# 机器人视觉三维成像技术综述

卢荣胜1\*, 史艳琼2\*\*, 胡海兵1

<sup>1</sup>合肥工业大学仪器科学与光电工程学院,安徽 合肥 230009; <sup>2</sup>安徽建筑大学机械与电气工程学院,安徽 合肥 230601

**摘要** 本文针对智能制造领域机器人视觉感知中的三维视觉成像技术进行综述,系统地总结了一些有代表性的机器人视觉成像方法的特点和实际应用中的局限性,内容涉及飞行时间三维成像、点线扫描三维成像、色散共焦成像、结构光投影三维成像、光学偏折成像、单目与多目立体视觉三维成像和光场成像等。绘制了各种视觉成像的图谱,并探讨了机器人手眼系统最佳三维成像方法。

关键词 机器视觉; 三维成像; 手眼系统; 结构光投影; 立体视觉

**中图分类号** TP74 文献标志码 A

doi: 10.3788/LOP57.040001

# **Review of Three-Dimensional Imaging Techniques for Robotic Vision**

Lu Rongsheng<sup>1</sup>\*, Shi Yanqiong<sup>2</sup>\*\*, Hu Haibing<sup>1</sup>

<sup>1</sup>School of Instrument Science and Opto-Electronics Engineering, Hefei University of Technology,

Hefei, Anhui 230009, China;

<sup>2</sup> School of Mechanical and Electrical Engineering, Anhui Jianzhu University, Hefei, Anhui 230601, China

**Abstract** The three-dimensional (3D) imaging techniques of robotic vision in the field of intelligent manufacturing robot vision perception are reviewed. The characteristics and limitation in practical applications of some typical robot vision imaging methods are systematically summarized. The content involves time-of-flight imaging, point and line scanning imaging, chromatic confocal imaging, structured light projection imaging, deflectometric imaging, monocular and multi-view stereo imaging, and light field imaging. The tree map of various robotic vision imaging methods are drawn. The best 3D imaging methods of eye-in-hand robotic system are discussed.

Key words machine vision; three-dimensional imaging; eye-in-hand system; structured light projection; stereo vision

OCIS codes 150.0155; 100.3010

# 1引言

在智能制造过程中,通过传统的编程来执行某 一特定动作的机器人(机械手、机械手臂、机械臂等, 未作特殊说明时,不作严格区分,统一称为机器人), 将难以满足制造业向前发展的需求。很多应用场合 下,需要为工业机器人安装一双眼睛,即机器视觉成 像感知系统,使机器人具备识别、分析、处理等更高 级的功能。这在高度自动化的大规模生产中非常重 要,只有当工业机器人具有视觉成像感知系统,具备 观察目标场景的能力时,才能正确地对目标场景的 状态进行判断与分析,做到智能化灵活地自行解决 发生的问题。

在智能制造领域,机器人视觉成像系统主要用 于测量、检测、识别与引导四个方面<sup>[1]</sup>。例如,在规 模化的生产过程中,视觉检测具有非接触、高效、低 成本、自动化程度高等优点,免除人工抽检带来的低 效、误差与漏检,正逐渐取代传统的检测方法。在精

收稿日期: 2020-01-06; 修回日期: 2020-02-08; 录用日期: 2020-02-11

基金项目:国家自然科学基金(51875164)、国家重点研发计划(2018YFB2003801)、国家重大科学仪器开发与应用专项 (2013YQ220749)

<sup>\*</sup> E-mail: rslu@hfut.edu.cn; \*\* E-mail: yqshi@ahjzu.edu.cn

密制造与组装生产线上,机器视觉成像系统通过精 密测量,可准确确定零部件的方向与位置,并传送给 机器人手臂(机械手臂),引导手臂末端正确地操控 零部件。

视觉成像的一个重要特点是从图像中获取目标 的信息。视觉成像最初从二维(2D)图像处理与理 解,即 2D 视觉成像发展起来的。2D 视觉技术主要 根据灰度或彩色图像中的像素灰度特征获取目标中 的有用信息,以及基于轮廓的图案匹配驱动,识别物 体的纹理、形状、位置、尺寸和方向等。2D视觉技术 距今已发展了 30 余年,在自动化和产品质量控制过 程中得到广泛应用,目前技术较为成熟,主要用于字 符与条码识读、标签验证、形状与位置测量、表面特 征检测等。但 2D 视觉技术难以实现三维高精度测 量与定位,二维形状测量的一致性和稳定性也较差, 易受照明条件等影响。尤其当前智能制造技术对机 器人视觉性能的要求越来越高,2D机器视觉技术的 局限性已经显现,机器人视觉系统集成商已经发现 越来越难以通过 2D 机器视觉系统来增值,迫切需 要发展三维(3D)视觉技术,因为 3D 视觉技术能够 产生 2D 视觉无法产生的形状或深度信息,因此使 用范围更广。

