

用于3D立体视觉测距技术原理详解和性能比较

莫良华¹ 李孝剑² 朱小军³ 王东生⁴ 李 威² 田金良² 杨 明⁵

1. 深圳阜时科技有限公司 广东深圳 518003
2. 深圳市大鱼智行科技有限公司 广东深圳 518109
3. 深圳市英卡科技有限公司 广东深圳 518000
4. 深圳盈达信息科技有限公司 广东深圳 518052
5. 深圳大学 广东深圳 518060

摘 要: 3D 立体视觉作为前沿的深度信息技术, 对未来机器人, 无人驾驶, 工业4.0等新兴的应用意义重大。3D 立体视觉以测距技术为基础, 配合空间方位角信息对整个空间进行成像。因此, 测距的精度, 准度, 抗噪能力尤为重要。按照测距原理的不同, 可将测距技术分为三角关系测距和飞行时间测距两大类; 按照测距距离的不同, 可以分为近距离和远距离。本文从测距原理展开, 并针对各种测距技术的性能进行对比, 以明确3D 立体视觉未来的技术路线发展。

关键词: 3D 立体视觉; 测距技术; 比较

Principle explanation and performance comparison of 3D stereo vision ranging technology

Lianghua Mo¹, Xiaojian Li², Xiaojun Zhu³, Dongsheng Wang⁴, Wei Li², Jinliang Tian², Ming Yang⁵

1. Shenzhen Fushi Technology Co., Ltd. Shenzhen, Guangdong 518003
2. Shenzhen Dayu Zhixing Technology Co., Ltd. Shenzhen, Guangdong 518109
3. Shenzhen yingka Technology Co., Ltd. Guangdong Shenzhen 518000
4. Shenzhen Yingda Information Technology Co., Ltd. Shenzhen, Guangdong 518052
5. Shenzhen University Shenzhen, Guangdong 518060

Abstract: as a cutting-edge depth information technology, 3D stereo vision is of great significance for emerging applications such as robots, driverless and industry 4.0 in the future. 3D stereo vision is based on distance measurement technology and coordinates with spatial azimuth information to image the whole space. Therefore, the accuracy, accuracy and anti noise ability of ranging are particularly important. According to different ranging principles, ranging techniques can be divided into two categories: triangular relationship ranging and time of flight ranging; According to the difference of ranging distance, it can be divided into short distance and long distance. This paper starts from the ranging principle, and compares the performance of various ranging technologies, so as to clarify the future technical route of 3D stereo vision.

Keywords: 3D stereovision; Ranging technology; compare

引言:

过去十年间可穿戴设备, 智能家电, 自动驾驶汽车, 机器人等新兴产业的高速发展, 而这些领域都在向模仿甚至是超越人类自身的智慧和效率这一终极目标而努力。虽然, 凭借人工智能, 大数据, 云计算, 无线, 惯性导航, 计算机视觉, 语音识别等方面的飞速发展, 机器在软体上已经获得超越人类的信息计算和传输的能力, 然

而在视觉、听觉、触觉, 味觉和嗅觉等硬件方面, 机器感知依然远远的落后于人类自身, 这些都需要通过多种传感器来实现^[1]。人能得到80%信息均来源于视觉, 相较于听觉单一通道感官不同, 人眼作为极其重要的视觉器官, 然而现在最先进的2D数码相机, 在还原真实世界上依旧与人眼有着较大差距, 人眼对三维的感知能力远超现有的3D摄像头^[2]。让机器获得与人类一样的视觉,

即俗称的“电子眼”、机器视觉或者3D立体视觉，目前存在多种技术路线，本文将结合文献与市面公开的资料对多种测距技术进行原理详解和性能比较。

1 近距离：双目/结构光/iToF

1.1 双目视觉

双目视觉是使用双红外传感器，这样能通过不同角度同时获得被测物的数字图像，再通过视差和外参数计算得到特征点的深度。双目估算深度是基于三角测距原理实现的，如下图1所示。Z值即为O点对应的深度，双蓝色的平面即相机拍摄图像。然而在实际中，我们并不知道左右相机中哪两个点对应的是同一点。为了计算视差，必须从左边的图像中找到每个像素，并将其与右边图像中的每个像素进行匹配，而这需要引入对极几何才能将左右图像中的每个像素进行匹配^[3]。此外，要得到真实的三维点云，除了需要知道每个点的距离的视差图，还需要知道真实世界中每个像素点的距离，并计算得到每个点在现实中的坐标。在摄像的过程中，先由世界坐标系下的物体通过刚体转变到摄像机坐标系，再通过透视投影变换到达图像坐标系，这些坐标系之间可通过矩阵变换计算得到（图1）。

