

光鉴分享 | 《ToF 深度相机技术白皮书》 完整版

摘要： 现行专业级或消费级的 3D 相机所采用的三角法(Triangulation)和飞时法(Time-of-Flight,ToF),现因苹果公司最新版 iPad Pro 的出现--搭载了 d-ToF 技术的深度相机--已然为 3D 视觉在消费场景的应用推动了新的机会。为了让读者更深入地了解 ToF 技术，我们根据行业现状、学术界的最新成果，编写了此版《ToF 深度相机技术白皮书》。全文可下载链接：http://www.deptrum.com/data/Deptrum_ToF_White_Paper.pdf

1.3D 视觉介绍

3D 视觉技术能够获取现实三维场景完整的几何信息，利用带有深度信息的图像来实现对于场景的精准的数字化的，从而实现高精度的识别、定位、重建、场景理解等机器视觉的关键功能。以 2010 年的 Kinect 和 2017 年的 iPhoneX 的

发布为标志，3D 视觉技术从传统意义上只应用于专业领域的高端技术变成了消费级产品。

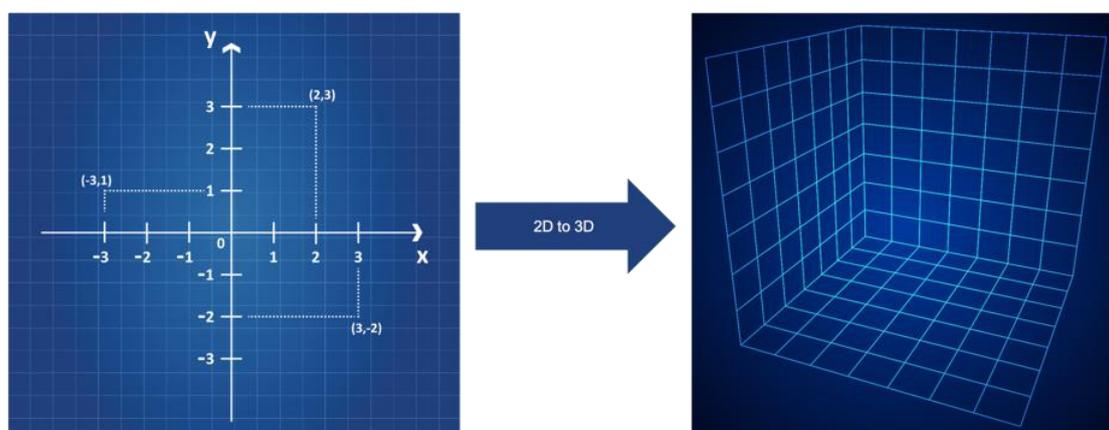


Figure 1-1 二维空间到三维空间示意图。

现行专业级或者消费级的 3D 相机采用两种主流技术，三角法(Triangulation)和飞时法(Time-of-Flight, ToF)[1]。采用三角法的 3D 视觉技术包括双目技术和结构光技术，基本原理采用三角几何视差来获得目标到相机的距离信息。这种方法在近距离有着很高的精度，但是误差会随着距离增大而快速变大。ToF 技术测量相机是指主动投射出的光束经过目标表面反射后被相机接收这个过程的来回的飞行时间，基于光速即可获得目标到相机的距离。ToF 技术在不同距离的误差相对三角法更稳定，在远距离有着更好的精度[2]。

在本文中，我们将介绍消费级的 3D 视觉技术的主要技术路径。针对 ToF 技术，我们将介绍其主要实现方法的具体工作原理，各自的优劣势以及技术挑战。根据行业的现状，我们将结合学术界的最新成果，介绍解决当前 ToF 相机痛点的一些方法。最后，我们也将结合现下行业需求，介绍 ToF 的一些重要的应用场景。

2. 3D 视觉方案介绍

常见的 3D 视觉方案主要包括双目、结构光和 ToF 三个技术方向。这三种方法各有优劣。虽然本文主要介绍的是 ToF 技术，本章节将简要地介绍和比较其他两种技术方案，帮助读者全面地了解 3D 视觉技术方案。

2.1 双目技术

双目深度重建利用的是三角测距法计算被测物体到相机的距离。具体的说，就是从两个相机观察同一物体，被观测物体在两个相机中拍摄到的图像中的位置会有一定位置差。正如将一只手指放在鼻尖前，左右眼看到的手指位置会有一个错位的效果。这个位置差称为视差，被摄物离相机越近，视差就越大；距离越远，视差就越小。在已知两个相机间距等相对位置关系的情况下，即可通过相似三角形的原理计算出被摄物到相机的距离。

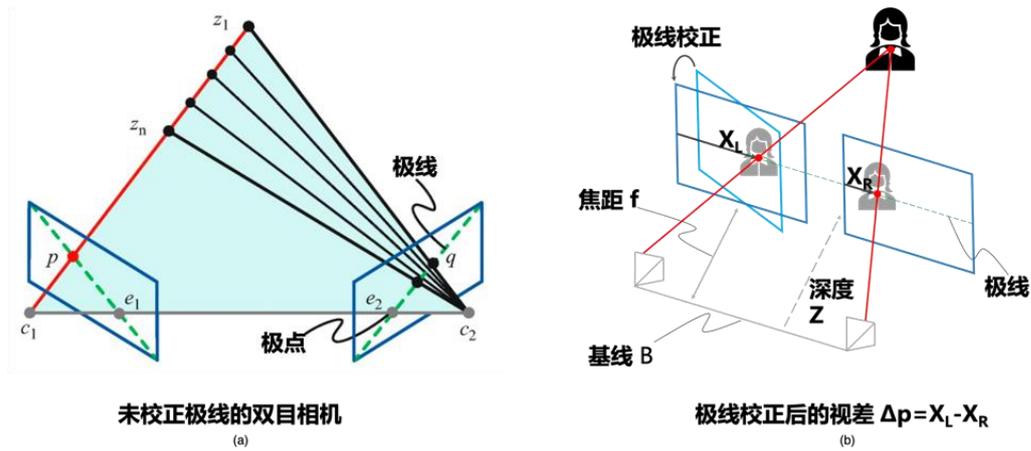


Figure 2-1 双目技术示意图。

双目深度重建的原理虽然简单，但在实际使用中遇到了两个挑战：计算量大，依赖被摄物的纹理及环境光照。下面对这两个挑战分别展开介绍。

要计算一幅图中每个像素的深度值，我们需要得到每个像素在两幅图中的一一对应关系。这个关系的建立通常是采用块匹配 (block matching) 的方法。具体的说，在一幅图中，以一个像素为中心，选取一个固定大小的窗口，在另一幅图中寻找最相似的窗口，从而得到该像素在另一幅图中的对应像素。块匹配算法有很高的计算复杂度，其计算量正比于 $O(NMWH D)$ ，其中 N, M 为图像的行数和列数， W, H 为匹配窗口的宽和高， D 为匹配寻找最相似像素的范围。为了达到更好的效果，会采用一些更复杂的改进算法（如

Semi-Global Block Matching, SGBM) ，这就更进一步提高了计算量和复杂度。鉴于此原因，业界常见的方法是将算法固化到特制的 ASIC 芯片中，从而解决计算量的大的问题，但这一增加了额外的硬件成本和迭代变化周期。

双目深度重建的另一大挑战是依赖于被拍摄物体的表面纹理和环境光照。利用双目原理重建表面没有任何纹理的物体时，例如拍摄一面白墙，会遇到无法找到匹配的对应像素的问题。另一方面，当拍摄环境的光照很弱的情况下，例如黑灯环境下，匹配也会遇到很大的挑战。结构光技术为解决这两个问题提供了新的思路。

2.2 结构光技术

结构光方案是一种主动双目视觉技术。每个结构光相机包括两个基本组件：一个红外激光投射端和一个红外摄像头。其基本思路是将已知的结构化图案投影到被观测物体上，这些结构化图案将根据物体的几何形状和拍摄距离而发生相应的形变。红外摄像头从另一个角度进行观察，通过分析观测图案与原始图案之间发生的形变，可以得到图案上各像素的视差，再根据相机的内外参恢复出深度。