近 30 年来,数字图像处理技术已经从简单的二 值图像处理发展到高分辨率多灰度(彩色、多光谱) 图像处理,在理论算法上取得了巨大进步;计算机技 术、并行处理、人工智能、神经元网络和深度学习等 学科的快速发展,也促进了对机器人视觉成像系统 与复杂视觉过程的深入研究,为 3D 视觉成像技术 研究及应用奠定了良好的基础。

当前,机器人视觉成像技术及系统正越来越广 泛地应用于视觉测量、检测、识别、引导和自动化装 配领域中。虽然很多机器人具备一定程度的智能 化,但还远未达到人类所需的智能化程度,一个重要 原因是机器人视觉感知系统中还有许多科学问题、 关键应用技术问题等,仍亟待解决。如:1)如何使机 器人像人那样,对客观世界的三维场景进行感知、识 别和理解;2)哪些三维视觉感知原理可以对场景目 标进行快速和高精度的三维测量,并且基于该原理 的三维视觉传感器具有小体积、低成本,方便嵌入到 机器人系统中;3)基于三维视觉系统获得的三维场 景目标信息,如何有效地自组织自身的识别算法,准 确、实时地识别出目标;4)如何通过视觉感知和自学 习算法,使机器人像人那样具有自主适应环境的能 力,自动地完成人类赋予的任务等。 在智能制造领域,机器人视觉主要被用来代替 人工视觉和拓展人工视觉的功能,达到高效、快速、 准确地完成测量与检测、识别与分类、引导与定位等 任务,或者对机器人本身进行基于视觉的伺服控制。 所有这些任务的实现都是基于 2D 或 3D 视觉成像 完成的。因此,本文重点针对智能制造领域机器人 视觉感知中的三维视觉成像技术原理、应用中的局 限性等问题进行综述,分析各种三维成像方法的特 点,探究机器人三维视觉成像技术的发展趋势。

# 2 机器人视觉成像系统的结构形式

机器人视觉系统的主要功能是模拟人眼视觉成 像与人脑智能判断和决策功能,采用图像传感技术 获取目标对象的信息,然后对图像信息提取、处理并 理解,最终用于机器人系统对目标实施测量、检测、 识别与定位等任务,或用于机器人自身的伺服控制。 视觉测量(如外形、尺寸、坐标、方位等的测量)更关 注视觉成像系统的精密度(简称精度),检测(如缺陷 检测)强调特征成像的分辨能力和灵敏度,而识别 (如目标上的二维码、文字、色彩等)侧重于特征辨识 的准确性,定位(目标的绝对位置,或机械手末端与 目标之间的相对位置)注重成像系统测量的准确性 (或精确性)。所谓机器人视觉伺服控制,就是采用 视觉感知来间接准确地获取机器人相对于基坐标系 的位姿或者相对于场景目标的位姿,在此基础上实 现机器人的定位控制或者轨迹跟踪。机器人视觉伺 服控制可分为基于图像(2D)、位置(3D)、位置和图 像(2.5D)三种策略<sup>[2-4]</sup>。但不论何种控制策略,都建 立在机器人视觉系统的测量、检测与识别的基础上, 视觉定位也是如此。

在工业应用领域,最具有代表性的机器人视觉 系统就是机器人手眼系统。根据成像单元安装方式 不同,机器人手眼系统分为两类:固定成像单元眼看 手系统(Eye-to-Hand)和随动成像单元眼在手系统 (Eye-in-Hand, or Hand-eye)<sup>[5]</sup>。如图1所示。

