(x) = \sigma_{z_i}(x) \frac{f_i}{z_i(x)^2}$$

(12)，其中 f_i 是体现在像素中的摄像头焦点长度。

将(12)和(11)代入(5)，我们得出了几何误差及其相关联的图像误差之间的联系：

$$\sigma_{g_i} = \sigma_{d_i} | [b * ((\nabla u \circ \tau_i) \cdot (c - c_i))] |, (13)$$

其中 σ_{d_i} 表征噪音差异。最后，公式(10)中的变形术语是：

$$|\det D\beta_i| = |\det D\tau_i|^{-1} = |1 + \nabla d_i \cdot (c - c_i)|^{-1}. (14)$$

数据集。为了验证该理论的贡献，我们比较两个光场数据集的结果：在HCI光场数据库[29]，以及斯坦福大学光场存档[25]。这些数据集提供了一个广泛的集合，它们是挑战性的合成和真实世界的场景。

在第一组实验中，我们以相同的分辨率呈现

数据集中现有视图，而无需使用各自的视点作为算法的输入。我们考虑几何代理的两个不同的特质：来自估计的视差映射的近似值，以及

一个由场景预估中心的一个无限正面平行的平面表示的极小值。我们因此适应 σ_{d_i} ，也就是在使用估计的视差时，我们使用一个对应于重建方法的预期精确度的值：

$$\sigma_{d_i} = \frac{d_{\max} - d_{\min}}{\text{nbLayers}},$$

其中 nbLayers 是由该方法考虑的数目差值。当裸平面在场景的中部被使用时，我们改用

$$\sigma_{d_i} = \frac{d_{\max} - d_{\min}}{4}.$$

在所有的情况下， $\sigma_s = 1/255$ 。

第二组实验是进行呈现一个取自一组 5×5 输入图像中的 3×3 超分辨率的图像。虽然超分辨率还不是文章中的主要目的，我们还提供与现有技术状态的比较。因为超分辨率依赖于子像素的差异值，我们只显示包含估计的视差地图的结果。

在表1中，我们展示我们方法中获得的数值结果，并将它与[28]取得的成果相比较。我们测量实际的和生成的图像之间的PSNR和DSSIM。虽然我们的方法看起来表现地很好，但数值应要仔细解释。在图4中，我们展示了详细的特写来说明这些好处或者我们的方法。因为对于大多数数据集来说高分辨率图像是不可得的，超分辨率图像的PSNR和DSSIM值通过输入图像来计算，生成新的超分辨率视图并且将它与原来的相比较。

当呈现精确的几何图形时，这两种方法大致与PSNR和DSSIM的值相当（表1中的第一和最后两行）。当代理的质量下降时（表1的第三和第四行），我们的方法明显优于以往的工作，就是利用深度不确定性的显式建模。如在图4的图中所示，我们在所有的配置方法中更好地重新构建颜色边缘。全分辨率图像在补充材料中提供。

当呈现8幅输入图像的目标分辨率时的计算

时间大致为 2 到 3 秒。超分辨率视图合成，采用的是 3×3 和 24 幅输入图像，它的计算

时间约是 2 至 3 分钟。所有的实验中使用了 nVidia GTX Titan GPU。

	HCI light fields, raytraced				HCI light fields, gantry				Stanford light fields, gantry					
	still life		buddha		maria		couple		truck		gum nuts		tarot	
<i>Estimated disparity</i>														
Wanner et al. [28]	30.13	58	42.84	17	40.06	53	26.55	226	33.75	408	31.82	1439	28.71	60
Proposed	30.45	55	42.37	18	40.10	53	28.50	178	33.78	407	31.93	1437	28.88	58
<i>Planar disparity</i>														
Wanner et al. [28]	21.28	430	34.28	74	31.65	144	20.07	725	32.48	419	30.55	1403	22.64	278
Proposed	22.24	380	37.51	44	34.38	99	22.88	457	33.79	386	31.30	1378	23.78	218
<i>Super-resolution</i>														
Wanner et al. [28]	24.93	230	34.50	122	35.18	129	25.54	287	33.11	378	31.80	1475	26.66	113
Proposed	25.12	228	34.44	123	35.20	129	25.34	289	33.08	379	31.89	1471	26.54	117

表 1: 两个网上不同的档案中的合成和真实世界的光场数值模拟结果。我们将我们的方法与华纳和 Goldluecke[28]比较，关于估计差值的相同分辨率视图合成的和平面代理，以及超分辨率视图合成。对于每一个光场，第一个值是峰值信噪比（越大越好），所述第二值是以为单位的 DSSIM 值（越小越好）。更优值由粗体突出显示。实验的详细说明见文章。