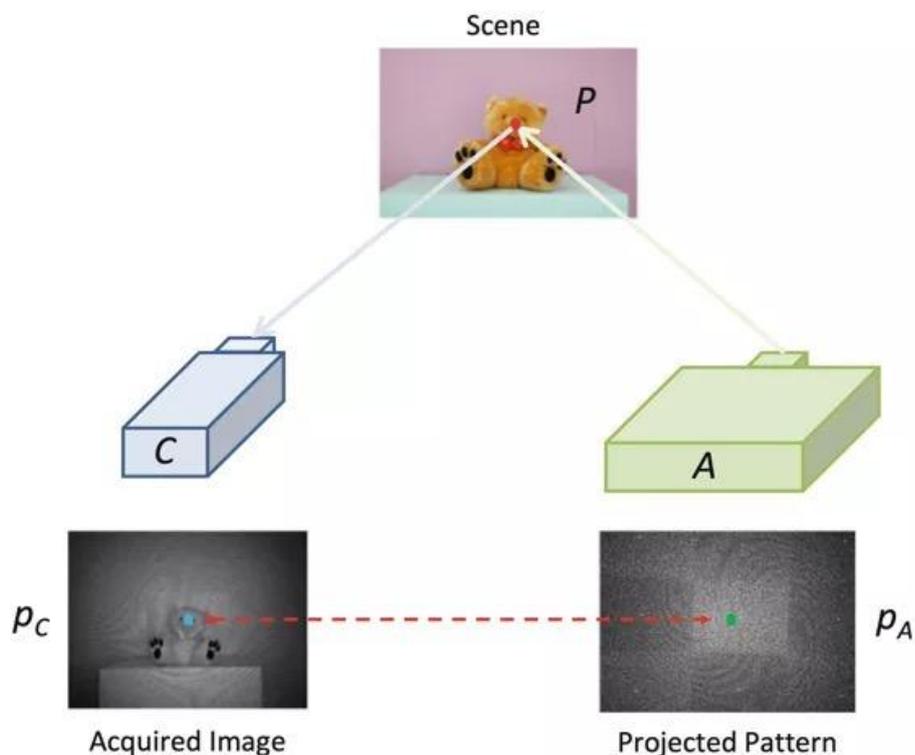


Figure 2-2 结构光技术示意图[1]。

结构光方案可以看成双目方案的一种特例。已知的投射端结构化图案和红外摄像头拍摄到的图案可视为左右双目的观测。结构光重建算法和双目重建算法采用了相似的思想，也面临着类似的挑战，主要包括高计算量和深度突变处的数据缺失。

为解决这两方面的挑战，光鉴创新地研发了一套高效软核重建算法，将计算量降低了两个数量级，只需一颗普通的 ARM 嵌入式处理器即可完成高精度深度重建。与此同时，利用多传感融合与深度学习，该算法大幅提升了常见的深度图缺失问题。

和标准的双目方案相比，结构光方案更为鲁棒，这得益于结构光方案采用的主动光源和投射的结构化图案。具体的说，投射端发出的红外激光照亮了被拍摄物体，这使得拍摄端无需依赖环境光源即可获得亮度稳定的图像输入；另一方面，投射的结构化图案为被拍摄物体增加了表面纹理，这使得拍摄表面没有任何图案的物体也能精准地重建出深度。

2.3 双目、结构光及 ToF 技术比较

为了更直观的比较双目、结构光和 ToF 技术路径的优劣势，我们汇总了各个每个方案的关键技术参数的比较。其中，i-ToF 和 d-ToF 技术将在之后的章节中具体介绍。

	双目	结构光	i-ToF	d-ToF
适用场景	近距离	近距离	中远距离	中远距离
基本原理	三角测距	三角测距	相位测距	时间测距
Sensor	RGB/IR CMOS Sensor	IR CMOS Sensor	i-ToF CIS	SPAD Array
工艺难度	容易	容易	中等	难
传感器信号	模拟	模拟	模拟	数字
发射光脉冲	无	低频率	中高频率	高频率
测量精度	近距离高，随测量距离平方下降	近距离高，随测量距离平方下降	距离呈线性关系	在工作范围内相对固定
功耗	低	中	高	中
多路径串扰	容易解决	难度适中	较难解决	容易解决
量产标定	简单	中等	难	中等

3. ToF 基本原理

相比双目视觉和结构光方案，ToF 的方案实现起来会相对简单，主要包括发射端和接收端，ToF 传感器给到光源驱动芯片调制信号，调制信号控制激光器发出高频调制的近红外光，遇到物体漫反射后，接收端通过发射光与接收光的相位差或时间差来计算深度信息。现大部分 ToF 传感器采用背照式 CMOS 工艺技术，该工艺大幅度提高了感光面积，提升了光子收集率和测距的速度，响应时间能够达到 ns 级，在远距离情况下也能保证高精度。

3.1 i-ToF 原理

i-ToF, 即 indirect ToF, 通过传感器在不同时间窗口采集到能量值的比例关系, 解析出信号相位, 间接测量发射信号和接收信号的时间差, 进而得到深度。i-ToF 根据调制方式的不同, 可分为两种: 连续波调制 (CW-iToF) 和脉冲调制 (PL-iToF), 分别发射连续的正弦信号和重复的脉冲信号; 前者是通过解析正弦信号相位解析深度, 而后者是解析脉冲信号相位来解析深度。

3.1.1 连续波调制 (CW-iToF)

通常采用正弦波调制方式, 接收和发射端正弦波的相位偏移和物体距离摄像头的距离成正比, 通过相位偏移来测量距离

$$\varphi_{TOF} = \text{atan} \left(\frac{C_1 - C_3}{C_2 - C_4} \right)$$
$$D = \frac{c}{2} * \frac{\varphi_{TOF}}{2\pi * f_m} + D_{offset}$$

相位偏移 (φ) 和 深度(D) 是由积分能量值从上述公式 C_1 、 C_2 、 C_3 、 C_4 解析得到, 这几个积分能量值, 是四个不同相位

延迟的接收窗口采集到的能量，分别对应于在相位采样点 0° 、 90° 、 180° 、 270° 采样，即：

$$C_1 = A \sin(\varphi)$$

$$C_2 = A \sin(\varphi + 90^\circ) = A \cos(\varphi)$$

$$C_3 = A \sin(\varphi + 180^\circ) = -A \sin(\varphi)$$

$$C_4 = A \sin(\varphi + 270^\circ) = -A \cos(\varphi)$$

其中 A 为接收到正弦信号的幅度。

精度方面，CW-iToF 精度主要受制于随机噪声和量化噪声，前者与接收光信号信噪比（Signal to Noise Ratio, SNR）成反比，后者与正弦波调制频率成反比。因此，为了提升精度，CW-iToF 一般采用大功率短积分时间采样，提高接受光信号 SNR；同时提高调制频率以抑制量化噪声。

量程方面，CW-iToF 可解析的相位范围为 $[0 \sim 2D_{\max} = c/(2f_m)]$ 。即频率越高，精度越高，量程也越小。超过量程的深度，将出现周期性的相位卷绕（Phase wrap），测量值错误的落在 $[0 \sim D_{\max}]$ 内。