### 2.1 眼看手视觉成像系统

在 Eye-to-Hand 系统中,视觉成像单元安装在 机器人本体外的固定位置,在机器人工作过程中不 随机器人一起运动,当机器人或目标运动到机械臂 可操作的范围时,机械臂在视觉感知信息的反馈控 制下,向目标移动,对目标进行精准操控。Eye-tohand 系统的优点是具有全局视场,标定与控制简 单、抗震性能好、姿态估计稳定等。但存在下列 不足<sup>[6]</sup>。



图 1 两种机器人手眼系统的结构形式。(a)眼在手机器人系统;(b)眼看手机器人系统

Fig. 1 Configurations of two hand eye systems. (a) Eye-in-hand robotic system; (b) eye-to-hand robotic system

 容易产生遮挡问题。当机械臂向目标移动 及操作时,在一定的空间位置会对目标图像造成遮 挡,不适合机器人实现基于图像的视觉控制和混合 视觉控制。

2)分辨率低,容易产生较大的定位误差。这种 系统中的相机通常与目标的距离较远,目标分辨率 较低,相机对目标的测量结果不会随机械臂向目标 移动发生变化。当相机标定精度不高时,会产生比 较大的绝对定位误差,甚至会导致机械臂不能准确 接近目标。

#### 2.2 眼在手视觉成像系统

在 Eye-in-Hand 系统中,成像单元安装在机器 人手臂末端,随机器人一起运动<sup>[7-12]</sup>。Eye-in-Hand 系统常用在有限视场内操控目标,但具有下列优点。

1)不会像 Eye-to-Hand 系统那样产生机械臂遮 挡成像视场问题。适合采用基于图像的视觉控制、 基于位置的视觉控制以及结合两者的混合视觉 控制。

2) 成像单元单次视场有限,但空间分辨率高。 对于基于图像的视觉控制,因在图像空间形成闭环, 成像单元模型参数的标定误差可以被有效地克服, 因而对标定的精度要求不高。对于基于位置的视觉 控制,虽然成像单元的标定误差不能在控制系统中 被有效地克服,但随着目标的接近,测量出的目标位 置的绝对误差降低,即使标定存在一定误差,一般也 能够满足应用要求。

#### 2.3 混合协同视觉系统

有些应用场合,为了更好地发挥机器人手眼系统的性能,充分利用 Eye-to-Hand 系统全局视场和 Eye-in-Hand 局部视场高分辨率和高精度的性能,可采用两者混合协同模式<sup>[13-18]</sup>,如利用 Eye-to-Hand 系统负责机器人的定位,利用 Eye-in-Hand 系统负责机器人的定向;或者利用 Eye-to-Hand 估

计机器人相对目标的方位,利用 Eye-in-Hand 负责 目标姿态的高精度估计等。

图 2 所示为一个协同合作的视觉成像系统,它 由一个固定眼看手成像单元和一个随动眼在手成像 单元组成。固定成像眼看手系统提供了一个目标场 景的全局视图,而眼在手随动成像单元随机器人一 起运动,提供了高分辨率局部图像。与开环视觉控 制模型相比,这种视觉伺服控制策略,可使用成像单 元对目标几何参数(外形、距离、坐标、方向等)进行 实时估计,精准操控目标物体,如焊接、抓取、装配、 码垛、拆垛等。

## 3 机器人视觉 3D 成像方法与特点

3D视觉成像是工业机器人信息感知的一种最 重要的方法,可分为光学和非光学成像方法。目前 应用最多的还是光学方法,包括:飞行时间法、结构 光法、激光扫描法、莫尔条纹法、激光散斑法、干涉 法、照相测量法、激光跟踪法、从运动获得形状 (Shape from motion)、从阴影获得形状(Shape from shading),以及其他的 Shape from X 等<sup>[19-23]</sup>。 下面对几种典型方法的特点和性能进行概述。

#### 3.1 飞行时间 3D 成像

飞行时间(TOF)相机每个像素利用光飞行的 时间差来获取物体的深度。如图 3 所示<sup>[24]</sup>,在经典 的 TOF 测量方法中,探测器系统在发射光脉冲的 同时启动探测接收单元进行计时,当探测器接收到 目标发出的光回波时,探测器直接存储往返时间  $\tau_{TOF}$ 。目标距离 z 可通过以下简单方程估算,

$$z = \frac{c}{2} \cdot \tau_{\text{TOF}}, \qquad (1)$$

式中:c=2.9979×10<sup>8</sup> m/s为光速,这种测距方式也称为直接 TOF(D-TOF)。D-TOF 通常用于单点测距系统,为了实现面积范围3D成像,通常需要采用



