**毕业设计（论文）题目： 基于深度学习的TOF深度图像优化**

**学 院：** 光电学院

**专 业：** 测控技术与仪器

**班 级：** 04111501

**姓 名：** 李颖婕

**指导教师：** 金伟其

**摘要**

飞行时间（TOF）相机通过使用低成本CMOS像素阵列和有源调制光源，由每个像素向场景发射调制信号光得到发射光与反射光之间的时间差来进行场景中对应点的深度结算，以获得场景深度图像，它可以不借助机器视觉算法直接进行三维重建。并且不受环境光照和场景特征数的影响，可探测的深度反围较广。结构紧凑、方便使用、大测量范围和对环境光照的低要求，使得这一技术在机器人导航、三维重建，以及人机交互等方面有很高的应用价值。但由于传感器接收到的反射光往往是来自多种反射路径的直接反射和间接反射的混合信号，因此对深度的计算准确度降低。使得生成的深度图像受到多路径干扰产生许多噪声和错误。为了降低这种影响，我们采用深度学习的方法，运用深度学习网络对深度图像进行优化。

**Instract**

**第1章 绪论**

**1.1 二维相机与三维相机发展简介**

由于三维相机加入了z轴信息，对场景的理解能力更强，可以根据获取的三维信息，对场景进行三维建模。将三维相机技术集成到便携设备上，结合当前非常火的虚拟现实、增强现实技术，在AR游戏、AR装潢、3D试装、体感游戏、全息影像交互等方面都有广阔的应用前景。

随着5G时代的到来，在高带宽的保证下，3D技术的发展和应用将会呈爆发式增长。虚拟现实、增强现实甚至是混合现实技术都将在一个更便利的平台上发展和实现，

未来的3D视频通话、虚景+虚景的远程虚拟现实（VR）、虚景+实景的远程增强现实（AR），甚至是虚景+实景+数字化信息的远程混合现实（MR）等在现实场景的使用将迎来爆发式应用。在拍照时对前景、背景的精细处理方面也有很大的发展空间。

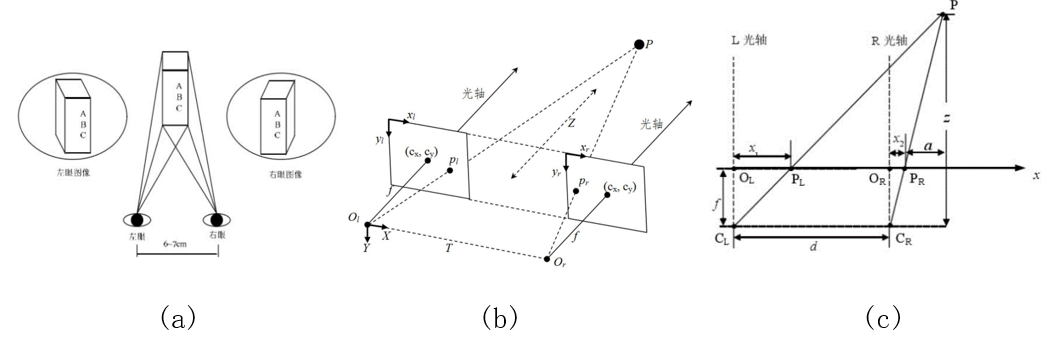
**1.2 深度相机成像方式**

当前技术较成熟、应用较广的深度相机可以分为三种，立体视觉相机、结构光和TOF相机。下面分别对这三类深度相机进行介绍。

**1.2.1 立体视觉**

当前最常见、应用最广泛的深度获取设备就是应用双目立体视觉技术，它基于视差原理，模拟人的双眼获取三维信息的方式。用两个相机对场景进行拍摄（或是同一相机进行移动旋转），在两相机内参一致的情况下，找到场景中同一点在两相机获取图像的位置，计算视差，再利用三角测量进行计算，得到该点的深度信息。

图1- （b）展示了双目立体视觉的测量原理。如图，和为位于同一平面的的两平行相机，两相机间的距离为基线理想情况下，任一相机通过沿水平方向平移都可与另一相机重合。两相机有各自的成像平面和坐标系，为空间内一点，它在两相机成像面上的投影分别为和，、、三点构成的平面与两成像面的交线称为级线。当相机水平放置时，左右两边只存在水平视差，即,因此我们只看平面（图1- （c）），为偏离原点的距离，为偏离原点的距离，为视差。



根据几何关系可以建立等式：

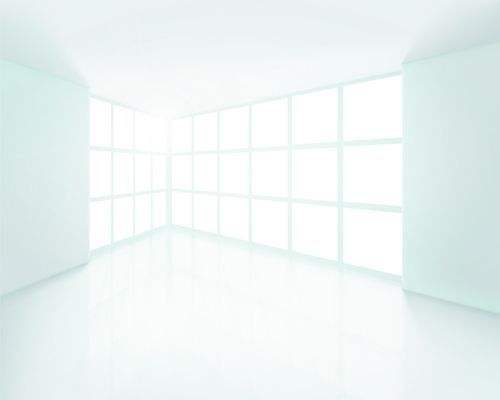
|  |  |  |
| --- | --- | --- |
|  |  | （1-1） |