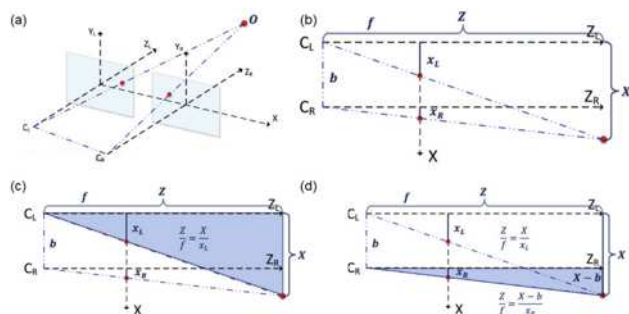


图1 双目视觉深度测量原理

(a) 双目视觉示意图；(b) 鸟瞰俯视图；(c) 左侧摄像头对应的三角形；(d) 右侧摄像头对应的三角形

由于双目立体视觉法会通过视觉特征对图像进行匹配，所以一些缺乏视觉特征的场景（如海洋、颜色墙、雪地等）易出现无法匹配或困难匹配，致使匹配误差大或匹配无法进行。由于双目对环境光的光照也十分敏感，所以拍摄出不同照片对匹配算法也是个很大的挑战，并光线较暗的条件下双目算法无法工作。由于是纯视觉方法，匹配需要使用传统方法，会出现算法中提高错误剔除策略^[4]，导致计算量过大，对硬件有着较高要求。双目视觉与人眼的视觉感受相似，越近的物体测的越准，越远越不准。因此随着距离的提高，测距精度的可靠性变低。除此之外，测距精度和标定也与双目紧密联系，因为双目在车上使用稳定性低，出现需要连续标定。由

于双目摄像头的硬件成本较低的，并且可输出的点云密度较高。目前，双目视觉的厂商主要有索尼，松下，麦格纳，Subaru EyeSight，Foresight，元橡科技，中科慧眼，阜时科技等。

1.2 结构光

结构光技术也是基于三角测距原理实现测距的，其主要是通过投影器投射出一个特定的图案，然后在摄像机拍摄的图像中识别出这个图案，通过计算这个图案投射到物体上的变形程度和位置^[5]，可推断出物体的三维形状，如下图2所示。

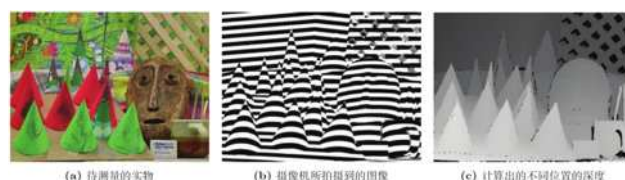


图2 结构光系统成像示例

与被动的双目相比，主动投射的结构光提高了测量的精准度和准确度^[6]。iPhone X的人脸识别就是用的结构光技术（图3）。该模组系统由红外相机和产生特征点图案的点阵投影仪构成。如图3所示，能在点阵投影仪的投影面上找到点亮P0点的投影，同时在相机CCD面上找到被P0点亮的成像点。图中所拍到的五个投影点，由于面形的存在因而改变相邻两点间距离，但不会因为面形的存在导致五个投影点相对顺序发生改变，因此能通过拓扑网格标号来确定相机CCD面上的投影点和投影仪投影面上点的相对对应关系。

结构光传感在近距离场景下的精度较高，但在远距离下，物体上投射图像以及光电的变大，导致精度变差。由于投射的经过编码的图像或散斑光点在室外会被强自然光淹没，所以结构光方案在室外并不好使用。并且，方案也容受光滑平面的反光影响，使得无法接收到信息或存在曝光过度的问题。目前，结构光技术的厂商主要有苹果，Mantis Visio，奥比中光，Opus，安思疆，华捷艾米，阜时科技等。