图 2 机器人协同视觉系统原理图

Fig. 2 Schematic of a cooperative vision system of robots



图 3 TOF 成像原理<sup>[24]</sup> Fig. 3 TOF imaging principle<sup>[24]</sup>

扫描技术<sup>[25]</sup>。无扫描 TOF 三维成像技术直到近几 年才实现<sup>[26]</sup>,因为在像素级实现亚纳秒电子计时是 非常困难的。与直接计时的 D-TOF 不同的方案是 间接 TOF(I-TOF),时间往返行程是从光强度的时 间选通测量中间接外推获得。I-TOF 不需要精确的 计时,而是采用时间选通光子计数器或电荷积分器, 它们可以在像素级实现。I-TOF 是目前基于 TOF 相机的电子和光混合器的商用化解决方案<sup>[27-28]</sup>。目 前已经有飞行时间 TOF 面阵相机商业化产品,如 Mesa Imaging AG SR-4000、PMD Technologies CamCube 3.0、微软 Kinect V2。

TOF 成像可用于大视野、远距离、低精度、低成本的 3D 图像采集。其特点是:检测速度快、视野范

围较大、工作距离远、价格便宜,但精度低,易受环境 光的干扰<sup>[29-30]</sup>。例如 Camcueb3.0 具有可靠的深度 精度(<3 mm@4 m),每个像素对应一个 3D 数据; Kinect V2,深度精度大约为 3~4 cm<sup>[31-32]</sup>。

#### 3.2 扫描 3D 成像

扫描 3D 成像方法可分为扫描测距、主动三角 法、色散共焦法等。其实,色散共焦法是扫描测距法 的一种,考虑到目前在手机、平板显示等制造行业应 用比较广泛,故在下文中用单独一节描述。 3.2.1 扫描测距

扫描测距是利用一条准直光束通过一维测距扫 描整个目标表面实现 3D 测量。典型扫描测距方法 有<sup>[25,33]</sup>:1)单点飞行时间法,如连续波频率调制 (FM-CW)测距、脉冲测距(激光雷达)等;2)激光散 射干涉法,如基于多波长干涉、全息干涉、白光干涉、 散斑干涉等原理的干涉仪;3)共焦法,如色散共焦、 自聚焦等。

单点测距扫描 3D方法中,单点飞行时间法适合远距离扫描,测量精度较低,一般在毫米量级。其他几种单点扫描方法有:单点激光干涉法、共焦法和单点激光 主动三角法,测量精度较高,但前者对环境要求高;线 扫描精度适中,效率高。比较适合于机械手臂末端执 行 3D测量的应是主动激光三角法和色散共焦法。 3.2.2 主动三角法

主动三角法是基于三角测量原理,利用准直光 束、一条或多条平面光束扫描目标表面完成 3D 测 量的。光束常采用以下方式获得:激光准直、圆柱或 二次曲面柱形棱角扩束,非相干光(如白光、LED光 源)通过小孔、狭缝(光栅)投影或相干光衍射等。主 动三角法可分为三种类型<sup>[34]</sup>:单点扫描、单线扫描 和多线扫描,如图4所示。



图 4 主动三角法扫描成像<sup>[34]</sup>。(a)准直光束一维测距;(b)光条投影二维轮廓测量;(c)多光条投影三维面形测量 Fig. 4 Active triangulation scanning imaging methods<sup>[34]</sup>. (a) Distance measurement using a collimating beam light; (b) 2D measurement with a sheet of light projection; (c) 3D measurement with a multi-sheet light projection

在未扫描的情况下,图 4(a)所示的主动三角 法沿着激光线的方向可实现一维测量,图 4(b)在 光平面上可实现二维测量,图 4(c)可实现三维测 量。通常单点扫描精度高于线扫描,但线扫描的 效率远大于前者。目前商业化的用于机械手臂末 端的产品大多数是单点和单线扫描仪<sup>[35]</sup>。在多线 扫描方法中,条纹极数可靠识别是难点。为了准 确识别条纹编号,通常采用两组垂直光平面高速 交替成像,这样还可以实现"Flying Triangulation" 扫描<sup>[36-37]</sup>,其扫描与三维重构过程如图 5 所示。 多线条投影一次频闪成像产生一幅稀疏 3D 视图, 通过纵横向条纹投影扫描生成若干幅 3D 视图序 列,再通过三维图像准配生成高分辨率的完整致 密的三维曲面模型<sup>[38]</sup>。