求得

|  |  |  |
| --- | --- | --- |
|  |  | （1-2） |

其中，为物体到相机的距离，即所求的物体深度，为相机焦距。可以看到，通过计算将场景深度与视差联系了起来。

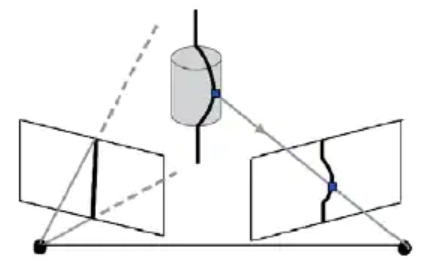
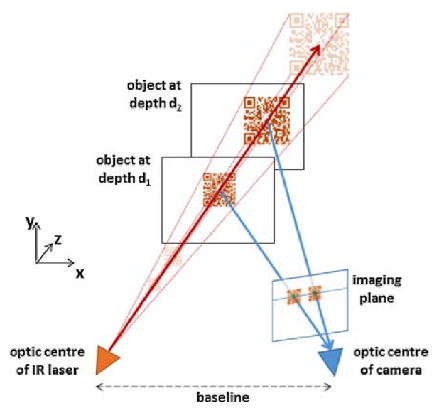
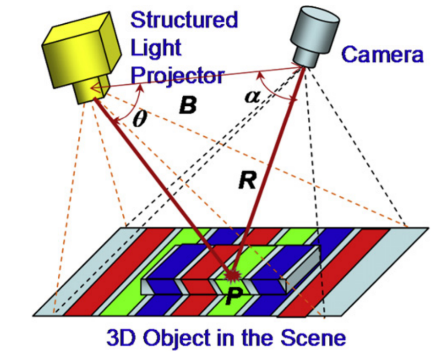
由于双目立体视觉技术并不向场景投射光源信号，而是完全依靠所获取的图像获取场景深度，因此属于被动式测量。但在实际情况中很难保证两个相机光心完全水平，这样就会导致两图的级线不共面，因此需要对图像进行前期矫正；同时，由于双目立体视觉技术主要通过图像处理技术，对图像进行特征匹配，因此对于光照角度变化、光照强度变化、缺乏视觉特征的场景会匹配错误（图）。



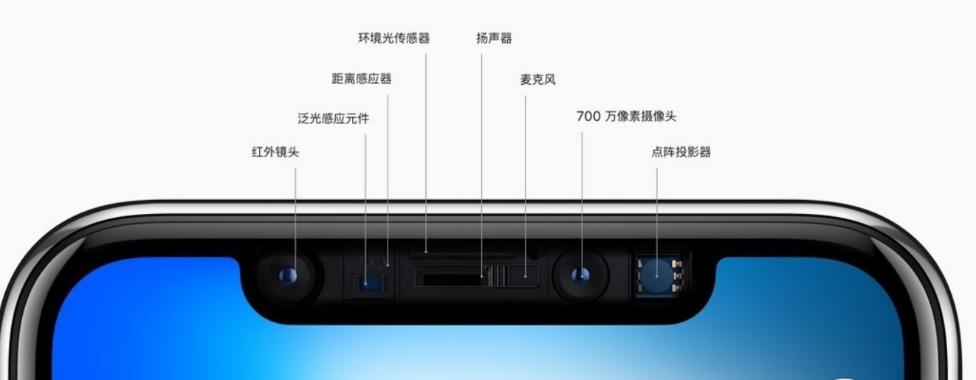
**1.2.2 结构光**

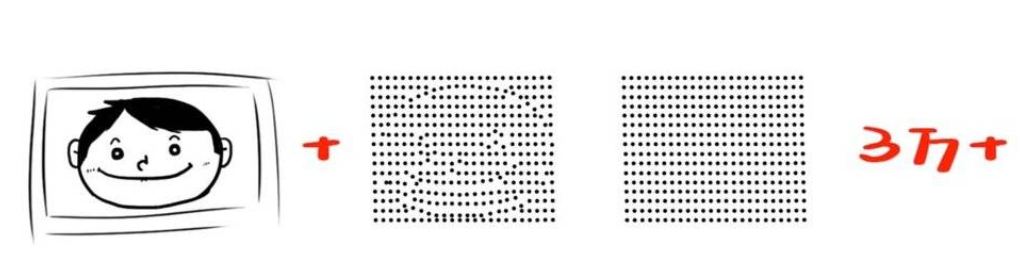
为了解决双目立体视觉的诸多问题，出现了用结构光获取深度的方法。结构光是主动式深度获取的一种，其主要原理是向场景中主动投影二维空间调制编码的结构光图案，投影的图案经过场景深度的调制，出现失真，相机拍摄到失真的投影图案，根据失真信息来计算空间点的深度。

图1-2- 展示了结构光照明的原理。结构光投射器向场景中发射红外二维编码图案，通常情况下，如果场景是一个没有三维信息的平面时，相机捕获的图像应与投射的结构光图案相似，但是当场景是非平面的，从相机来看，场景的几何形状会使投射的结构光图案产生失真效果[1]，而红外相机通过捕获失真图案，再利用三角测量等进行计算，将场景的深度信息从失真图中提取出来[2]。



我们熟知的iphoneX正是应用了结构光技术，实现面部解锁，并将其称为Face ID，应用于购物、人像特效拍摄、创建动画表情等方面。虽然“刘海”屏出现之初一直不被接受，但其中恰恰囊括了强大的结构光技术。如图1-2-所示，iphoneX的面部解锁功能主要由红外镜头、泛光照射器、距离感应器、点阵投影器来配合实现。当人脸面对屏幕，首先唤醒距离传感器，再由距离传感器启动泛光照射器，泛光照射器向外发射非结构的红外光，由红外镜头获取图像，返回到手机中进行人工智能识别，判别为人脸后，启动点阵投影器产生大约3万多个投射点，投射到人脸，再由红外镜头捕获经三维深度调制过的点阵图，计算出各点深度信息后，再与预存的人脸信息对比，实现身份识别。