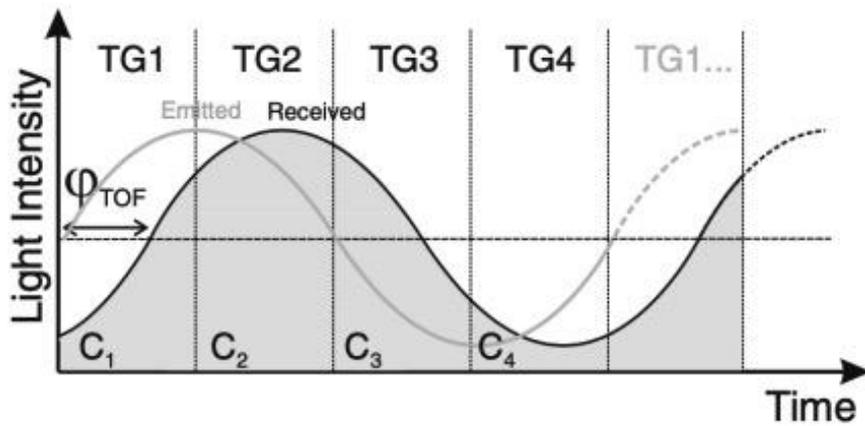


Figure 3-1 CW-iTOF 工作示意图[3]。

3.1.2 脉冲调制 (PL-iToF)

在 PL-iToF 系统中，激光光源发射带有振幅信息 A 和时间 T_P 的光脉冲，根据光的飞行速度 C ，可计算得到最远探测距离 $d_{MAX}=T_P \cdot C/2$ 。反射光信号、背景光以及探测器的噪声集成在三个不同时间段内（见 Figure 3-2）。PL-iToF 通过双采样技术提高精度，同激光脉冲同步的第一个窗口 W_0 ，同激光信号正交的第二个窗口 W_1 ，是累积反射光信号的两部分，且每个窗口与目标距离成比例；第三个窗口 W_B 在没有光脉冲发射时开启，仅收集背景光信号。如果 C_0, C_1, C_B 分别表示在窗口 W_0, W_1, W_B 的光子数，目标距离 D ，接收到的有效光强度 A_R ，背景光 B 可由以下公式得到

$$D = \frac{c * T_p}{2} * \left(\frac{C_1 - C_B}{C_0 + C_1 - 2 * C_B} \right)$$

$$A_R = \frac{C_0 + C_1 - 2 * C_B}{T_p}$$

$$B = \frac{C_B}{T_p}$$

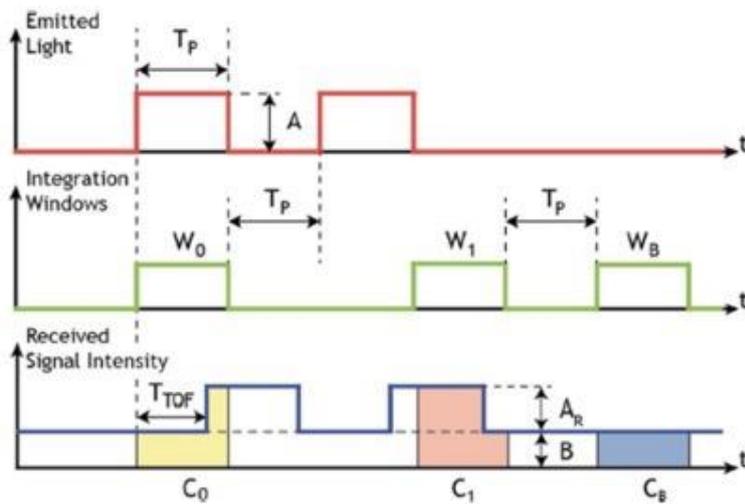


Figure 3-2 PL-iToF 的基本工作原理[4]。

3.1.3 CW-iToF 与 PL-iToF 对比

CW-iToF 在工作过程中，不论目标物体的距离是多少，系统都采集了完整时长的反射光。相比之下，PL-iToF 在两个窗口内采集的信号的信噪比与距离直接相关。在有背景噪声的情况下，如果目标距离的很近， W_1 窗口的能量几乎为零，因此， W_1 信噪比非常差；类似的，在较远的距离， W_0 中的信号很弱，导致 W_0 的信噪比差。这种效应会导致 PL-iToF 在近和远距离都有比较大的误差。

相比 CW-iToF 连续波调制方式，PL-iToF 解算深度更简单、计算量更低，对于平台后端处理能力要求也相应更低。然而，PL-iToF 的精度取决于发光次数，发光次数越多，精度越高，但同时也会带来功耗的增加。即使在相同平均功耗的情况下，PL-iToF 不仅精度弱于 CW-iToF，而且对于背景噪声和暗噪声更加敏感[5]。

因此，现下的主要手机厂商，包括华为、三星、Oppo 等，以及 ToF 芯片厂商，包括索尼，三星，英飞凌等都采用了 CW-iToF 的方案。

3.2 d-ToF 原理

d-ToF 即 direct ToF，相比于 i-ToF 技术用测量信号的相位来间接地获得光的来回飞行时间，d-ToF (direct time-of-flight) 技术直接测量光脉冲的发射和接收的时间差。由于激光安全的限制以及消费类产品的功耗限制，ToF 相机发射的脉冲能量有限，但是需要覆盖完整的视场区域。光脉冲在经过反射回到接收器时，能量密度降低了超过一万亿倍。于此同时，环境光作为噪声，会干扰接收器对于信号的检测和还原。在这种情况下，探测器获取的信噪比不足以直接还原脉冲的模拟信号，进而导致直接测量深度存在很大

的误差。因此，d-ToF 方法需要有灵敏度极高的光探测器来检测微弱的光信号。

单光子雪崩二极管(Single Photon Avalanche Diode, SPAD)具有探测单个光子的灵敏度。SPAD 在工作状态是一个偏置了高反向电压的二极管。反向偏压在器件内部形成了一个强大的电场。当一个光子被 SPAD 吸收转化为一个自由电子时，这个自由电子被内部的电场加速，获得足够的能量撞击其他原子时产生自由电子和空穴对。而新产生的载流子继续被电场加速，撞击产生更多的载流子。这种几何放大的雪崩效应使得 SPAD 具有几乎无穷大的增益，从而输出一个大电流脉冲[6]，实现对于单个光子的探测。

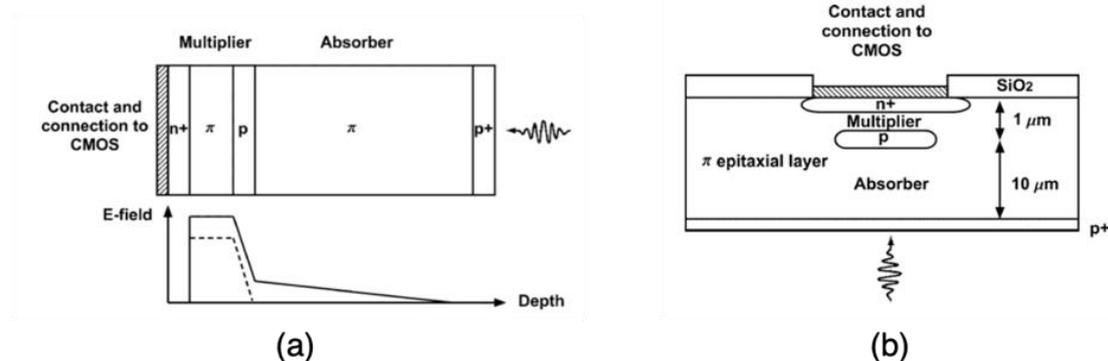


Figure 3-3 SPAD 雪崩效应示意图. (a)雪崩二极管示意图，光子在吸收区被吸收，转化为自由电子。自由电子在穿过 PN 结的过程中被电场加速。在获得足够能量时，在放大区产生雪崩效应，使得器件输出一个大电流脉冲。(b)SPAD 的 CMOS 示意图。具体的、结构参数取决于器件采用的 CMOS 工艺。 [7]

d-ToF 技术采用 SPAD 来实现高灵敏度的光探测, 并且采用时间相关单光子技术方法 (Time-Correlated Single-Photon Counting, TCSPC) 来实现皮秒级的时间精度[8]。光脉冲的第一个被 SPAD 捕获的光子即可出发 SPAD, 产生电流脉冲信号。系统的时间数字转换器 (Time-to-Digital Converter, TDC) 可以转换这个电流脉冲相对于发射时间的延时。SPAD 捕获一段脉冲内哪一个瞬间到达的光子具有一定的随机性, 这种随机性的概率与光脉冲在该瞬间的能量近似成正比。因此, d-ToF 相机重复很多次 (比如数千次) 发射和探测相同的脉冲信号即可获得每次探测的电流脉冲延时的统计分布。这个统计直方图即恢复了发射脉冲能量随着时间的变化, 进而得到了脉冲来回的飞行时间。