图 5 飞行三角三维测量原理[36]

Fig. 5 3D measurement principle of flying triangulation<sup>[36]</sup>

3.2.3 色散共焦法

色散共焦似乎可以扫描测量粗糙和光滑的不透明和透明物体,如反射镜面、透明玻璃面等,目前在 手机盖板三维检测等领域广受欢迎。

色散共焦扫描有三种类型:单点一维绝对测距 扫描<sup>[39]</sup>、多点阵列扫描<sup>[40-41]</sup>和连续线扫描<sup>[42-46]</sup>,图 6 和图 7 分别列出了绝对测距和连续线扫描两类示 例,其中连续线扫描也是一种阵列扫描,只是阵列的 点阵更多、更密集。在商业产品上,目前较为知名的 扫描光谱共焦传感器是法国的 STIL MPLS180,采 用 180 个阵列点形成一条线,最大线长 4.039 mm (测量点 11.5 μm,点与点间距为 22.5 μm),另一款 产品是芬兰的 FOCALSPEC UULA,采用的是色散 共焦三角法技术。



图 6 两种色散共焦单点测距方法。(a)基于小孔和分光镜的结构;(b)基于 Y 型光纤分光的结构

Fig. 6 Two types of chromatic confocal measurement technology. (a) Structure based on pinhole and beam splitter; (b) structure based on Y-type optical fiber





Fig. 7 Two type schematics of chromatic confocal line scan imaging techniques. (a) Chromatic confocal microscope<sup>[43]</sup>; (b) chromatic confocal triangulation<sup>[44]</sup>

#### 3.2.4 小结

扫描 3D 成像的最大优点是测量精度高,其中 色散共焦法还有其他方法难以比拟的优点,即非常 适合测量透明物体、高反与光滑表面的物体。但缺 点是速度慢、效率低;当用于机械手臂末端时,可实 现高精度 3D 测量,但不适合机械手臂实时 3D 引导 与定位,因此应用场合有限;另外主动三角扫描在测 量复杂结构形貌时容易产生遮挡,需要通过合理规 划末端路径与姿态来解决。

#### 3.3 结构光投影 3D 成像

结构光投影三维成像目前是机器人 3D 视觉感知的主要方式,结构光成像系统是由若干个投影仪和 相机组成,常用的结构形式有:单投影仪-单相机、单 投影仪-双相机<sup>[47]</sup>、单投影仪-多相机、单相机-双投影 仪<sup>[48-49]</sup>和单相机-多投影仪<sup>[50]</sup>等典型结构形式。

结构光投影三维成像的基本工作原理是:投影

仪向目标物体投射特定的结构光照明图案,由相机 摄取被目标调制后的图像,再通过图像处理和视觉 模型求出目标物体的三维信息。常用的投影仪主要 有下列几种类型:液晶投影(LCD)、数字光调制投影 (DLP)[如数字微镜器件(DMD)]<sup>[51-52]</sup>、激光 LED 图案直接投影<sup>[53]</sup>。根据结构光投影次数划分,结构 光投影三维成像可以分成单次投影 3D 和多次投影 3D 方法<sup>[54-55]</sup>。

#### 3.3.1 单次投影成像

单次投影结构光主要采用空间复用编码和频率 复用编码形式实现<sup>[56-57]</sup>,常用的编码形式有:彩色编 码<sup>[58-60]</sup>、灰度索引、几何形状编码<sup>[61]</sup>和随机斑 点<sup>[62]</sup>。目前在机器人手眼系统应用中,对于三维测 量精度要求不高的场合,如码垛、拆垛、三维抓取等, 比较受欢迎的是投射伪随机斑点获得目标三维信 息,其 3D 成像原理如图 8 所示。