由于结构光采用主动照明，将含有特征信息的光源图案投影到场景中，使得相机拍摄到的图像中包含了很多特征点，再通过找到特征点和投影点的匹配关系，确定空间点的坐标，相比于双目立体视觉对于缺乏特征场景的难以匹配，结构光对于照明不足和缺乏特征的场景能够非常好地进行深度获取。与TOF相机相比，结构光的分辨率较高，但也存在一些不足，虽然结构光对于光照变化和环境的特征要求不高，但是在强光照情况下，投影图案会被淹没在场景中，使得对于投影到场景中的图案的拍摄难度增加，因此结构光易受到反光和强光照的影响。

**1.2.3 TOF相机**

TOF相机即飞行时间相机，其工作原理是由相机的传感器发射源发射调制信号光线到场景中，通过接收由场景反射回的光线，通过计算发射光线与接收到光线之间的时间差，进而计算出场景深度。其具体的工作原理将在下一章进行详细分析。

通常情况下，TOF的调制照明来自于固体激光器，或是工作于近红外波段的LED。这就使得TOF对场景深度的捕获不会受到环境照明和场景特征的影响，同时，由于TOF并不通过相机对场景进行拍摄，因此面对强光情况依然能正常工作，而且，理论上讲，TOF能够探测的场景的深度范围要远大于结构光和双目立体视觉，因此适合在户外工作。

但是，由于TOF往往使用特殊的传感器和快门，每个像素点独立向场景发射红外信号光并接受反射光进行深度计算，其复杂程度大大增加，这使得传感器单个像素的尺寸要远大于一般的图像传感器的像素尺寸，因此TOF相机的分辨率通常较低；同时，传感器接收到的反射光除了直接反射回的光，还包含大量经过多次反射的光信号，这使得对于场景深度的计算准确度大大降低，这也促使大量的研究致力于实现TOF深度图像的超分辨。

**1.2.4 立体视觉、结构光、TOF相机对比**

**TOF与双目立体视觉** TOF相机属于主动式深度获取，其发射的信号光源为高频近红外光，双目立体视觉为被动式深度获取技术，不投射光源，直接对拍摄到的RGB图像进行特征匹配；由于TOF的每个像素都要配置光源发射和传感器，因此像素尺寸较大，受物理器件和功耗限制，分辨率很难有大幅提高，双目立体视觉对相机要求较低，因此其分辨率随相机的分辨率改变，因此可以达到很高的分辨率；由于受基线约束，双目立体视觉的基线越长，可探测的深度范围越广，但两个相机可拍摄的场景重合度降低，因此其探测范围一般不会太远，且对环境光照和场景特征量较敏感，TOF相机由于其只探测发射信号与接收信号的时间差，并不对场景采图，因此受环境影响较小[3]，探测范围较广，适合户外工作；最后，双目立体视觉由于需要进行场景匹配和三角解算，同时易受环境光照的影响，其算法较为复杂，需要进行大量的情况预设和筛选，而TOF计算较简单，软件复杂度较低，但硬件要求高。

**TOF与结构光** TOF与结构光同为主动式深度获取方式，结构光透过投射编码图案再根据图案进行特征匹配来获取深度，需要对场景采图，虽然通过主动投射特征图案到场景中，降低了对场景特征量的要求，但在场景中存在强光或镜面反射情况时受干扰较大，在户外场景中，投射光源易被淹没在场景中，因此适用于室内工作，相比之下，TOF对环境要求较低，且探测范围更广；结构光的由于需要投射编码光源，因此硬件复杂度较双目立体视觉略高，但远低于TOF，因此分辨率仍高于TOF，可达到1080P；由于TOF需要向全场景发射信号光源，每个像素都发射高频红外光源，对功耗要求大，而结构光只需投影部分场景，功耗远小于TOF。

总之，由于TOF硬件方面使用低成本的CMOS传感器技术，且传感器照明基线非常小[4]，易于集成在便携设备上，在获取场景深度时受环境影响小，探测范围广，且软件复杂度低，因此具有更加广泛的应用前景。

**1.3 本文主要研究内容和贡献**

TOF对深度的测量是通过使用周期性振幅调制光源照射场景，光线经场景反射，沿直接或间接反射路径回到相机，相机测量入射信号相对于发射信号的相位差。为了从原始测量数据中提取深度信息，必须解决几个重要的重建问题。首先，在深度断层处和场景材料反射率出现明显偏差时，对应像素形成的深度值在可探测范围内任意取值，导致重建时出现飞行伪像；其次，对于单个像素来说，只能在一个周期内进行明确深度解算，否则会出现周期性模糊问题，这需要通过相位展开来解决[4]；最后，在存在全局照明的情况下，沿直接和间接路径的多路径光形成多路径干扰[5][6]，使深度图产生严重失真。

本文的主要研究内容是采用深度学习的方法，通过构建条件生成对抗网络对深度图像进行优化，并用基于光线追踪的物理渲染（PBRT）[7]进行仿真数据集的建立，解决了真实数据集无法获取地面实况作为参考集的问题，通过对网络的训练和调整，改善了深度图的成像质量，提高了深度重建的准确度。