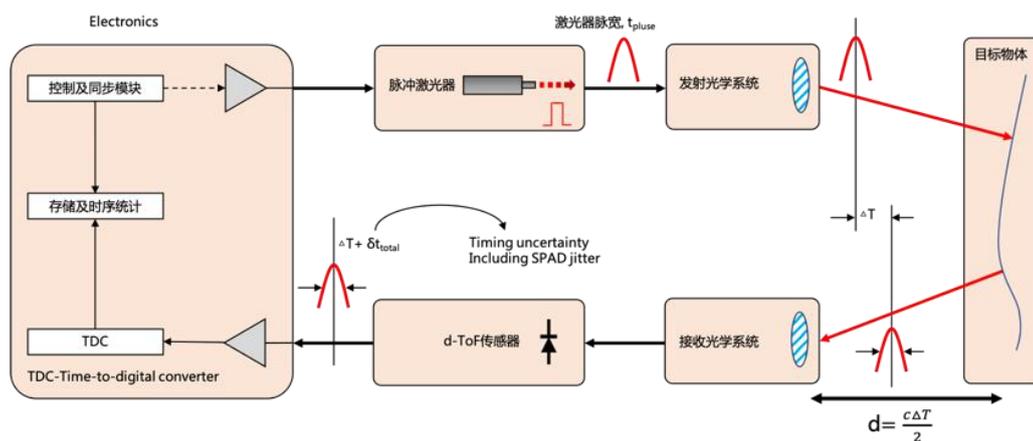


Figure 3-4 TCSPC 方法. 系统控制激光器发射出激光脉冲, 通过光学系统投射到目标物体表面。反射回的光脉冲被接受器的光学系统成像到

d-ToF 传感器上。光脉冲触发 SPAD，输出电流脉冲。TDC 根据电流脉冲的时间来输出数字化的脉冲时序。一次成像会重复几千到几十万次的脉冲，从而获得 TDC 输出的统计直方图，重建光脉冲及获得飞行时间。[9]

4. ToF 技术挑战

4.1 i-ToF 挑战

在实际应用中，i-ToF 技术面临着诸多的挑战，真实环境的复杂多变，给深度测量引入了大量的干扰和噪声。这也是 i-ToF 技术提出已经有数十年的时间，但实际应用却十分有限的主要原因。本章节对 i-ToF 技术面临的诸多问题进行简要的原理定性分析。

4.1.1 飞点噪声 (Flying pixels)

在 i-ToF 测量的深度图中，物体边缘处往往存在大量错误的深度测量值，生成 3D 点云后，视觉上表现为飞在空中的无效点（如 Figure 4-1 所示），称为飞点噪声。飞点噪声使得 i-ToF 无法有效获取物体边缘的 3D 信息，这也是当下 i-ToF 能否得到广泛应用的一大挑战。

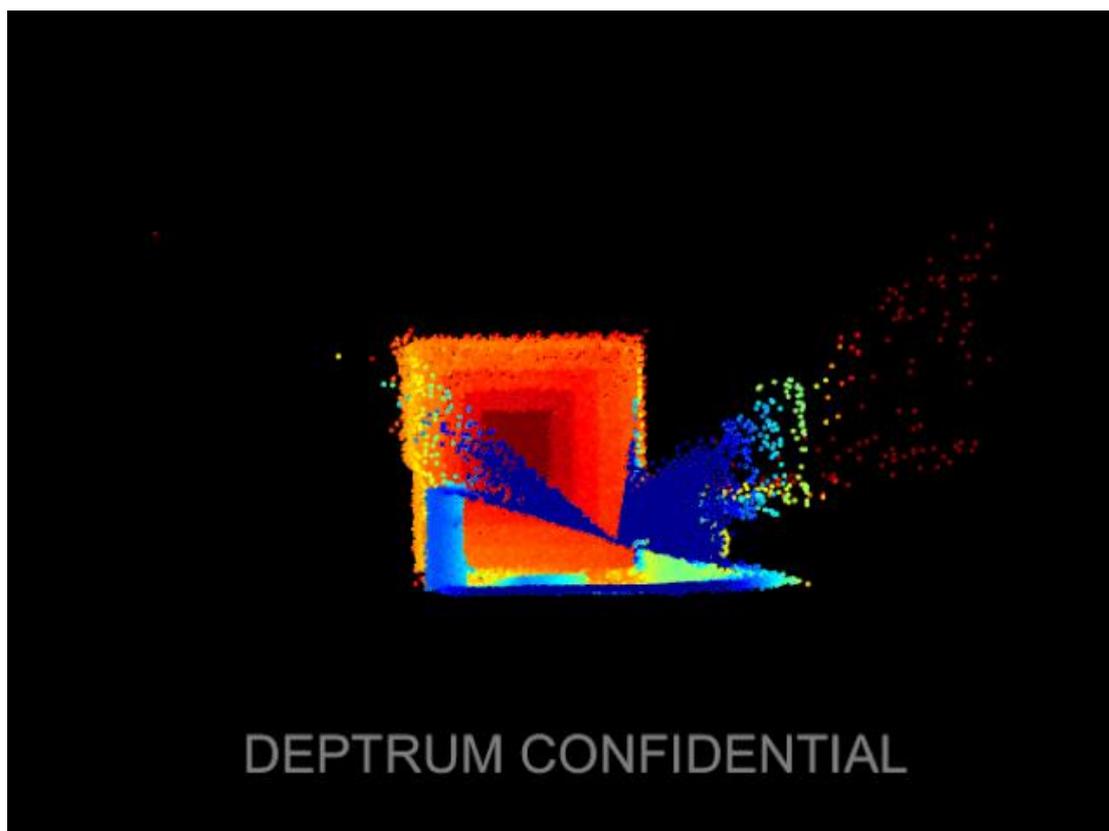


Figure 4-1 典型的 i-ToF 测量点云，边缘处存在飞点噪声

如 Figure 4-2 所示，飞点噪声产生的主要原因在于：i-ToF 传感器上，每个像素都具有一定的物理尺寸，在测量物体边缘时，单个像素会同时接收到前景和背景反射回来的光线；二者产生的能量叠加在一起，使得传感器获取的原始数据中包含多个距离的信息，使用 3.1 节原理解算相位计算距离时将得到错误的深度测量值。另外，镜头散射及像素间串扰，有时也会引起飞点噪声，甚至造成背景的大范围变形[10, 11]。