图 8 伪随机散斑投影 3D 成像原理与算法过程

Fig. 8 Principle and algorithm process of pseudo-speckle projection 3D imaging

该方法的难点是如何准确和快速地计算两幅图 像中的匹配点,即开发合适的立体匹配算法,求出稀 疏或稠密深度图。其中,稀疏匹配只选择关键特征 点,如工业产品零部件的几何角点、边缘点等,计算 立体视差值。稀疏算法计算匹配点少、速度快,但后 续三维重构是需要通过插值计算缺失的像素点视 差,因此应用场合受限。稠密匹配是对几乎所有像 素生成确定视差值,计算量相对比较大,但获得的匹 配点多。稠密匹配算法很多,比较经典的算法有:快 速零均值归一化自相关局部算法[63]、灰度绝对值差 和算法(SAD)、BT (Birchfield and Tomasi)算 法<sup>[64]</sup>、AD-Census ( absolute difference-Census Transform)<sup>[65]</sup>、半全局匹配算法(SGM)<sup>[66]</sup>、半全局 算法(SGBM)<sup>[67]</sup>、图割全局算法(GC)<sup>[68]</sup>等,其算法 的准确性和速度等性能比较可参看 vision. middlebury.edu 网页上的 Middleburry 测试平台对

比评价结果[69]。

3.3.2 多次投影成像

多次投影 3D 方法主要采用时间复用编码方式 实现<sup>[70]</sup>,常用的图案编码形式有:二进制编码<sup>[71-72]</sup>、 多频相移编码<sup>[73-75]</sup>和混合编码法(如格雷码+相移 条纹)<sup>[76]</sup>等。

条纹投影 3D 成像基本原理如图 9 所示,利用 计算机生成结构光图案或用特殊的光学装置产生结 构光,经过光学投影系统投射至被测物体表面,然后 采用图像获取设备(如 CCD 或 CMOS 相机)采集被 物体表面调制后发生变形的结构光图像,利用图像 处理算法计算图像中每个像素点与物体轮廓上点的 一一对应关系;最后通过系统结构模型及其标定技 术,计算得到被测物体的三维轮廓信息。在实际应 用中,常采用格雷码投影、正弦相移条纹投影或格雷 码+正弦相移混合投影 3D 技术。



图 9 多次投影 3D 成像。(a)多次投影 3D 系统机构示意图;(b)二进制格雷码投影 3D 基本原理; (c)二进制格雷码+正弦相移混合编码投影 3D

Fig. 9 Multi-shot 3D imaging. (a) Configuration schematics of a multi-shot 3D imaging; (b) principle of binary Gray code 3D imaging; (c) mixed encoding with binary Gray code and phase shifting for 3D imaging

格雷码是贝尔实验室的 Frank Gray 在 1940 年 基于 Posdamer 等<sup>[77]</sup>提出的普通二进制编码改进得 到的。图 9(b)给出了格雷码编码投影 3D 的原理示 意,对于任意两个相邻码值,有且只有一位二进制数 不同,无权重高低,其解码误差较小,可以得到绝对 编码值,将这种编码方式应用于三维形貌测量,可以 准确地获取被测物体轮廓上各采集点相对于传感器 坐标系下的绝对位置<sup>[78]</sup>。但是格雷码方法仅能在 投射空间内进行离散的划分,空间分辨率受到成像 器件的限制。为了提高空间分辨率,需要增加投影 条纹幅数,投射条纹宽度更小的格雷码条纹图,但条 纹宽度过小会导致格雷码条纹的边缘效应,从而引 起解码误差。

正弦光栅条纹投影克服了格雷码空间离散划分的缺点,成为使用率最高的结构光类型之一。众所周知,对于复杂外形,如有空洞、阶梯、遮挡等,采用 正弦单频相移法条纹投影时,存在相位解包裹难题。 另外为了能够从系列条纹图中求出相位绝对值,需 要在条纹中插入特征点,比如一个点、一条线作为参 考相位点,但是这些点或线特征标志有可能投影在 物体的遮挡或阴影区域,或受到环境光等干扰等,发 生丢失,影响测量结果的准确性。因此,对于复杂轮 廓的物体,常采用多频相移技术。