**1.4 论文组织结构**

针对上文所述的研究内容，本文的组织结构如下：

第1章是文章的绪论，主要介绍了机器视觉相机的发展，对现行的几款深度相机进行介绍和对比，并介绍了研究内容和目标

第2章详细介绍了TOF相机的成像原理，并对几种主要的影响深度重建的噪声进行了分析。

第3章介绍了一种基于深度学习的TOF图像优化方法，介绍了网络构架、损失函数的设置，以及数据集的建立。

第4章为具体的训练过程和实验结果。

第5章为实验结果的分析。

第6章为全文工作总结和展望。

**第2章 TOF相机成像原理与噪声分析**

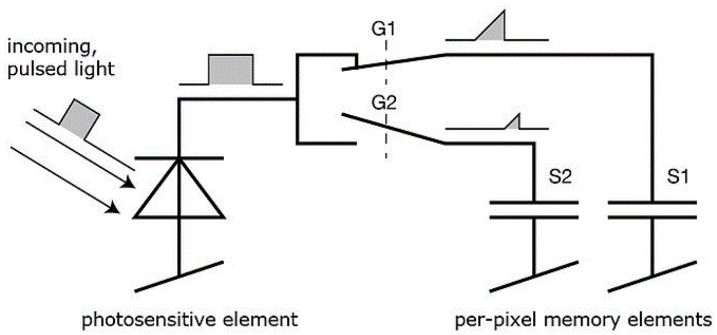
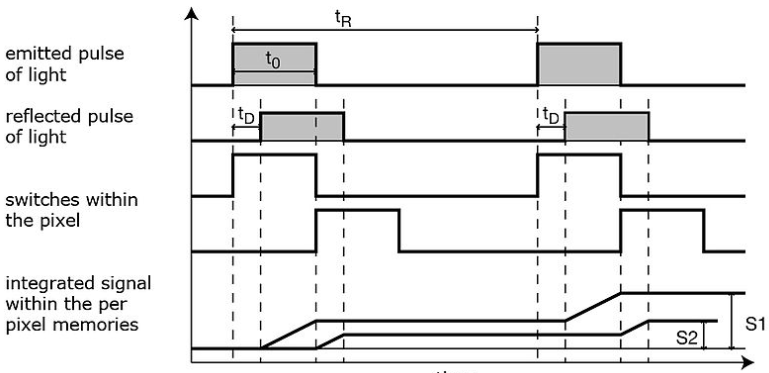
**2.1 TOF相机成像原理**

飞行时间相机的基本原理是由发射器向场景发射调制过的光信号，通过测量光信号从发射器发出到被检测目标并反射回到接收器的时间，计算出检测目标的深度。由于每个像素都会向场景发射光信号，因此TOF可以同时得到整个场景的深度信息。根据调制方式，TOF可以分为脉冲调制和连续波调制。

**2.1.1 脉冲调制**

脉冲调制法向场景发射调制脉冲（通常是皮秒宽度的光脉冲），直接测量发射脉冲通过场景目标反射回来的时间差，进而计算脉冲走过的距离，属于直接测量法。

图2-1-1展示了基于脉冲调制的TOF原理，它应用发射离散光脉冲的光源，这些脉冲经过场景中的目标后向散射，并由系统检测[8]。由于脉冲在场景中的往返，后向脉冲存在时间延迟，图像传感器在一定的曝光时间内对入射光强进行积分。通过使用极快的快门，可以确定入射脉冲与积分窗口之间的错位，这样有助于根据传感器检测到的光量来估计脉冲延迟量。

每个像素中存在两个开关、和两个电荷存储器、，开关窗口与脉冲同宽。工作时，开关与脉冲发射器同时开启和关闭，开关在脉冲发射器关闭时开启，也就是恰好延迟了一个脉宽的时间。由于后向脉冲存在延迟，只有部分的光脉冲通过进入存储器，剩下的部分则通过进入，定义脉宽为，则脉冲的延迟时间可借助、中的电荷量来计算：

|  |  |  |
| --- | --- | --- |
|  |  | （2-1） |

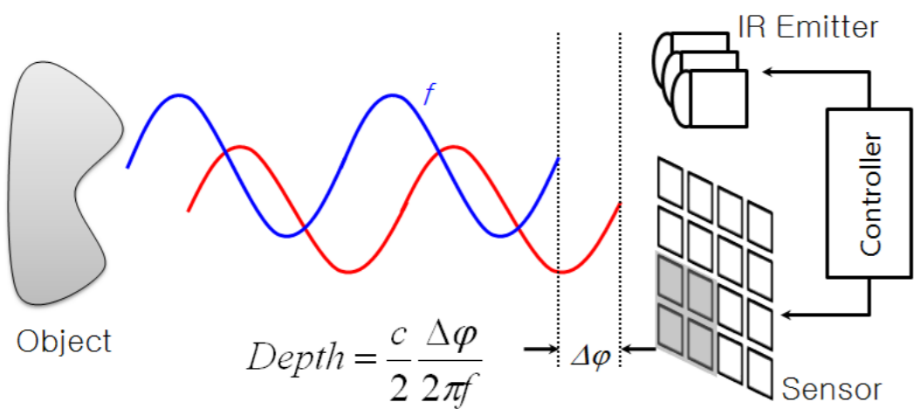
进而根据延迟时间的一半计算出深度信息：

|  |  |  |
| --- | --- | --- |
|  |  | （2-2） |

脉冲调制法发射端要发生高频、上升和下降时间非常快的高功率极短脉冲，一般是激光或激光二极管，由于需要严格控制开关，因此对时间的检测要十分精确，对于快门和传感器的响应速度要求也较高。