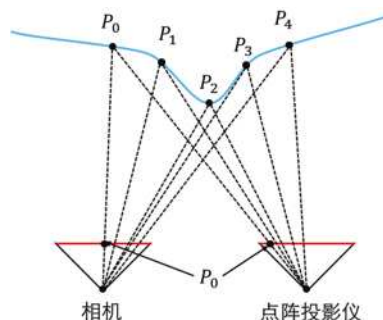


图3 结构光传感示意图

1.3 间接飞行时间测距 iToF

iToF, 是一种间接利用光飞行时间的测距技术^[7], 如4图所示。iToF通过传感器在不同时间窗口收集到能量值的比例关系, 解析出信号相位差 $\Delta\phi$, 从而间接测量发射信号和接收信号的时间差, 进而计算得到深度信息(图4)。

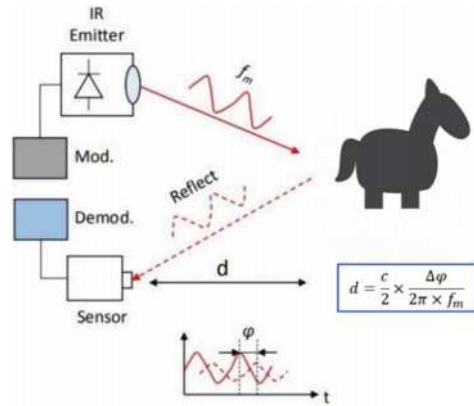


图4 iToF连续波调制测距原理 (f_m 为调制信号频率)

iToF可解析的相位范围为 $[0 \sim 2\pi]$, 对应的最大量程为 $d_{\max} = c/2f_m$ 。频率越高时, 精度越高, 量程越小。当超过量程的深度, 会导致周期性的相位卷绕。例如, 当前目标的距离是0.15m, 那么整个发射和反射光的飞行时间就为1ns。在调制光的周期为10ns时, 1ns的飞行时间差转化为相位差就是36度, 而如果调制光的周期为100ns时, 1ns的飞行时间差转化为相位差就为3.6度。显然, 36度的相位差要更容易检测, 因此iToF的调制光调制频率越高, 则测距精度越高。但调制频率增大限制住最大测距距离。例如, 当调制频率为100MHz时, 那么无论飞行时间长短反应在调制相位差上都是36度, 因此其最大测距距离被调制周期所限制了; 例如, 在10ns调制周期时其最大测距距离是飞行时间等于调制周期时对应的距离(1.5m)。因此, iToF的最大问题就在于最大测距距离和测距精度之间存在矛盾^[8]。此外, iToF需多次采样积分, 测量时间过长, 限制住了相机的帧率, 并且在测量运动物体时会产生运动模糊的风险。iToF技术的厂商主要有英飞凌, Melexis, 飞芯电子, 光微科技, 光鉴科技。

2 远距离: dToF/FMCW/毫米波

2.1 直接飞行时间测距 dToF

dToF, 是一种直接利用光飞行时间来实现测距的技术^[8, 9], 它通过直接测量发射脉冲与反射脉冲的飞行时间来计算出光源与物体之间的距离(如下图3-5所示)。dToF测距原理简单, 深度算法简易, 被认为是目前激光雷达发展的主流技术路线。dToF通过计时飞行时间计算目标距离, 而计时的精度并不会因距离变远而发生变化,

所以理论上在整个量程中, dToF测距的距离分辨率不会存在实质性的变化。dToF测距范围很大, 从数厘米到数百公里; 在200米的测距范围内, 测距精度一般为毫米或厘米级别。

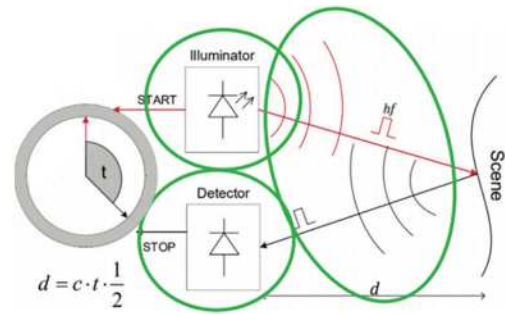


图5 dToF测距原理

由于涉及精确的时间测量, dToF对接发端和计时系统的要求都比较高, 因此在技术实现上存在诸多难点。激光安全的限制及消费类产品的功耗限制, 使得dToF系统发射的脉冲能量十分有限, 因此通常使用纳秒甚至皮秒级的短脉冲激光来提升瞬时发射光功率。发射的激光脉冲在经过物体反射回到接收器时, 能量密度随光传播距离大幅度降低, 因此dToF系统需要灵敏度极高的光电探测器(如APD, SPAD器件等)来检测这些微弱的光信号。而环境光作为噪声, 会干扰高灵敏度接收器对于信号的检测与还原, 并且其他激光雷达发射出的激光信号也会对系统造成干扰, 所以当前dToF需要解决的一大难点便是抗干扰问题。从整体上看, dToF系统测距条件较极端, 脉冲发射峰值巨大、脉冲回波很弱、信号很快, 这就对驱动能力和噪声抑制水平提出了更高的要求。