由于硬件光栅制作的成本高、灵活性低,随着数 字投影仪的迅速发展,利用数字投影仪在计算机控 制下投射条纹,已成为最流行的投影方式。该方法 是利用计算机产生精确的数字光栅相移条纹图像, 经过数字投影设备投影到被测物体表面。但数字投 影仪的非线性响应不可避免地导致相位测量误差, 这种非线性响应误差是众多相位误差源中最主要的 一种。另外由 CCD 或 CMOS 相机、镜头组成的光 学成像系统,本身也是一种非线性滤波器,用它采集 被物体表面调制后发生变形的结构光图像,再从图 像中解调出条纹相位,也不可避免地导致相位测量 的非线性误差。第三种非线性相位误差来源于测量 环境与被测表面的多样性,如环境光照明的变化、被 测物体表面色彩与纹理复杂多样、表面超光滑性以 及反射率动态范围变化大等,归根结底这种非线性 误差是由反射率变化造成的。常规的数字相移条纹 投影三维测量方法针对漫反射表面具有较好的测量 结果,对表面反射率变化较大的大动态范围被测对

象,采集图像的灰度值局部容易达到饱和,造成相位 测量误差。然而相位的准确测量是精确三维形貌恢 复的重要保证,任何相位误差都将直接影响最终三 维测量结果,如产生三维轮廓拉丝现象等。因此,对 于高精度三维成像测量,减小正弦条纹非线性误差 带来的 3D 重建模型误差,可采用相位非线性误差 修正技术<sup>[79]</sup>。

对于多频相移投影,例如采取3频3步相移,再加上非线性误差校正技术,一共至少需要投射12幅条纹图。但目前基于数字微镜器件(DMD)开发的 微型投影仪,如 TI 公司的 DLP VisionFly4500,使用 DLPC350 控制器,内部 RAM 只能烧48个1 bit 图,如果是8位灰度图,只能烧6幅图。因此,实现3频3步相移和非线性误差校正,只能利用外存储器或者由上位机输送,这样难以实现高速多频相移3D 成像。为了解决此问题,可以采取格雷码+相移的条纹投影相结合的方式<sup>[72,80-82]</sup>。投影 N 幅正弦条纹图,其条纹周期与格雷码最小黑白条纹周期相同,可以得到连续单值的绝对相位值,从而大大提高了测量准确度、精度,同时兼顾了测量速度,如图9(c)所示。

3.3.3 偏折法成像

对于粗糙表面,结构光可以直接投射到物体表 面进行视觉成像测量;但对于大反射率光滑表面和 镜面物体 3D 测量,结构光投影不能直接投射到被 测表面,3D 测量还需要借助镜面偏折技术<sup>[83-86]</sup>,如 图 10 所示。在这种方案中,条纹不是直接投影到被 测轮廓上,而是投射到一个散射屏上,或用液晶显示 屏代替散射屏把条纹直接显示出来。相机通过光亮 表面折返光路,获取被光亮表面曲率变化调制的条 纹信息,然后解算出三维轮廓形貌。如图11所示,



图 10 光亮表面偏折测量方法示意图。(a)偏折法机器人手眼系统示例<sup>[83]</sup>;(b)偏折法测量过程示例<sup>[89]</sup> Fig. 10 Schematics of deflectometry for bright surface. (a) Demonstration of an eye-in-hand robot for deflectometry<sup>[83]</sup>; (b) explanation of measurement process of deflectometry<sup>[89]</sup>



图 11 偏折法测量原理 Fig. 11 Principle of deflectometry

当被测面为平面时,参考屏幕上一点 Q<sub>i</sub>,在镜面 S<sub>i</sub>处经反射后,对应于相机图像平面上的点 P<sub>i</sub>, 入射光线为 w<sub>i</sub>,反射光线为 v<sub>i</sub>。当被测面在该点 处存在倾角θ时,对于同一像点,入射光线将偏折 角度 2θ,并与参考屏幕交于点 Q'<sub>i</sub>。由此可得待测 面的表面倾角会使光线传播方向发生偏折,导致 相机图像平面上的点 P<sub>i</sub> 对应于参考屏幕上的光 源点位置发生偏移。该偏移量与被测曲面的梯度 (或法向量)相关,对表面梯度进行积分处理可实 现待测面的面型测量,或通过双目视觉建模进行 求取<sup>[87-92]</sup>。