**2.1.2 连续调幅波调制**

连续调幅波法通过测量发生信号与接收信号之间的相位差，来计算深度，属于间接测量法。信号的波形可以是正弦波或方波，在已知调制频率的前提下，可以通过相位差与场景深度之间的相关关系来解算对应深度[8][10]。



由TOF的基本原理可知，场景深度由时间延迟量计算

|  |  |  |
| --- | --- | --- |
|  |  | （2-3） |

在连续波调制中，时间延迟不能直接得到，但发射信号与接收信号的相位差是可测的，其关系为

|  |  |  |
| --- | --- | --- |
|  |  | （2-4） |

在连续波调幅法中我们令发射信号为：

|  |  |  |
| --- | --- | --- |
|  |  | （2-5） |

则经过场景的接收信号为：

|  |  |  |
| --- | --- | --- |
|  |  | （2-6） |

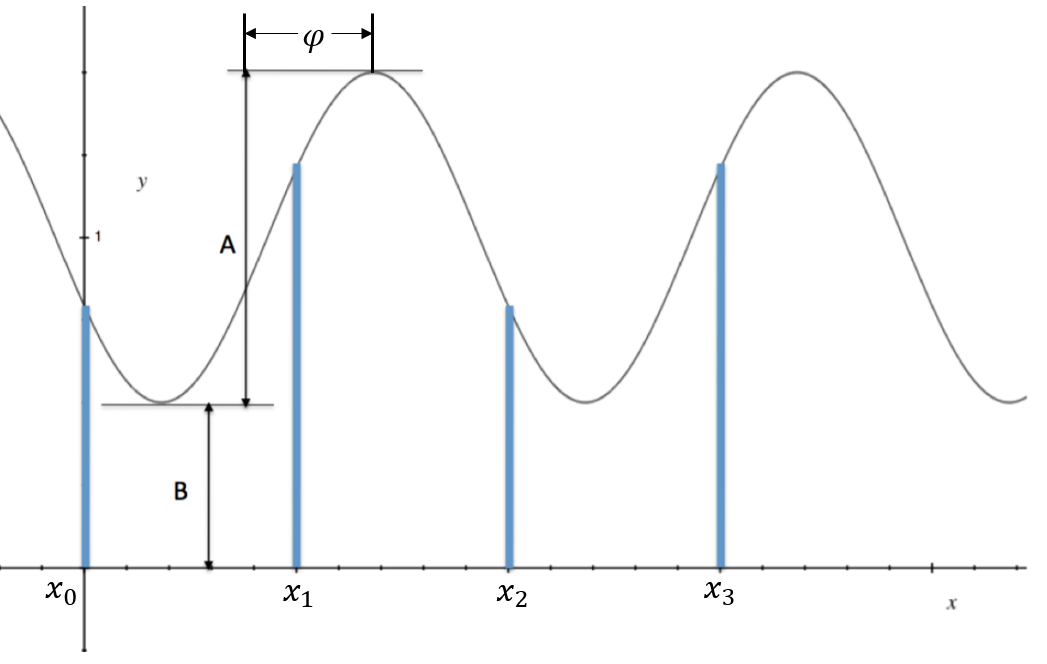
发射信号与接收信号之间的相关关系为：

|  |  |  |
| --- | --- | --- |
|  |  | （2-7） |

化简后：

|  |  |  |
| --- | --- | --- |
|  |  | （2-8） |

利用“四点法”可以计算出则入射信号与接收信号之间的相位差。如图，在一个周期内，每隔进行一次采样，对取四个采样点[10]，



则相位差计算公式为：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | | （2-9） |
|  |  | （2-10） | |
|  |  | （2-11） | |

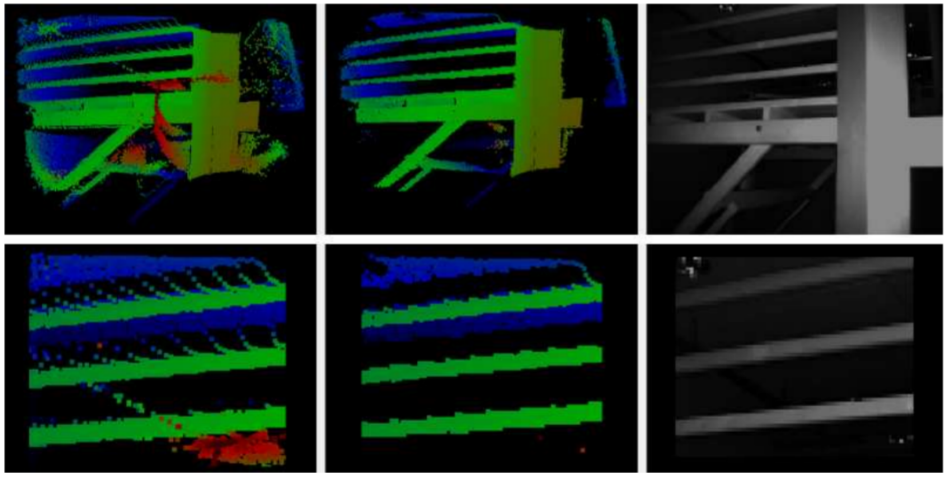
A是所接收信号的幅值，它受物体反射率和传感器灵敏度的影响，B是环境光照带来的偏移量。