通过边缘检测等图像算法[12],可以在一定程度上检测并去除边缘飞点噪声,但对散射和串扰引起的变形难以根除,同时,误检也会造成大量有效深度测量值的丢失。

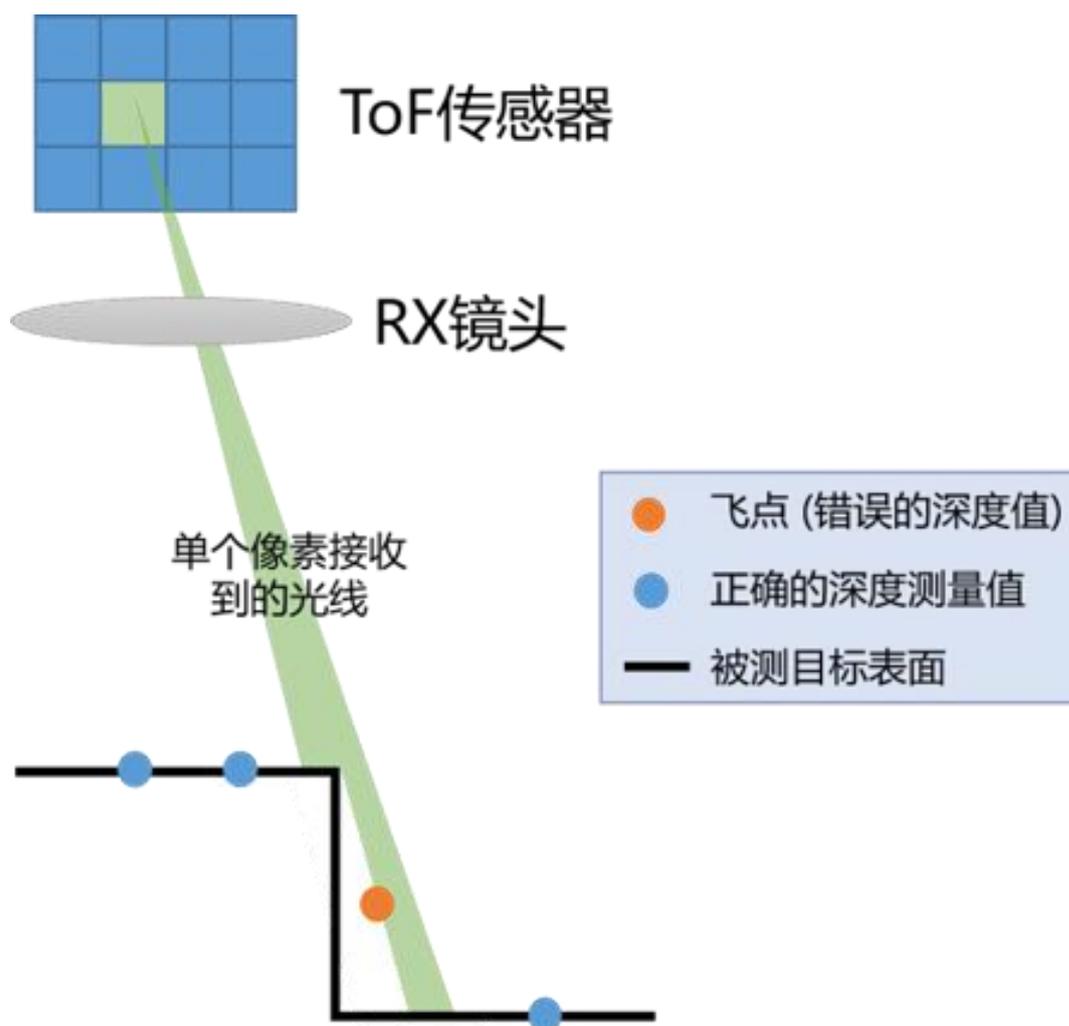


Figure 4-2 飞点噪声产生原理示意图: ToF 传感器上的蓝色像素仅接收到单一深度信息(前景 or 背景), 可以获取正确的测量值(蓝色点); ToF 传感器上的绿色像素同时接收到前景和背景反射的光线,两个深度信息叠加在一起无法区分, iToF 测量得到错误的深度值(橙色点)。

4.1.2 多径干扰 (Multi-Path Interference, MPI)

真实场景中存在复杂的漫反射甚至镜面反射，MPI 在原理上会使得测量值变大，严重影响三维重建的效果。

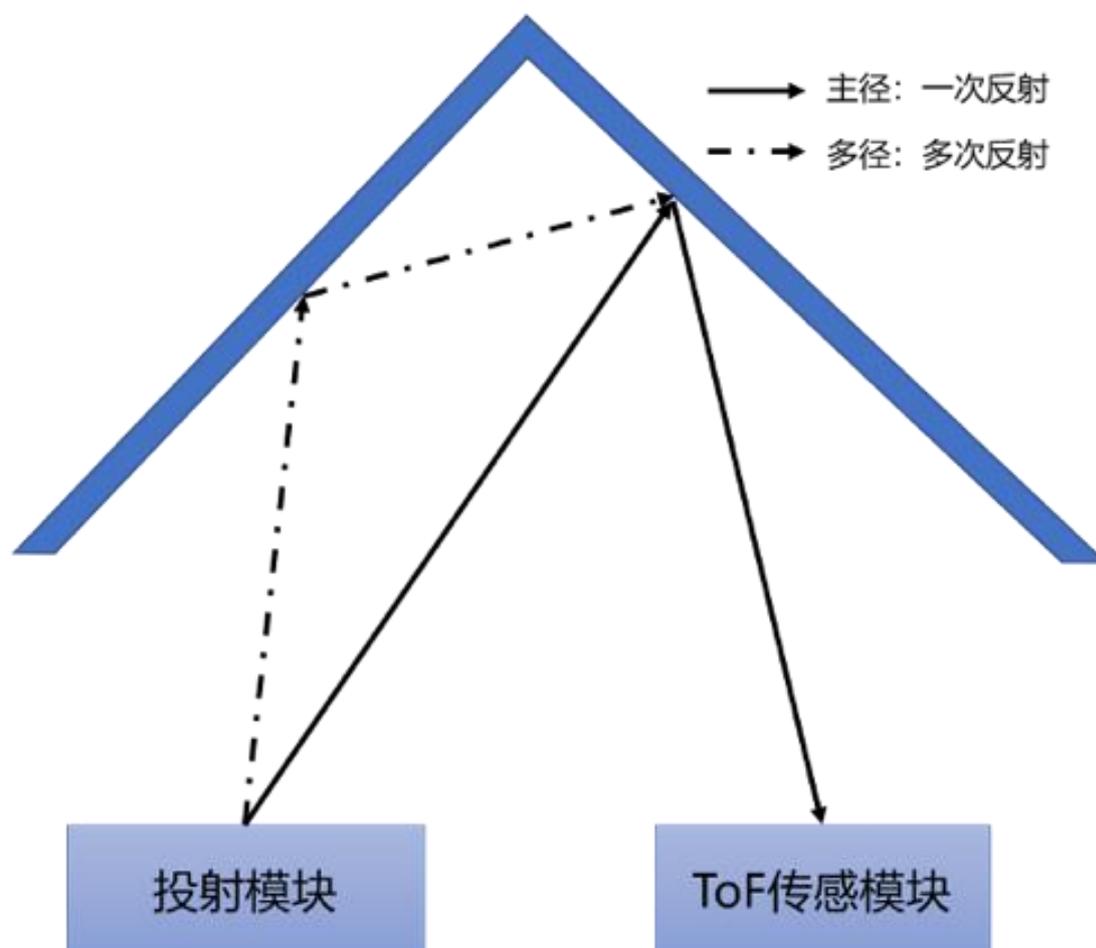


Figure 4-3 多径干扰产生原理示意图：图示以测量墙角为例,投射模块投向左边的光线(虚线)经两次反射,与投向右边的光线(实线)同时被 ToF 传感器接收到。双重的深度信息导致 ToF 测量值错误。

