ＴＯＦ———ｔｉｍｅ　ｏｆ 　 ｆｌｉ ｇｈｔ，有别与利用２Ｄ图像来推导３Ｄ信息这一传统作法．它 是的原理是给被测目标发出连续的脉冲信号，再通过接收返回的信号，然后根据脉冲信号的往返飞行时间， 来计算被测目标的距离．３ＤＴＯＦ技术的发明是机器视觉学科的重要突破，此技术结构并不复杂，易于使用， 且对环境光并无要求［ １］，精度和帧率也达到很高的程度，这些特点使得３ＤＴＯＦ的应用范围愈加广泛．

基于光子飞行时间(Time-of-Flight，TOF)测量 原理的三维场景感知技术是通过测量主动/反射光 程传输时间计算场景深度信息，从而获取场景三维 结构的分布数据。

在过去的几年中，深度感知领域引起了广泛关注。通过提供对三维信息的直接访问，深度传感器使得许多计算机视觉任务变得相当容易。例如包括对象跟踪和识别，人类活动分析，手势分析和室内3D绘图。

目前大多数机器视觉系统都是二维的，在照明受到严格控制的情况下，这是一种经济高效的方法，它们非常适用于图像处理技术检测应用，如边缘检测，模板匹配等。这些算法通过提取关键特征参数，与数据库进行比较以确定是否合格。为了检测ｚ轴上的缺陷，通常会部署一个额外的一维传感器。虽然二维视觉可以借助先进的图像处理算法来避免因照明和阴影条件变化而引起的复杂情况，但相比之下，三维视觉克服了二维视觉的许多问题，因为深度测量可以用来将前景与背景分开。这对于场景理解特别有用，其中第一步是分割感兴趣的主题（前景）从图像的其他部分（背景）。例如，手势识别涉及场景的理解。使用距离作为鉴别器，TOF传感器使得脸部，手部和手指与图像的其余部分分离，因此可以高度置信地实现手势识别。

深度图像的获取主要分为主动和被动两种方式。被动的深度获取方式主要是基于同一场景的二维图像组间的立体匹配，但其算法无法获取遮挡区域的信息而存在着不可克服的缺陷。因此近年来业界出现了一种新的深度信息获取方式称为主动立体获取，它通过采用特定的深度摄像机如Kinect直接测量场景的真实深度，这一方式在过去的几年里获得了广泛的关注并且逐渐成熟。深度摄像机按照成像原理划分主要有飞行时间法(Time of Flight，ToF)、结构光(Structured Light)、三维激光扫描(Laser Scanner)等。目前，全3D数据可以在视频的帧率下实时可用，由此打开了通往3D视觉系统的更广泛应用，如测量、机器人、3D游戏以及手势接口、人体姿态捕捉与行为分析等人机交互。

在深度传感技术中，飞行时间（ToF）成像显示出了巨大的潜力。ToF是Time of Flight的缩写，与传统彩色摄像机不同，ToF摄像机通过使用周期性调幅的光线照射场景来测量深度，该光线沿着直接光路和间接光路反射回相机。然后，摄像机测量入射信号相对于照明信号的相移，从而获取深度信息。由于ToF传感器使用红外光谱中的光在低照度设置下操作，因此通常能够实现非常高的精度。但是，为了从这些原始相位测量中提取深度，必须解决许多具有挑战性的重建问题，例如，对于场景中的单个漫反射器，相位测量仅明确地编码深度直到整数相位包裹，这通过相位展开来解决；在全局照明的情况下，多个光路沿着直接和间接路径干扰，导致严重深度图的多径干扰（MPI）失真；最后，由于IR调制的低吸收深度，以及与RGB CMOS图像传感器相比不成熟的传感器技术，原始ToF测量受到严重噪声的影响。这些影响因素不像普通噪声，并且可能非常大，导致深度估计非常不准确。。