**2.2 TOF深度图像噪声分析**

**2.2.1 飞行伪像**

在场景中材料具有不合适的反射率时，以及深度不连续处往往会出现飞行伪像[3]。在这些具有较大的深度断层处，根据深度不同，场景被分为前景和后景，TOF相机发射的近红外信号光一部分由前景中的目标反射，另一部分由后景中的目标反射，这一部分的像素收集来及前景和后景的调制光[8][9]，然后计算得到的深度值既不属于前景的成像面也不属于后景的成像面，由于我们对深度的重建并不是线性的，这个值当然也不局限于两者之间，这意味着，它不同于二维图像中的边缘模糊问题，相机可探测深度内的任何值都是可能的，这取决于在这些像素对应的场景内，那些未知的描述场景信息的各项参数的值（比如反射率）。

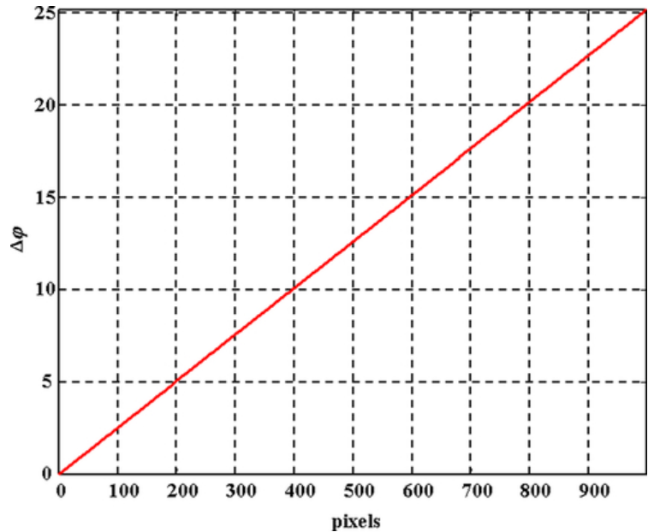
如图2-2-1所示，通过给不同深度上色来体现图像的深度信息，红色距离相机最近，蓝色最远，最右边是场景的灰度图作为参考。可以明显看到深度图中存在飞行伪像，这些部分严重干扰了对于场景的重建，因此需要对伪像进行消除。第二行是对第一行图像的局部放大。



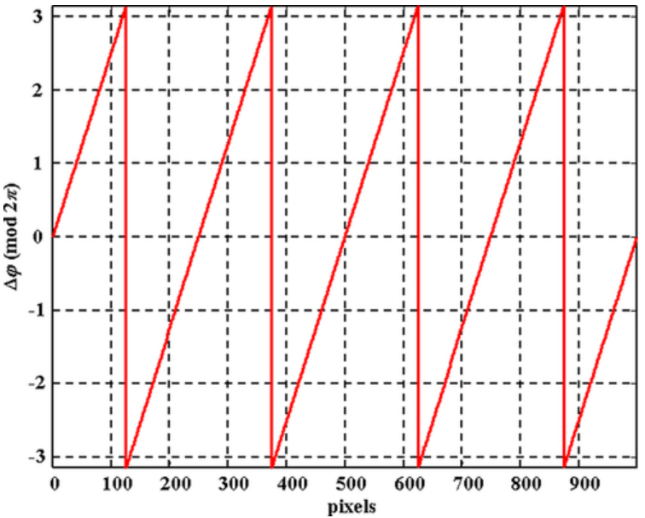
**2.2.2 相位包裹**

在连续波调制的TOF相机中，TOF通过计算相位差来获得深度信息，但由于使用的是周期性光源，对于单个传感器来说，只能在一个周期内进行明确的深度解算，因此在距离测量中存在周期性模糊，这就涉及到相位展开问题。二维信号的相位展开是一个长期存在的问题，在核磁共振成像、自适应光学，以及我们现在探讨的TOF成像中都是经典的问题。

由于相位的周期性，原本连续的相位信息在离散的包裹相位中被采样，当绝对相位值被包裹在一个周期[]内，就会出现模糊问题。以某个物理量（如距离）为例，当调制频率一定时，距离与相位成比例，我们假定理想情况下，相位应沿着某一行像素线性变化（如图2-2-）。

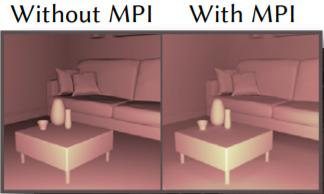


但由于原始计算出的并不是绝对相位，而是包裹相位，因此，在周期两端会出现相位跳变（如图2-2-）。

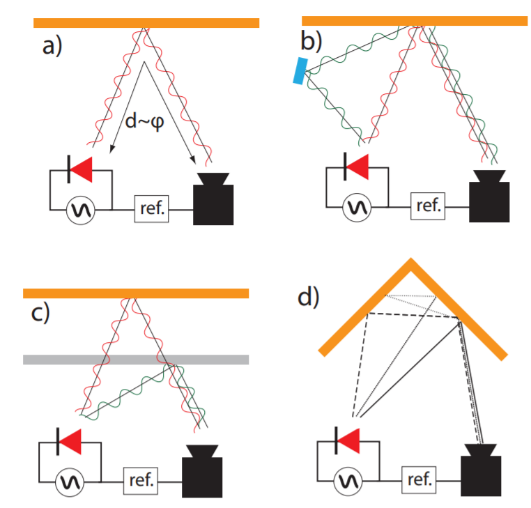


为了得到绝对相位，我们需要寻找应该添加到相位值上的正确的整数个相位周期，以得到真实的相位值。

**2.2.3 多路径干扰**



TOF相机对于场景深度的测量是全场景每个像素点同时发射信号光源并根据接收信号进行深度解算，因此可以同时得到全场景的深度信息。理想情况下，每个像素点都应只接收到由该点发射的信号光源经过直接反射得到的反射光，但在实际情况中，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，对于某一像素点而言，除了这一点的直接反射光外，很多在场景中经过多次反射的光也会回到这一像素中，因此该点接收到的反射光是来自多个路径的反射光的线性组合。这就使得在计算深度时得到的发射光与接收光之间的信息是错误的，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，严重影响了场景的重建。