以 Figure 4-4 的场景为例,投向桌面的光线经标准件二次反射后被 i-ToF 传感器接收到, MPI 效应导致测量到的标准件

形状扭曲；投向标准件的光线经桌面二次反射后被 i-ToF 传感器接收到，MPI 效应导致桌面测量值的错误，桌面近似于一个镜面，桌面测量值接近于标准件的镜像。

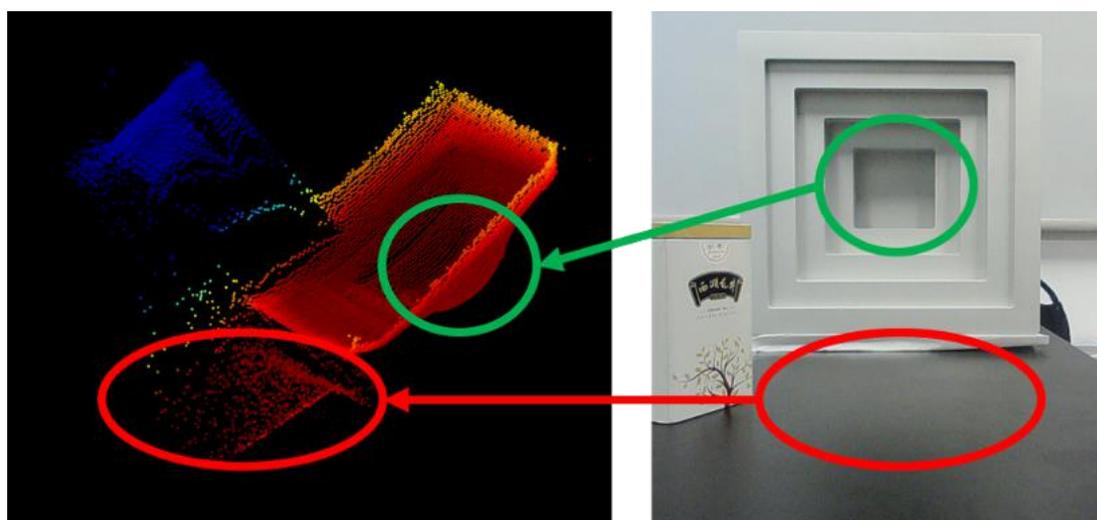


Figure 4-4 多径干扰示意图：多径干扰导致标准件测量点云形状扭曲（绿色），以及桌面错误地测量成标准件镜像（红色）

MPI 是困扰 i-ToF 多年的重要问题，一直是 i-ToF 广泛应用的障碍。在过去的十年中，微软，MIT，Waikato 大学等诸多研究机构在解决 MPI 问题上做出了大量算法和系统层面的尝试[13]，但仍无法根除该问题。

4.1.3 强度误差(Intensity Related Error)

在 i-ToF 传感器测量到的深度图上，存在一类特殊的误差，即同一平面上不同反射率的区域体现出不同的深度，如 Figure 4-5 所示。



Figure 4-5 强度误差示意图，同一平面上不同反射率区域呈现不同深度，黑色和灰色区域从平面上凸起[14]。

i-ToF 的强度误差与距离、反射率、积分时间等因素都存在关联，究其产生的原理，就笔者了解所限，误差的解析尚未完全明确。PMD 等研究机构在该问题上做出了一些分析和尝试[14]，能够缓解强度误差，但难以在全场景消除强度误差的影响。

4.1.4 远距离-高精度矛盾(Trade-off between range and precision)

在 i-ToF 的两种类型上，量程和精度都存在典型的矛盾：

CW-iToF：调制频率决定量程，频率越低量程越远；同样的相位解析精度下，深度测量精度随频率降低而降低；

PL-iToF：脉冲宽度决定量程，脉宽越大量程越远；同样的相位解析精度下，深度测量精度随脉宽增加而降低；

同时，i-ToF 往往采用泛光投射，传感器探测到的能量随距离的平方快速衰减，远距离测量的信噪比极差，进一步恶化上述矛盾。远距离和高精度这一对矛盾，在原理上很难调和，通常需要根据实际应用进行权衡，选取最合理的模式配置。

4.1.5 高频驱动

i-ToF 投射端需要特定驱动芯片（driver IC）驱动激光器发出调制的光信号。一般情况下，为了保证测量精度，CW-iToF 采用提高调制频率的方式，PL-iToF 则采用窄脉冲高功率的驱动方式。综合起来，iToF 对于驱动芯片的主要需求是高调制频率和高功率；同时，驱动芯片的温度系数与 i-ToF 测量的温漂紧密相关，需要尽可能保证线性。这些需求对芯片工艺，尤其是 CMOS 工艺有着较高的要求。

4.1.6 片上集成

片上集成对于 i-ToF 芯片的设计提出了较高的要求。一方面，为保证足够的探测灵敏度和测量动态范围，i-ToF 往往需要有足够的像素尺寸；另一方面，i-ToF 芯片相对于普通图像传感器，增加更复杂的时序控制电路和相关采样电路，整体集成难度更大。

而消费电子行业，尤其是手机等产品，对芯片有着很高的尺寸限制。在上述要求之下，i-ToF 芯片很难集成较高的分辨率，目前市面上主流的 i-ToF 传感器像素一般在 QVGA (320x240) 上下；近几年开始，逐步有 VGA 分辨率的 i-ToF 传感器进入市场，但其像素尺寸一般不大于 7 μm ，且性能上会有一定折扣。

4.2 d-ToF 挑战

d-ToF 技术的误差在正常工作范围内不随距离变化，并且受到多径等因素的干扰较小。在远距离、复杂环境的应用有着优势。然而，d-ToF 的技术成熟面临着一系列的挑战，需要在芯片设计、系统设计、制造工艺等方面全面突破才能真正兑现 d-ToF 技术承诺的优势，并实现在消费场景的普及。本章节将从原理上分析 d-ToF 技术存在的技术挑战及优化方向。

4.2.1 暗计数率(Dark Count Rate, DCR)

相比于传统的摄像头图像传感器(Camera Image Sensor, CIS), SPAD 输出的是数字化的脉冲, 因此受到电子噪声的影响较小。但是由于在雪崩区域内出现的单个自由电子即可触发计数, 从而导致错误的计数, SPAD 受到暗噪声的影响较大。

暗计数率 DCR 的主要来源包括探测器内由于热产生的自由电子。由于制造工艺和掺杂工艺, 器件内部存在 Shockley-Read-Hall(SRH)缺陷, 释放和捕获载流子。在尺寸为深亚微米(Deep sub-micron, DSM)的 CMOS 工艺下, 电压下降要求更高的掺杂浓度以及更小的 PN 结尺度。这导致更高的缺陷密度以及更强的加速电场, 使得暗计数效应更为严重。相比于高电压的定制工艺, DSM 的 CMOS 工艺暗计数率要高几个数量级[6]。因此, DCR 主要取决于制造工艺的特性和优化。

此外, 在雪崩过程中, 存在自由电子被缺陷捕获的情况。这些被捕获的电子基于 SRH 统计速度被重新释放, 制造 SPAD 的暗计数。这被称为 AP(after-pulsing)现象。这种现象可以通过在 SPAD 配合的主动猝熄电路(quenching circuit)来设置一个适合的关闭时间(hold-off time)来解决。在 SPAD

被触发输出脉冲后，保持一段时间不开启接收新的光子触发，让这些被捕获的电子有足够时间重新释放而不会重新引起雪崩。这个关闭时间通常需要几十到几百纳秒。这段时间被称为 SPAD 两次探测状态之间的空滞时间 (dead time)[15]。空滞时间成为了单位时间内重复测量的脉冲次数的限制。

通常更大尺寸的工艺有更低的 DCR，这与 CMOS 集成的要求是相悖的。特别对于 SPAD 阵列，每个像素的 SPAD 都要配合一个独立的淬灭电路，大尺寸的 CMOS 工艺会导致淬灭电路占据像素可观的面积，并且产生更高的功耗。而手机等产品有着很高的尺寸限制，VGA 像素的 ToF 相机的像素尺寸不大于 7 μm 。因此，3D 集成工艺是 SPAD 阵列优化暗计数率和感光率的必然要求。即，用大尺寸的制造工艺制作 SPAD 的感光部分，用更小尺寸的制造工艺制作高集成度的辅助电路，然后将不同工艺模块 3D 堆叠。这对 SPAD 阵列的制造提出了更高的要求。

4.2.2 光子探测效率(Photon Detection Efficiency, PDE)

光子探测效率是另一个重要的技术挑战。光探测效率是感光面积的占比 FF(Filling factor)与吸收率以及雪崩触发率的乘积。

对于尺寸受限的 SPAD 阵列，每个像素的面积尺寸非常有限。这些有限的面积也无法完全用来感光。首先，每个像素之间需要设置保护区域，用于防止像素间的串扰。此外，淬灭电路也将占据可观的面积。另外，对于 FSI(Front Side Illumination)工艺，用于吸收光子的厚度非常有限，限制了光子的吸收率。