当前, 基于dToF技术的激光雷达主要使用APD或SPAD作为接收, 这两种接收端所采用的电路架构是不同的。下图6和7分别展示了基于APD和SPAD的测距系统。使用APD作为接收端的dToF激光雷达, 通过跨阻放大器(TIA)将产生电流信号转换为电压信号, 放大器将该输出信号进一步放大, 并通过比较器、时间数字转换器(TDC)等器件来实现飞行的时间的采集(图6)。而使用SPAD作为接收端的dToF激光雷达, 往往采用的是

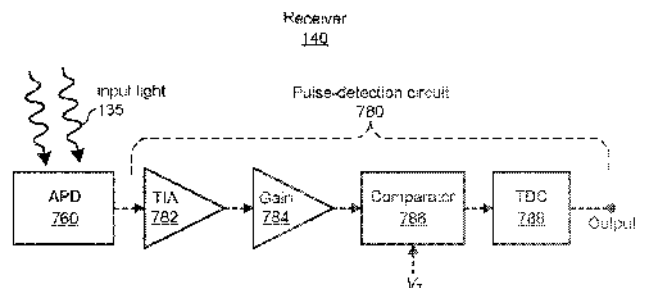


图6 基于APD的dToF测距接收系统

数字电路架构,不需要模数转换,通过TDC采集数字信号并经过时间相关单光子计数法(TCSPC)来确定发射信号到接收信号间的时间(图7)。当前,激光雷达的厂商大多都采用dToF测距技术,如Velodyne, Luminar, Ibeo, Innoviz, Ouster, Leddar Tech, 禾赛科技, 镭神智能, 速腾聚创, 华为, Livox, 一径科技等, 部分厂家专注于FMCW激光雷达的研发。

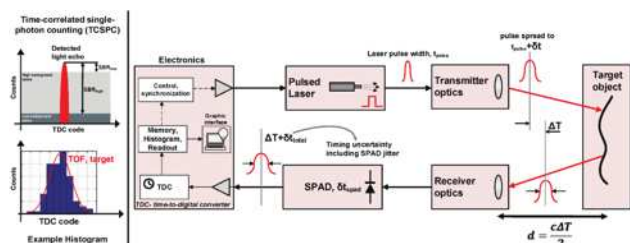


图7 基于SPAD接收端、TCSPC计数的DTOF系统

2.2 调频连续波FMCW

FMCW, 即调频连续波, 是基于发送和接收信号之间的频率差来实现距离测量的一种技术。频率调制就是发射的信号被反复调制和“啁啾”, 信号的频率一次又一次改变, 如图8所示, 该调频的连续波既可以是光波, 也可以是毫米波电磁波, 这里先介绍基于FMCW的激光光波雷达。

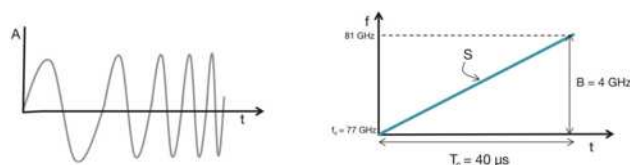


图8 线性调频信号以幅度(左)和频率(右)作为时间的函数

FMCW激光雷达是发射功率恒定但载波频率(或相位)存在持续周期性变化的雷达^[10], 通过相干检测的方式, 测量回波信号与发射信号间由距离延时引入的调制频率差和相对速度引入的多普勒频率差, 从而解调出目标的距离与速度。图9介绍了FMCW的测距原理, 被啁啾调频的光波一部分传输到目标, 另一部分保留在本地作为参照, 把回光和本地光做干涉, 并利用混频探测技术来测量发送和接收的频率差异, 再通过频率差换算出目标物的距离和相对速度。不同于dToF技术对所有进入探测器的光都发生响应, 基于相干原理的FMCW激光雷达只接收自己发出去的频率相近的光, 因此不会受到环境光或其他雷达激光的干扰。另一方面, dToF需要多次测距才能间接估算出障碍物的速度信息, 而FMCW激光雷达在单次测距中便可直接得到物体的速度信息, 降低了后端处理对算力的要求, 节省了反应时间。