图像的去噪方法通常是对通用的滤波器进行改进，而深度图像不同于自然图像，该类通用的滤波器缺乏针对性，导致重建效果的不太理想。虽然己有学者开始利用深度神经网络对深度图像进行去噪恢复，但由于没有利用额外的先验信息（如同场景的彩色图像），忽略了它们对于重建深度图像的边缘结构的重要指导作用，因此重建效果提升不明显，尤其对图像的边缘结构保持不理想

传统上，在给定来自双频测量的相位和幅度图像的情况下，这三个重建问题，相位展开，MPI减少和去噪，都是在流水线方法中解决的，其中每个步骤孤立地处理单个子问题，虽然这种设计有利于分而治之的算法，但它忽略了各个子模块之间的耦合，并在重建流水线中引入了累积误差和信息损失。

当前，用于解决相位模糊的方法是获得两个不同调制频率的测量，通过用其较低频率对应展开高频相位来保持长距离范围。虽然对于直接场景有效，但是在存在MPI的情况下，这种双频采集方法变得不准确，导致明显的展开错误并且随后形状恢复不准确。

MPI失真通常在后处理步骤中减少。大量的工作探索了简化的双路径或仅弥散问题的解析解决方案，或者试图将MPI作为计算成本高昂的优化问题单独解决，对场景稀疏性有很强的假设。

许多工作提出了混合结构光ToF系统，由于投影机摄像机基线，需要编码和精心同步的照明，显著增大的面积，从而消除了ToF技术的许多固有优势。

虽然MPI，相位展开和去噪是耦合的，但是现有的重建方法都没有以联合和计算上有效的方式解决它们

新兴3D技术及其应用的关键在于可信深度图的获取。各类3D应用的后端技术如模式识别、立体视点映射、人机交互、三维成像等技术的基础就是高质量的场景深度图。传统的被动立体获取方式存在着不可克服的缺陷，难以产生高质量的深度图，而近年出现的主动立体获取设备产生的深度图具有实时、真实、无遮挡区域等优点，并且在许多方面能够填补被动立体获取的缺点，是未来3D技术采集设备的主要发展方向，因此具有很高的研究和实用价值。

当前主流的主动立体获取设备——ToF深度摄像机所采用的测距原理是计算从深度相机发射的光线照射到场景物体上反射回深度相机经过的时间，通过计算获得该物点的深度值。

目前主流的ToF深度相机采用的是间接飞行时间测量法。其原理是由ToF深度相机发射经过余弦波调制的近红外光线，照射到物体上面之后反射回相机的接收镜头，通过计算发射光线和接收光线的调制波相位差来确定光线飞行的时间，从而测得目标物体到深度相机的距离。而使用近红外光线的测距方法存在着不少的缺陷，比如：易受环境光源的影响、场景随机误差大、测距范围较小等等。这些由于测距原理导致的缺陷会使最后测量生成的ToF深度图含有许多误差和噪点，这样的原始深度是无法直接提供给后端使用的。因此必须具体分析ToF相机产生误差的原因，研究建立适当的模型来减少ToF相机的测量误差，使之符合后端深度图的需求。

ToF传感器的主要问题是它们遭受多路径干扰。由于光线是针对每个像素发出的，并且由于光可以以无数种方式反射离开表面，因此特定像素也可以接收最初为其他像素发出的光子。理论上，发射信号经一次反射沿直接路径由传感器接收，传感器处理发射和反射信号的相位信息获取场景深度，但在实际情况中，部分信号会多次反射经直接或间接路径被接收，这些多路径干扰体现在重建的场景深度图像上是大量的噪点。因此我们的主要目标是通过优化减少多路径干扰，提高场景重建的准确度。

当前，通过深度学习解决多路径问题取得的效果相对较好，因此后续将主要通过建立合理的深度学习框架，用TOF原始测量图像训练深度学习网络，优化深度图像，减少多路径干扰。其关键问题在于：选择效果较好的深度学习网络，建立模拟训练数据集训练网络，以及实际采图检验深度学习网络的效果。