因此，采用 3D 集成和 BSI 工艺将大幅提升 PDE。BSI 工艺用晶圆的背面接受照射作为光吸收层，吸收厚度得到了数量级的提升。此外，每个像素的辅助电路和光吸收不再共享同一个表面积，因而大幅提升了 FF[16]。

提升 PN 结偏置电压可以帮助提升 PDE，但是高电压会带来更高的功耗和发热以及更高的 DCR。针对于具体的应用场景，偏置电压可以作为权衡优化具体某个指标的杠杆。

4.2.3 串扰(Cross talk)

采用 CMOS 工艺的 SPAD 阵列有共享的电极，帮助提升集成度。但是，一个像素的自由电子被加速后，有可能渗透到相邻的像素，进而出发相邻像素触发计数，造成图像的模糊。传统的 CIS 也有串扰效应，但是每个像素的自由电子没有被强电场加速，所以渗透到其他像素的情况要远少于 SPAD。

在 CMOS 设计中, 护环(guard ring)被用于防止不同像素的自由电子的串扰。护环有多种设计方法, 包括采用 STI(Shallow Trench Isolation)和 buried n-well 等。这些护环的方法和工艺在集成电路设计中都很成熟。但是, 由于护环本身占据面积, 并且其掺杂会带来附近区域更高的缺陷, 因此会影响 SPAD 的 DCR 和 PDE。

4.2.4 时间精度

d-ToF 要求亚厘米级或厘米级的测量精度, 考虑到光速为每秒 30 万公里, 对应的时间精度要求为皮秒级。时间测量误差主要来源于系统的时序抖动(timing jitter)。激光发射器、系统电路等部分的 jitter 有很多优化的办法。SPAD 中自由电子被加速和渗透的过程是一个随机过程, 渗透时间必然存在 jitter, 是 d-ToF 系统的时间精度的极限。常见的 jitter 在 100ps 左右, 对应于 1cm 左右的误差[17]。更薄的吸收层和有源层可以减小 jitter, 但如之前 PDE 的分析所述, 降低吸收层厚度会降低 PDE。

4.2.5 SPAD 阵列 3D 集成

为了有足够的重复次数来获得时序的统计, 一帧深度图需要有几千到几十万次的统计数据来实现。SPAD 阵列的每个像

素都在以 1MHz 左右的速度重复测量脉冲,而每次触发, TDC 都会产生一个时间的数字输出。以 100X100 像素的 SPAD 阵列为例,数据量就达到了几十 Gbps。而 VGA 像素的 SPAD 阵列则将产生 1Tbps 以上的数据流。在芯片以外去处理这样的数据量是不现实的,要求有 25 条以上的 40Gbps 的 PCIE 通道和几瓦甚至 10 瓦以上的功耗。所以,提高 d-ToF 的像素就需要在 d-ToF 芯片上集成存储和数字处理的能力。芯片需要能够存储每一帧测量过程中的每个像素和每个脉冲对应的 TDC 的数据,在一帧测量完成后,进行数据处理,计算出每个像素的时序统计,然后输出其来回的飞行时间。

片上集成对于 d-ToF 芯片的设计提出了非常高的要求。在一个芯片上需要堆叠用于光探测的 SPAD、淬灭电路, TDC、存储单元以及运算单元。这要求设计团队同时具有 SPAD 器件的设计能力和 SoC 的设计能力。此外,消费类的应用要求芯片的功耗控制在几百毫瓦以内,否则功耗和散热都将成为应用的瓶颈。在这些复杂度之上,需要有足够好的良率,以保证个位数美元的芯片单价。

5. ToF 发展方向

当前 ToF 行业现状, d-ToF 技术在激光功耗、抗干扰、远距离精度等方面有明显优势,但在工艺和产业链均离成熟尚

远，仍需较长时间打磨；i-ToF 芯片在工艺和产业链虽已趋于成熟，但达到的效果却不尽完美，从而导致其应用受阻。

随着 2020 年发布的 iPad Pro 等高端消费电子领域的持续关注，d-ToF 技术将进入快速迭代发展阶段，技术发展方向可能会集中在：SPAD 工艺升级（包括 DCR、PDE、jitter 等），片上集成度提升（包括片上直方图/深度图算法，I/O，Memory 等），TRX 系统协同设计等方面；随着工艺和产业链的成熟，d-ToF 的技术优势也会逐步释放，占据一定市场空间。

与此同时，i-ToF 仍有很大潜力可以持续挖掘，不论是在算法端，亦或是系统端和应用端均有望通过软硬件的协同设计，弥补原理上的非理想效应[10]。以光鉴科技的 mToF (modulated ToF) 方案为例，通过在系统端结合软硬件，引入调制光场的概念，通过空域、频域、时域上的巧妙设计，创新硬件协同前沿算法，在物理上提升 i-ToF 抗干扰、抗噪声能力，解决 i-ToF 在实际应用场景中面临的关键痛点，一定程度上可以媲美 d-ToF 的性能。

综上所述，我们认为：在 d-ToF 产业链成熟之前，i-ToF 还有很大的潜力可以挖掘，有望先一步抢占 3D 行业市场份额；而随着工艺和产业链的成熟，d-ToF 将逐步从高端消费

电子往下渗透，在较长的时间周期中，与 i-ToF 平分秋色，各自占据重要的市场份额。在 d-ToF 方案成熟之后，i-ToF 在像素、成本等方面有着优势，而 d-ToF 在功耗、距离以及抗干扰方面有着优势。而不论基于何种技术路线，ToF 系统的成像芯片只能解决如何探测和处理返回的光信号；而作为一个 3D 成像系统，光学系统的设计、投射光的调制和控制、图像数据的算法处理等因素也将决定了一种技术方案是否能够充分发挥出其原理的优势，实现真正适合应用需求的方案。

6. ToF 技术应用

ToF 的精度取决于其脉冲持续时间，相比双目视觉、结构光方案，ToF 精度不会随着距离增长而显著降低，d-ToF 是远距离应用的关键技术。

随着 2020 年苹果 iPad Pro 的发布，采用了激光雷达扫描仪 ToF 传感器，势必会带动 ToF 在消费类电子应用的进一步爆发。目前消费电子中 ToF 应用以手机为主，华为、三星已在前后摄都搭载 ToF 摄像头，今年苹果机型有望也开始搭载 ToF 技术。

3D ToF 技术在其他领域应用也开始逐步渗透，目前还是主要依赖头部终端厂商的推动，主要的应用领域包括以下场景：

01 消费电子



(a) 人脸识别



(b) 体感互动



(c) 手势识别



(d) AR

Figure 6-1 ToF 在消费类电子领域应用: (a) ToF 体积小, 在对于精度要求不高的场景下可以用于简单的人脸活体识别 (b) 3D 感应人体关键部位, 体感互动游戏 (c) 追踪手部位置和姿势, 进行手势控制 (d) 构建三维信息, 虚拟与真实环境进行交互

02 机器人



(a) AGV避障



(b) vSLAM定位建图



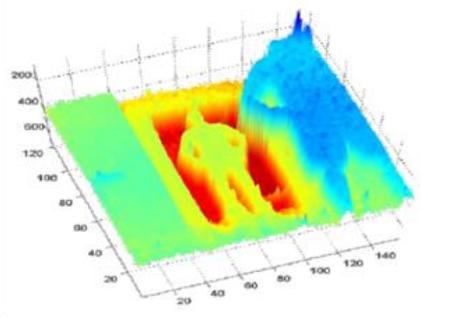
(c) 机器人导航



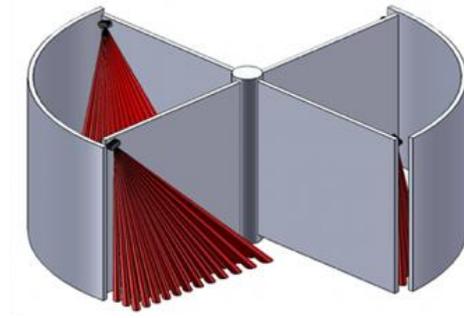
(d) 无人机定高

Figure 6-2 ToF 在机器人领域应用 (a) ToF 低速激光雷达可精确识别障碍物，进行自动避障 (b) 测量得到周围环境深度信息，定位自身位置构建地图 (c) 应用于服务型机器人，智能导航 (d) 无人机得到 ToF 稳定、精准的距离信息定高悬停