尽管FMCW激光雷达不易被干扰并且能直接获取

速度信息, 但目前实现起来还具有一定难度。在收发端, 由于前端和后端的要求, 接收器的整体成本比ToF更高, 而窄线宽高速线性调频激光光源则更是FMCW激光雷达中的一大痛点^[11]。光学元件方面, FMCW系统所有光学器件表面的公差都必须控制在 $\lambda/20$ 以内, 而大部分供应商的技术不能达到如此精度。电子元器件方面, FMCW对ADC转换速率的要求是ToF系统的2~4倍, 对FPGA的要求是能够接收数据并进行超高速FFT转换, 即使使用片上集成的ASIC, FMCW系统所需的处理系统复杂度(和成本)也是ToF系统的几倍。以FMCW为技术路线或将FMCW技术应用到车载激光雷达领域的企业主要有Aeva、Insight、Blackmore(被Aurora收购)、SiLC, 光匀科技, 洛微科技。

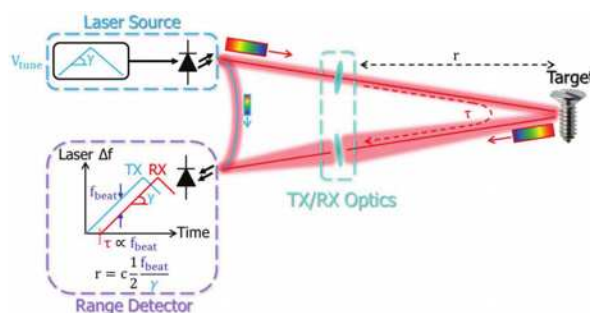


图9 FMCW测距原理

2.3 毫米波

毫米波技术也是基于FMCW原理实现的, 只是发出的调频连续波不是光波而是毫米波。毫米波指波长介于1~10mm(频率30~300GHz)的电磁波^[12]。相较于光波, 毫米波穿透烟、雾、灰尘能力强, 传输距离远, 不易受到较多因素的干扰。虽然毫米波对于外界环境要求较低, 但自身存在着一些致命缺陷。如今汽车搭载的毫米波雷达, 电磁波依旧以一个锥状的波束发射出去, 其信号天线只在二维方向上排布, 因此无法精确感知前方物体的高度信息, 并且毫米波雷达的水平分辨率也较差, 所以无法对障碍物建立精确模型。鉴于以上原因, 毫米波雷达在自动驾驶应用中不得不将这些静止物体反射的信号过滤掉以防止“幽灵刹车”, 而通过物体的速度信息来判断前方是否有障碍物存在。但毫米波雷达只可对前方径向运动的物体(即可以产生多普勒频差的物体)产生感知, 而对横向穿过的目标敏感度低, 此外对于反射波较弱的物体和行人毫米波雷达的识别也较差。

目前, 毫米波雷达频段主要集中在24GHz和77GHz, 汽车领域被广泛使用^[13]。其中, 24GHz频段严格意义来讲其应该被称为厘米波, 用于中短距离。77GHz频段在距离分辨率和精度上都更具优势, 并且体积可以更小, 主要用

于长距离。目前,毫米波技术的厂商主要有德尔福、电装、博世,隼眼科技,德赛西威,智波科技,行易道。此外,还有一些一些厂商,如大陆集团、采埃孚、安波福、傲酷、华为、华域汽车、森思泰克、纵目科技、纳瓦电子等,也在积极开发4D的FMCW毫米波雷达,但4D毫米波雷达与lidar相比在角度分辨率上可能还是会存在差距。

3 结语

上述从测距分类出发分析了不同测距范围的3D立体视觉的技术原理和性能特点^[14]。这些技术各自的优势都比较突出,如双目式低成本,结构光高精度,TOF简单快速,毫米波远距离。但劣势同样明显,如双目式计算量大,结构光抗强光弱,TOF高成本易干扰,毫米波分辨率差等,因此面对简单的应用场景可扬长避短选择合适的传感装置来满足需求。而当面对一些复杂场景的应用时,单一的传感装置可能较难满足需求,目前可将不同的传感器配合使用,发挥各自长处。如当前较热门的无人驾驶,就是纯视觉计算构成,构成是由摄像头主导、再配合毫米波雷达等成本较低的元件组装而成。因此,在复杂或要求较高的场景下,多传感器融合可能是未来发展趋势。