深度卷积神经网络已经在监督图像重建问题方面取得了巨大进步，包括去模糊，去噪/修复和超分辨率。虽然这种前馈体系结构对于自然图像上的局部操作很有效，但是对于非局部逆问题，例如MPI移除，通常需要大的感受域。最近，条件生成对抗网络（cGAN）已经在监督和无监督设置下显示出高质量的图像翻译结果。与传统的GAN不同，在cGAN中，发生器G和鉴别器D都观察输入图像。通过将GAN损失与传统像素损失相结合，人们可以学会惩罚输出和目标图像之间的结构差异，而不依赖于领域知识

通过深度学习方法，拟选用条件生成对抗网络进行深度图像的多路径优化。

方案路线：

通过blender 3D建模软件建立虚拟场景；利用基于光线追踪的物理渲染系统PBRT来模拟TOF过程并采集测量图像，建立训练集；采用条件生成对抗网络进行深度学习，并通过建立跳过链接来实现输入与输出几何信息的一致，减少不必要的信息传输量；用模拟训练集训练深度学习网络；用网上共享的实际TOF图像检验深度学习网络，根据结果进行调整。

3月11日-3月17日 文献翻译

3月18日-3月31日 开题。利用blender建立虚拟场景，并用PBRT文件进行TOF模拟和采集瞬态图。

4月1日-4月14日 通过matlab对生成的瞬态图进行处理生成相位和幅度图，建立模拟数据训练集。搭建深度学习环境。

4月15日-4月21日 建立cGAN网络。

4月22日- 建立多个虚拟场景，生成模拟数据训练集

训练深度学习网络，根据结果不断进行修改。并检验网络效果

写结题报告。

1. S. Su, F. Heide, G. Wetzstein, W. Heidrich. Deep End-to-End Time-of-Flight Imaging. In The IEEE conference on computer vision and pattern recognition, 2018.
2. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to image translation with conditional adversarial networks. In The IEEE conference on computer vision and pattern recognition, 2017.
3. Julio Marco, Quercus Hernandez, Adolfo Munoz, Yue Dong, Adrian Jarabo, Min H. Kim, Kaist Xin Tong, Diego Gutierrez, DeepToF: Off-the-Shelf Real-Time Correction of Multipath Interference in Time-of-Flight Imaging. ACM Transactions on Graphics, Vol. 36, No. 6, Article 219(Nov.2017).
4. Freedman, D., Krupka, E., Smolin, Y., Leichter, I., and Schmidt, M. 2014. SRA: Fast removal of general multipath for tof sensors. In Proc. ECCV (to appear).
5. N. Naik, A. Kadambi, C. Rhemann, S. Izadi, R. Raskar, and S. B. Kang. A light transport model for mitigating multipath interference in tof sensors. In The IEEE conference on computer vision and pattern recognition, 2015.
6. 唐贤伦, 杜一铭, 刘雨微, 李佳歆, 马艺玮. 基于条件深度卷积生成对抗网络的图像识别方法. 自动化学报, 2018, 44(5): 855-864
7. 王雪娇. 基于深度神经网络的TOF深度图像去噪研究[D]．浙江大学，2016．
8. 陈宝亮. 基于卷积神经网络和融合TOF与立体数据的高质量深度图获取算法研究 [D]．西安电子科技大学，2018．
9. 吴翔骅． 基于3D TOF技术的深度图像超分辨率重建算法[D]．合肥工业大学，2018．