03 安防监控&轨道交通



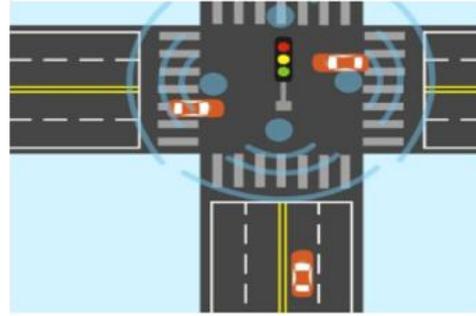
(a) 人数统计



(b) 安全自动



(c) 智能停车场



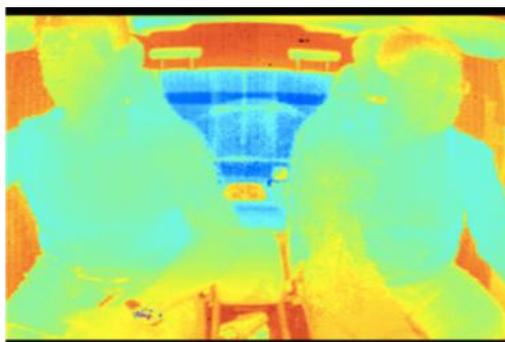
(d) 智能交通

Figure 6-3 ToF 在安防监控&轨道交通领域应用: (a) ToF 获得人体深度数据, 结合人数统计算法, 相比传统监控可实时统计、跟踪人员数量 (b) 通过智能方向识别忽略交错人流 (c)智能停车, 广覆盖精准识别车位车辆信息 (d) 实时监控路口车辆。增加监控环境三维信息

04 无人驾驶&工业自动化



(a) 无人驾驶



(b) 车厢检测



(c) 智能仓储



(d) 体积测量

Figure 6-4 ToF在无人驾驶&工业自动化领域应用:(a)随着面阵 dToF 工艺的成熟,未来纯固态激光雷达将与其他雷达相融合用于无人驾驶中 (b) 车厢监控,监测驾驶员疲劳状态,监控车内人员情况 (c) 仓储分拣,智能识别货物信息 (d) 物流包裹体积测量,可快速识别包裹长宽高

7. 总结

本文介绍了 ToF 深度相机的基本工作原理和不同技术路径的优势和挑战。我们还比较了 ToF 与双目和结构光技术方案的优势劣势。ToF 技术的成熟将带来其在消费电子、机器人、工业自动化、物流等领域的大量应用和突破。

如果有关于光鉴科技的 3D 视觉产品的问题，请联系 info@deptrum.com 。

8. 参考文献

- [1] P. Zanuttigh, G. Marin, C. Dal Mutto, F. Dominio, L. Minto, and G. M. Cortelazzo, Time-of-Flight and Structured Light Depth Cameras. Springer, 2016.
- [2] H. Sarbolandi, D. Lefloch, and A. Kolb, "Kinect range sensing: Structured-light versus Time-of-Flight Kinect," Computer Vision and Image Understanding, vol.139,pp.1–20,2015,doi:10.1016/j.cviu.2015.05.006.
- [3] D. S. Fabio Remondino, ToF Range-Imaging Cameras. Springer, 2013.
- [4] H. Sarbolandi, M. Plack, and A. Kolb, "Pulse Based Time-of-Flight Range Sensing," Sensors (Basel), vol. 18, no. 6, May 23 2018, doi: 10.3390/s18061679.
- [5] D. Bronzi, Y. Zou, F. Villa, S. Tisa, A. Tosi, and F. Zappa, "Automotive Three-Dimensional Vision Through a Single-Photon Counting SPAD Camera," IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 3, pp. 782–795, 2016, doi: 10.1109/TITS.2015.2482601.

- [6] D. P. Palubiak and M. J. Deen, "CMOS SPADs: Design Issues and Research Challenges for Detectors, Circuits, and Arrays," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 6, pp.409–426, 2014, doi: 10.1109/jstqe.2014.2344034.
- [7] B. F. Aull, E. K. Duerr, J. P. Frechette, K. A. McIntosh, D. R. Schuette, and R. D. Younger, "Large-Format Geiger-Mode Avalanche Photodiode Arrays and Readout Circuits," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 2, pp. 1–10, 2018, doi: 10.1109/jstqe.2017.2736440.
- [8] J. S. Massa, G. S. Buller, A. C. Walker, S. Cova, M. Umasuthan, and A. M. Wallace, "Time-of-Flight Optical Ranging System Based on Time-Correlated Single-Photon Counting," *Appl. Opt.*, vol. 37, no. 31, pp.7298–304, Nov 1 1998, doi: 10.1364/ao.37.007298.
- [9] P. Padmanabhan, C. Zhang, and E. Charbon, "Modeling and Analysis of a Direct Time-of-Flight Sensor Architecture for LiDAR Applications," *Sensors (Basel)*, vol. 19, no. 24, Dec 11 2019, doi: 10.3390/s19245464.

- [10] Y. He and S. Chen, "Recent Advances in 3D Data Acquisition and Processing by Time-of-Flight Camera," *IEEE Access*, vol. 7, pp. 12495–12510, 2019.
- [11] A. Sabov and J. Krüger, "Identification and correction of flying pixels in range camera data," in *Proceedings of the 24th Spring Conference on Computer Graphics*, 2008, pp. 135–142.
- [12] M. Reynolds, J. Doboš, L. Peel, T. Weyrich, and G. J. Brostow, "Capturing time-of-flight data with confidence," in *CVPR 2011*, 2011: IEEE, pp. 945–952.
- [13] R. Whyte, L. Streeter, M. J. Cree, and A. A. Dorrington, "Review of methods for resolving multi-path interference in time-of-flight range cameras," in *SENSORS*, 2014 IEEE, 2014: IEEE, pp. 629–632.
- [14] M. Lindner, I. Schiller, A. Kolb, and R. Koch, "Time-of-flight sensor calibration for accurate range sensing," *Computer Vision and Image Understanding*, vol. 114, no. 12, pp. 1318–1328, 2010.
- [15] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection," *Appl. Opt.*, vol. 35, no. 12, pp. 1956–1976, 1996.

[16] T. Al Abbas, N. Dutton, O. Almer, S. Pellegrini, Y. Henrion, and R. Henderson, "Backside illuminated SPAD image sensor with 7.83 μm pitch in 3D-stacked CMOS technology," in 2016 IEEE International Electron Devices Meeting (IEDM), 2016: IEEE, pp. 8.1. 1–8.1. 4.

[17] H. Xu, L. Pancheri, G. D. Betta, and D. Stoppa, "Design and characterization of a p+/n-well SPAD array in 150nm CMOS process," *Opt Express*, vol. 25, no. 11, pp. 12765–12778, May 29 2017, doi: 10.1364/OE.25.012765.

本文为深圳市光鉴科技有限公司原创并依法对其享有完整的著作权。未经书面授权，任何个人、机构或其它主体不得编辑、修改、复制、转载、摘编或另作其它商业使用。本公司亦不接受任何主体在未经书面授权而作出的类似“本文来源于网络，版权归原作者”之一切表达。若需获得书面授权使用者，请联系我司并支付相应报酬。对不遵守本声明或其他违法、恶意使用本文任何内容者，我司保留追究其法律责任的权利。