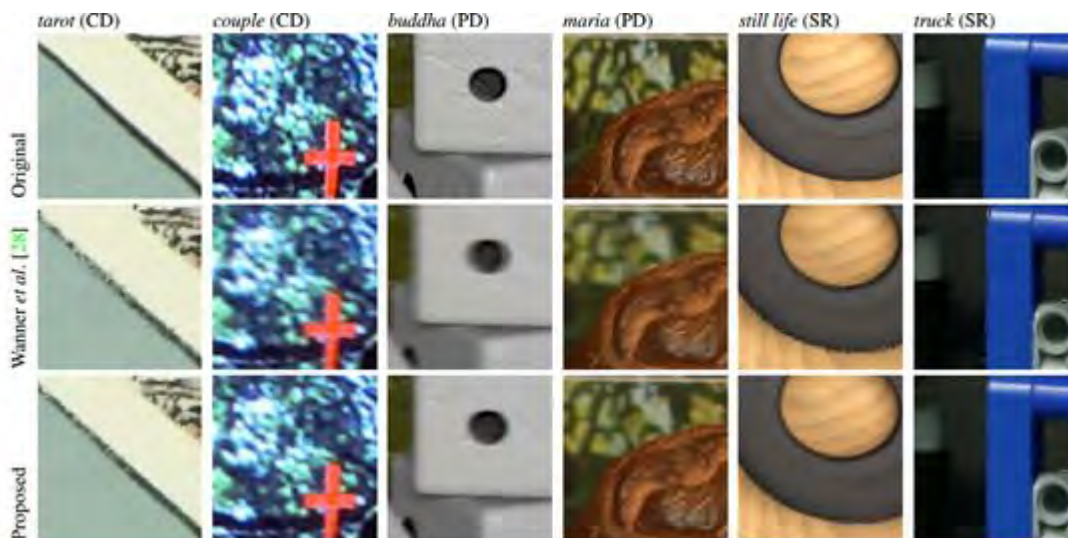


图 4.不同的光场下的新视图的视觉比较。从上到下，每行表现出真实图像的特写，由[28]中获得的结果，以及我们的结果。 CD 代表计算的差异，PD 表示平面差距，SR 表示超分辨率，详见文章。全分辨率图像可以在附加材料中找到。由该方法所获得的结果在视觉上很明显，特别是沿边缘的颜色。

6 讨论与总结

本文的主要贡献是为在近几十年中为新视图合成提出的启发式方法和通过基于物理的生成模型而扣除的能量建立第一次正式

的联系。

该模型可以用来解决一般的问题，其中包括从一组输入异构的图像生成新的图，和一个场景（称为几何代理）的几何描述，它可以

是明确的（即 3D 场景的预估几何形状）或隐式的（即一组新旧视图之间对应的图）。我们贡献的一部分是分析标乐等人[4]提出的模型如何满足几乎所有制定的准则。提出的生成模式提供隐藏在这些准则之后的直观的启发式正式描述。当呈现新图像时，这种统一关键的元素是考虑到估计几何代理的误差。我们广泛讨论了我们的物理模型是如何解释了为什么一些重要的启发当初被发现。该模型优于现有的技术因为它能克服自己的局限。此外，在合成和真实图像上进行的实验表明，我们的方法在呈现图像的质量方面提高了先进的性能。

今后的工作应更好地处理该模型的可见项。在这项工作中，可见度是从深度计算出来的，

但深度本身含有错误，应该传送到可视图。这可能是一个关键的解决方案来把过去的缺失连续性启发式收录到基于物理的贝叶斯框架。同样，把模型延长至非朗伯场景也是至关重要的，但很辛苦。其中需要包括广大的 BRDF 和有价值的信息，以正确模拟输入图和新视图之间的转换。

一个重要的发现是，如果 3D 重建方法或 2D-2D 图像对应方法不仅提供深度估算，而且还有相关联的深度不确定性，这个基于图像的绘制方法可以受益于这种信息以便创造更好的新视图。因此，当发展旨在 IBR 的新的（隐式或显式）重建方法时，这应该是一个新的目标。

参考资料

[1] S. Baker and T. Kanade. Limits on super-resolution and how to break them. PAMI, 24(9):1167 - 1183, 2002. 4

[2] A. Beck and M. Teboulle. A fast iterative shrinkagethresholding algorithm for linear inverse problems. SIIMS, 2(1):183 - 202, 2009.5

[3] T. Bishop and P. Favaro. The Light Field Camera: Extended Depth of Field, Aliasing, and Superresolution. PAMI, 34(5):972 - 986, 2012.3

[4] C. Buehler, M. Bosse, L. McMillan, S. Gortler, and M. Cohen. Unstructured Lumigraph rendering. InProc. SIGGRAPH, pages 425 - 432. ACM, 2001. 1,2,4,5,6,7