在存在全局照明的情况下，接收信号包括直接反射和来自场景中其他多个点的反射，我们将这些间接反射量表示为来自个场景点的相位为幅值为的正弦信号,此时接收信号就是这些信号的混合，可以写为：

|  |  |  |
| --- | --- | --- |
|  |  | （2-12） |

其中，不限于有限的数目，因此，将含有多路径的式子代入之前求相位和幅值的公式中，得到新的公式：

|  |  |  |
| --- | --- | --- |
|  |  | （2-13） |
|  |  | （2-14） |

通过上面的式子我们可以看到，混合了多路径的相位与直接反射的相位不再相等，这就导致深度值被改变。

**第3章 基于深度学习的TOF图像优化**

**3.1 TOF深度图像优化现状**

图像优化去噪的常用方法是对现行的经典滤波器进行优化升级，使其满足去噪要求，但深度图像与传统的二维图像不同，由于加入了深度信息，很难找到具有针对性的滤波器来进行优化。传统方法对相位展开和多路径等问题分别优化。

**相位展开** 相位展开方法可分为单次拍摄和多次拍摄两类。单次拍摄通过全局优化的方法处理单一调制频率下捕获到的数据，多次拍摄方法则是利用多个频率下捕获的数据来计算包裹数。目前应用较广泛的相位解包裹方法是双频测量方法。

相位展开的目标是从包裹相位中找到相对相位跳变，相位跳变一般在图像x方向和y方向的相邻两像素之间定义。通常情况下，大部分算法会预先设定判别条件以区分正常的相位变化和由相位包裹引起的相位跳变。通常认为相邻像素间的相位应该是接近的，通过这种假设筛选相位跳变，使展开值尽可能准确。然而即使没有噪声，也可能存在无法单纯根据相位梯度来判断相位跳变的情况，有些场景中两个相邻像素的实际相位差的确大于调制波长，在存在多路径干扰的情况下，这种并非由相位包裹引起的相位跳变的情况就更多了。

**多路径优化** 当前对于多路径问题的优化方法大致可以分为两大类：

第一类从多路径组成成分下手，对一般性的多路径问题进行了简化，一种方法侧重于解决由朗伯曲面引起的漫反射的多路径情况；另一种方法侧重于解决“双路径”型的多路径情况，也就是我们所说的镜面反射。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。

第二类通过提高计算成本来优化多路径问题。

介绍了一种“编码”照明技术用于TOF成像，通过对光源编码，使不同的场景反射出现峰值差异，并演示了在包含半透明物体的场景上进行MPI校正的结果。这种技术需要对硬件进行修改，且对硬件要求较高，由于硬件限制，会妨碍其实时性能。

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，因此依然无法适用于真实场景的多路径（比如存在次表面反射情况时）。

**3.2 深度学习网络**

**3.2.1 卷积神经网络**

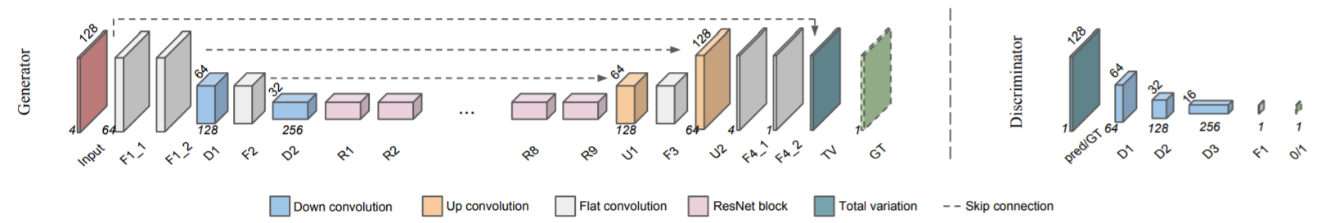
当前，卷积神经网络是解决各种图像处理、预测问题的主要工具，卷积神经网络通过最小化一种评估结果好坏的标准，也就是我们所说的损失函数来进行学习，虽然这个学习的过程由网络自主实现，但仍然需要人为设计有效的损失函数。也就是说，我们需要告知网络应该最小化哪些损失函数，从而得到想要的结果。

**3.2.2 生成对抗网络**

如果能够找到一种方法，我们只需给网络一个类似“使输出图像接近真实图像”这样的最终目标，让网络自主学习能够实现这一目标的损失函数，那么不仅将大大减少工作量，同时也能使学习结果更加接近预期。近期提出的条件生成对抗网络（GAN）正是以这样的方式工作。GAN网络通过学习一种判别图像是真实图像还是伪图像的损失函数，在训练生成器的过程中最小化这一损失。由于GAN网络能够自主学习与数据相匹配的损失函数，它被广泛应用于那些原本需要设置大量损失函数的问题。

**3.3 基于深度学习的TOF图像优化**

**3.3.1 网络构架**

****

生成器G：

鉴于MPI可以表示为空间变化的卷积，MPI补偿可以建模为一组深度空间中的卷积和反卷积。换言之，MPI原则上可以通过设计卷积神经网络（CNN）来解决。具体而言，因为不正确和正确的深度只是稍微不同（但结构相似）卷积自动编码器（CAE）将是一个有效的解决方案。