通过二维图像提取图像中物体的三维信息受外部环境影响较大，三维视

觉［１＿２１克服了二维视觉的许多问题，可以直接将前景与背景分开。３ＤＴＯＦ（ＴｉｍｅｏｆＦｌｉｇｈｔ）技术［３］差别于２Ｄ图中提取３Ｄ的信息，而是基于调制的调制光在场面中的传播时间，测得并算出场面目标物的间距。３ＤＴ〇Ｆ改变了图像视觉行业，利用使用自动调好的调光手段，以及低费用ＣＭＯＳ阵列的像素面来供应３Ｄ场面的间距景深。差别于单点逐点扫描方式［４］，而是任意像素都能测量指定目标物的亮光度和反射回来的调制光的传播时间，从而计算出该点对应的距离景深。３ＤＴＯＦ原理简单，易于便用，不依托于外部环境，不错的准确性和高帧速率使得ＴＯＦ相机应用更有吸引力。三维图像己经是电子技术在视觉领域的重点发展方向，三维图像的用户体验受到重视。获取高质量的场景深度图是各种３Ｄ应用的基本要求，如模式识别、自动尺寸标注、人机交互、虚拟现实增强等。ＴＯＦ相机是未来三维视觉技术集成设备的重要发展方向，具有很高的研宄价值。

大量的先前工作将多路径情况进行了简化，大致可分为两类：

第一类算法侧重于解决由朗伯曲面引起的漫反射多路径的情况；

第二类算法侧重于解决“双路径”多径的情况，一般是我们所说的镜面反射情况。

但是实际情况中，多路径的情况远远比这些要复杂，它可以是两个以上的镜面多路径，也可以是漫反射与镜面反射的组合，甚至还有许多次表面反射等复杂情况，因此无法简单归类。

但是多径通常可以比这更通用：具有两个以上路径的镜面多径是可能的，漫反射和镜面多径的组合也是可能的。

MPI失真通常在后处理步骤中减少。

大量的工作探索了简化的双路径或仅弥散问题的分析解决方案[17,14]，或者试图孤立地解决MPI作为计算成本高昂的优化问题[

29, [29] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and

13, [13] D. Freedman, Y. Smolin, E. Krupka, I. Leichter, and M. Schmidt. Sra: Fast removal of general multipath for tof sensors. In European Conference on Computer Vision, pages 234–249. Springer, 2014. 2, 7

7 [7] A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar. Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization. Optics letters, 39(6):1705– 1708, 2014. 2]，并对其进行了强有力的假设。

场景稀疏。

虽然MPI，相位展开和去噪是耦合的，但是现有的重建方法都没有以联合和计算有效的方式解决它们。

我们在表1中总结了关于MPI校正的重要相关出版物.Fuchs [11]和Jiminez等[29]

使用优化问题和TOF成像的正演模型解决无MPI阶段。

Dorrington等。

Kirmani等人。

Kadambi等人。

[32[32] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. ACM TOG, 32(6):167, 2013. 1, 2, 8

]介绍了用于TOF成像的“编码”照明技术，并演示了在包含半透明物体的场景上进行MPI校正的结果。

这种技术虽然很有前景，但需要对硬件进行修改，并包括妨碍实时性能的硬件限制。

最近，Freedman等人。

[10[10] D. Freedman, Y. Smolin, E. Krupka, I. Leichter, and M. Schmidt. SRA: Fast removal of general multipath for ToF sensors. In ECCV, pages 234–249, 2014. 1, 2

]和Bhandari等人。

[3

2[2] A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar. Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization. Optics Letters, 39(6):1705–1708, 2014. 1, 2, 3

]

[33Spumic: Simultaneous phase unwrapping and multipath interference cancellation in time-of-flight cameras using spectral methods.

]使用二次反射来对MPI进行近似建模，或应用多频测量（或频率扫描）的方法来解决多路径问题。这些方法都对场景有很强的稀疏性假设，当这些假设无法成立时，比如场景中存在次表面反射，将会出现严重的误差。

提出了用多频测量来解决一般性的MPI的迭代方法，但其局限性在于，他们的方法局限于K稀疏模型，其中，K是有限量

[32] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. ACM TOG, 32(6):167, 2013. 1, 2, 8

]

介绍了一种编码照明技术

介绍了用于TOF成像的“编码”照明技术，并演示了在包含半透明物体的场景上进行MPI校正的结果。

这种技术虽然很有前景，但需要对硬件进行修改，并包括妨碍实时性能的硬件限制。