[5] A. Chambolle. An algorithm for total variation minimization and applications. Journal of Mathematical imaging and vision, 20(1-2):89 - 97, 2004. 4

[6] C.-F. Chang, G. Bishop, and A. Lastra. LDI tree: A hierarchical representation for image-based rendering. InProc. SIGGRAPH, pages 291 - 298. ACM, 1999. 2

[7] S. E. Chen and L. Williams. View interpolation for image synthesis. InProc. SIGGRAPH, pages 279 - 288. ACM, 1993.2

[8] T. S. Cho, C. L. Zitnick, N. Joshi, S. B. Kang, R. Szeliski, and W. T. Freeman. Image restoration by matching gradient distributions. PAMI, 34(4):683 - 694, 2012. 5

[9] P. Debevec, Y. Yu, and G. Borshukov. Efficient viewdependent image-based rendering with projective texturemapping. Springer, 1998. 2

[10] B. Goldluecke and D. Cremers. Superresolution Texture Maps for Multiview Reconstruction. InProc. ICCV, 2009. 3

[11] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The Lumigraph. InProc. SIGGRAPH, pages 43 - 54. ACM, 1996.2,6

[12] S. Laveau and O. D. Faugeras. 3-d scene representation as a collection of images. InProc. CVPR, volume 1, pages 689 - 691. IEEE, 1994.2

[13] M. Levoy and P. Hanrahan. Light field rendering. InProc. SIGGRAPH, pages 31 - 42. ACM, 1996. 2

[14] L. McMillan and G. Bishop. Plenoptic modeling: An imagebased rendering system. InProc. SIGGRAPH,

pages 39 - 46. ACM, 1995.2

[15] R. Raskar and K. -L. Low. Blending multiple views. In Proc. of Pacific Graphics, pages 145 - 153. IEEE, 2002. 5

[16] S. Roth and M. J. Black. Fields of experts: A framework for learning image priors. In Proc. CVPR, volume 2, pages 860 - 867. IEEE, 2005.4,6

[17] J. Shade, S. Gortler, L. -W. He, and R. Szeliski. Layered depth images. In Proc. SIGGRAPH, pages 231 - 242. ACM, 1998.2

[18] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. In Proc. TOG, volume 27, page 73. ACM, 2008. 4

[19] H. -Y. Shum, S. -C. Chan, and S. B. Kang. Image-based rendering. Springer, 2007. 1, 2

[20] H. -Y. Shum and L. -W. He. Rendering with concentric mosaics. In Proc. SIGGRAPH, pages 299 - 306. ACM, 1999. 2

[21] A. Smolic, K. Muller, K. Dix, P. Merkle, P. Kauff, and T. Wiegand. Intermediate view interpolation based on multiview video plus depth for advanced 3D video systems. In Proc. ICIP, pages 2448 - 2451. IEEE, 2008. 2

[22] K. Takahashi. Theory of optimal view interpolation with depth inaccuracy. In Proc. ECCV, pages 340 - 353. Springer, 2010.3

[23] K. Takahashi and T. Naemura. Super-resolved freeviewpoint image synthesis using semi-global depth estimation and depth-reliability-based regularization. In Advances in Image and Video Technology, pages 22 - 35. Springer, 2012.3

[24] M. Tanimoto. Overview of free viewpoint television. Signal Processing: Image Communication, 21(6):454 - 461, 2006. 2

[25] V. Vaish and A. Adams. The (New)

Stanford Light Field Archive. <http://lightfield.stanford.edu>, 2008. 6

[26] P. Vangorp, G. Chaurasia, P. -Y. Laffont, R. W. Fleming, and G. Drettakis. Perception of visual artifacts in image-based rendering of faces. In Computer Graphics Forum, volume 30, pages 1241 - 1250. Wiley Online Library, 2011.3

[27] S. Vedula, S. Baker, and T. Kanade. Image-based spatiotemporal modeling and view interpolation of dynamic events. Proc. TOG, 24(2):240 - 261, 2005. 2

[28] S. Wanner and B. Goldluecke. Spatial and angular variational super-resolution of 4D light fields. In Proc. ECCV, pages 608 - 621. Springer, 2012.1,2,3,4,5,6,7

[29] S. Wanner, S. Meister, and B. Goldluecke. Datasets and benchmarks for densely sampled 4D light fields. In Proc. VMV, 2013. 6

[30] G. Wolberg. Image morphing: a survey. The visual computer, 14(8):360 - 372, 1998. 2

[31] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski. High-quality video view interpolation using a layered representation. In Proc. TOG, volume 23, pages 600 - 608. ACM, 2004.3