卷积自动编码器是一个功能强大的工具，它采用相同的输入和输出通过无监督学习来学习低维特征向量的隐藏表示，从而产生两个对称的网络：编码器和解码器。这样可以更深入地构建网络架构，并在构建这些表示时保留空间局部性，较低维特征向量在输入上保留相关结构信息并消除现有错误，有效地返回恢复（参考）图像。

网络概述：我们首先学习编码器网络作为传统的无监督CAE培训之前的深度，其中未标记的真实深度图像（具有未知错误）用作输入和输出。生成的编码器获得不正确深度图像的低维特征向量。第二步，与稀疏编码不同，其中原始信号被重建为字典原子的线性乘积，训练可以从这些特征向量重建参考深度图的解码器。为此，保持编码器网络不变并将其与残余解码器网络级联。通过受监督的CNN训练解码器网络的权重。

由于输入和输出数据呈现相似的结构，我们将网络建立在自动编码器上，我们分两个阶段进行训练。首先，使用编码器（卷积滤波器）来学习表示MPI干扰的深度图像;然后，训练解码器（反卷积滤波器）以从使用基于物理的时间分辨渲染器生成的合成场景中校正深度。这种方法允许我们通过使用具有MPI损坏深度的大规模捕获训练集来训练编码器来解决ToF中的关键问题，缺乏地面实况数据；并用具有地面实况信息的较小的合成训练集来训练解码器。

鉴别器网络D由3个向下卷积层组成，在重叠的片中对G的预测进行分类。

**3.3.2 损失函数**

**3.3.3 仿真数据集的建立**

**第4章 实验结果**

**4.1 实验环境**

**4.2 网络训练过程**

**4.3 实验结果比较**

**第5章 结论**

**第6章 工作总结与展望**

1. Geng J . Structured-light 3D surface imaging: a tutorial[J]. Advances in Optics & Photonics, 2011, 3(2):128-160.
2. Han J , Shao L , Xu D , et al. Enhanced Computer Vision With Microsoft Kinect Sensor: A Review[J]. IEEE Transactions on Cybernetics, 2013.
3. M. Reynolds, J. Dobos, L. Peel, T. Weyrich, and G. Brostow. Capturing time-of-flight data with confidence[J]. In Proc. CVPR, 2011: 945–952.
4. S. Su, F. Heide, G. Wetzstein, and W. Heidrich. Deep end-to-end time-of-flight imaging[J]. In Proc. CVPR, 2018，1:6383–6392.
5. Bhandari A , Kadambi A , Whyte R , et al. Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization[J]. Optics Letters, 2014, 39(6):1705.
6. Freedman D, Smolin Y, Krupka E, et al. SRA: Fast Removal of General Multipath for ToF Sensors[M]// Computer Vision – ECCV 2014. 2014.
7. Mattpharr D, Humphreys N G. Physically Based Rendering[M]// Physically based rendering. 2004.
8. M. Schmidt. Analysis, Modeling and Dynamic Optimization of 3D Time-of-Flight Imaging Systems[D]. Germany:Ruperto-Carola University of Heidelberg, 2011.
9. Kahn, Svenja & Wuest, Harald & Fellner, Dieter. Time-of-flight based Scene Reconstruction with a Mesh Processing Tool for Model based Camera Tracking[J]. VISAPP 2010 - Proceedings of the International Conference on Computer Vision Theory and Applications. 2010, 1. 302-309.
10. Lange, R. 3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology. Ph.D. Dissertation, University Siegen, Siegen, Germany, 2000.
11. Hansard M , Lee S , Choi O , et al. Time of Flight Cameras: Principles, Methods, and Applications[M]. Springer Publishing Company, Incorporated, 2012.
12. Naik N , Kadambi A , Rhemann C , et al. A Light Transport Model for Mitigating Multipath Interference in TOF Sensors[J]. 2015.
13. Marco J , Hernandez Q , Muñoz, Adolfo, et al. DeepToF: Off-the-Shelf Real-Time Correction of Multipath Interference in Time-of-Flight Imaging[J]. Acm Transactions on Graphics, 2018, 36(6):1-12.
14. Isola P , Zhu J Y , Zhou T , et al. Image-to-Image Translation with Conditional Adversarial Networks[J]. 2016.
15. Mirza M, Osindero S. Conditional Generative Adversarial Nets[J]. Computer Science, 2014:2672-2680.
16. Droeschel D , Holz D , Behnke S . Multi-frequency Phase Unwrapping for Time-of-Flight cameras[C]// Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on. IEEE, 2010.
17. TI C, YANG R, DAVIS J. Single-Shot Time-of-Flight Phase Unwrapping Using Two Modulation Frequencies[C]//International Conference on 3D Vision, 2016:667-675.
18. Kadambi A , Whyte R , Bhandari A , et al. Coded time of flight cameras[J]. ACM Transactions on Graphics, 2013, 32(6):1-10.
19. Chou P A . SPUMIC: Simultaneous Phase Unwrapping and Multipath Interference Cancellation in Time-of-Flight Cameras using Spectal Methods[C]// IEEE International Conference on Multimedia & Expo. IEEE, 2013.
20. Mei, J., Kirmani, A., Colaco, A., Goyal, V.K. Phase unwrapping and denoising for time-of-flight imaging using generalized approximate message passing[C] // IEEE International Conference on Image Processing. IEEE, 2013.
21. Ronneberger O , Fischer P , Brox T . U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. 2015.