目前,3D立体视觉市场有超过70个玩家,许多厂商均有布局芯片化技术,较知名的厂家有Velodyne, Ouster, Sick, Luminar等。按照代差估计,国外厂商可能比我们先进1-2年。图10列出不同测距原理下的一些激光雷达厂商(有些厂家同时掌握了多种原理测距方案,因此信息统计可能不全面)。

测距方法	厂商
双目	索尼,松下,法雷奥,麦格纳,阜时科技,苏州智华,辉创电子,同致电子
结构光(三角)	苹果,乐动,思岚科技,镭神智能,阜时科技,奥比中光,国科光芯
DTOF	苹果, Velodyne, Luminar, Ibeo, Innoviz, 禾赛科技, 阜时科技, 芯视界、灵明光子、览沃(Livox), 北醒光子, 镭神智能, 芯辉科技、速腾聚创, 乐动, 蓝海光电, 思岚科技, 飞芯电子, 光微科技, 力策科技
ITOF	北醒光子, 镭神智能, 英飞凌, Melexis, 飞芯电子, 奥比中光, 光微科技
FMCW激光雷达	Mobileye, Aeva, Aurora, 光勺科技
FMCW毫米波雷达	德尔福、电装, 博世, 隼眼科技, 德赛西威, 智波科技, 行易道

图10 不同测距原理对应的3D立体视觉厂商

当前,短距内,双目、结构光、iToF、dToF各有所长。市场上的激光雷达大多基于TOF原理,安卓手机厂商多

采用3D iToF技术,3D结构光技术在移动设备端也有较多应用,但苹果已在ipad中率先使用到了dToF技术。长距范围,从技术优缺点出发,dToF和FMCW可能将会是未来自动驾驶激光雷达的主流方案。但是出于成本考虑,当前L2级别的自动驾驶依旧是以摄像头+毫米波雷达为主,辅之以激光雷达。此外,在一些要求较低的场景或者室内短距离场景下,结构光等3D立体视觉技术在扫地机器人或者服务性机器人中都会有用武之地。

参考文献:

- [1]Lu, Yang. "Artificial intelligence: a survey on evolution, models, applications and future trends." Journal of Management Analytics 6.1 (2019): 1-29.
- [2]Gu, Leilei, et al. "A biomimetic eye with a hemispherical perovskite nanowire array retina." Nature 581.7808 (2020): 278-282.
- [3]王荣本, 张明恒, 石德乐. 双目视觉技术在目标测量中的应用[J]. 公路交通科技, 2007, 24 (2): 122-125.
- [4]黄向东, 姬智华. 基于SIFT特征与区域相结合的双目视觉匹配算法[J]. 计算机应用与软件, 2014, 31 (4): 4.
- [5]Rocchini C, Cignoni P, Montani C, et al. A low cost 3D scanner based on structured light[C]//computer graphics forum. Oxford, UK and Boston, USA: Blackwell Publishers Ltd, 2001, 20(3): 299-308.
- [6]赵庆祥, 曾丹, 张之江. 编码结构光技术在人脸建模中的应用[J]. 电子测量技术, 2010 (1): 5.
- [7]Bamji C, Godbaz J, Oh M, et al. A Review of Indirect Time-of-Flight Technologies[J]. IEEE Transactions on Electron Devices, 2022.
- [8]陈千颂, 杨成伟, 潘志文, 等. 激光飞行时间测距关键技术进展[J]. 激光与红外, 2002, 32 (1): 4.
- [9]霍玉晶, 陈千颂, 潘志文. 脉冲激光雷达的时间间隔测量综述[J]. 激光与红外, 2001, 31 (3): 4.
- [10]Martin A, Dodane D, Leviandier L, et al. Photonic integrated circuit-based FMCW coherent LiDAR[J]. Journal of Lightwave Technology, 2018, 36(19): 4640-4645.
- [11]Behroozpour B, Sandborn P A M, Wu M C, et al. Lidar system architectures and circuits[J]. IEEE Communications Magazine, 2017, 55(10): 135-142.
- [12]石星. 毫米波雷达的应用和发展[J]. 电讯技术, 2006.
- [13]向敬成, 张明友. 毫米波雷达及其应用[M]. 国防工业出版社, 2005.
- [14]赵远, 蔡喜平, 陈锺贤, 等. 成像激光雷达技术概述[J]. 激光与红外, 2000, 30 (6): 3.