3.3.4 小结

由于单次投影曝光和测量时间短,抗振动性能 好,适合运动物体的 3D 测量,如机器人实时运动引 导,手眼机器人对生产线上连续运动产品进行抓取 等操作。但深度垂直方向上的空间分辨率受到目标 视场、镜头倍率和相机像素等因素的影响,大视场情 况下不容易提升。

多次投影方法(如多频条纹方法)具有较高空间 分辨率,能有效地解决表面斜率阶跃变化和空洞等 难题。不足之处在于:①对于连续相移投影方法, 3D重构的精度容易受到投影仪、相机的非线性和环 境变化的影响;②抗振动性能差,不合适测量连续运 动的物体;③在 Eye-in-Hand 视觉导引系统中,机械 臂不易在连续运动时进行 3D 成像和引导;④实时 性差,不过随着投影仪投射频率和 CCD/CMOS 图 像传感器采集速度的提高,多次投影方法实时 3D 成像的性能也在逐步改进<sup>[54,93-94]</sup>。

偏折法对于复杂面型的测量,通常需要借助多次投影方法,因此具有多次投影方法相同的缺点。 另外偏折法对曲率变化大的表面测量有一定的难度,因为条纹偏折后的反射角的变化率是被测表面 曲率变化率的2倍,因此对被测物体表面的曲率变 化比较敏感,很容易产生遮挡难题。

#### 3.4 立体视觉 3D 成像

立体视觉字面意思是用一只眼睛或两只眼睛感 知三维结构,一般情况下是指从不同的视点获取两 幅或多幅图像重构目标物体 3D 结构或深度信 息<sup>[95-96]</sup>。深度感知视觉线索可分为 Monocular cues 和 Binocular cues(双目视差)。目前立体视觉 3D 可 以通过单目视觉、双目视觉、多(目)视觉、光场 3D 成像(电子复眼或阵列相机)实现。

3.4.1 单目视觉成像

单目视觉深度感知线索通常有:透视<sup>[97]</sup>、焦距 差异<sup>[98-100]</sup>、多视觉成像<sup>[101-102]</sup>、覆盖、阴影<sup>[103-104]</sup>、运 动视差<sup>[105-107]</sup>等。在机器人视觉里还可以用镜 像<sup>[108]</sup>,以及其他 shape from X<sup>[109]</sup>等方法实现。 3.4.2 双目视觉成像

双目视觉深度感知视觉线索有:眼睛的收敛位 置和双目视差<sup>[95,111]</sup>。在机器视觉里利用两个相机 从两个视点对同一个目标场景获取两个视点图像, 再计算两个视点图像中同名点的视差获得目标场景 的 3D 深度信息。典型的双目立体视觉计算过程包 含下面四个步骤:图像畸变矫正、立体图像对校正、 图像配准和三角法重投影视差图计算<sup>[112]</sup>,如图 12 所示。



图 12 双目立体视觉系统与计算过程示意图[113]

Fig. 12 Schematic diagram of binocular stereo vision system and calculation process<sup>[113]</sup>

3.4.3 多(目)视觉成像

多(目)视觉成像,也称多视点立体成像,用单个 或多个相机从多个视点获取同一个目标场景的多幅 图像,重构目标场景的三维信息。其基本原理如图 13 所示<sup>[114]</sup>,给定空间 n 个固定的 3D 点和 m 个视 点图像,根据下式从  $m \times n$  个匹配点  $p_{kj}$ 估计 m 个 投影矩阵  $H_k$  和 n 个 3D 点  $P_j$  坐标。

 $\boldsymbol{p}_{kj} = \boldsymbol{H}_k \boldsymbol{P}_j, k = 1, \cdots, m, j = 1, \cdots, n.$ (2) 多视点立体成像主要用于下列几种场景:

1) 使用多个相机从不同视点,获取同一个目标



图 13 多视点成像基本原理[114]

Fig. 13 Principle of multi-view imaging<sup>[114]</sup>

场景多幅图像,然后基于特征的立体重构等算法求 取场景深度和空间结构信息<sup>[114-116]</sup>。

2)从运动恢复形状(SfM)的技术[117]。使用同



一相机在其内参数不变的条件下,从不同视点获取 多幅图像,重构目标场景的三维信息。该技术常用 于跟踪目标场景中大量的控制点,连续恢复场景的 3D结构信息、相机的姿态和位置。

3.4.4 光场成像

光场 3D 成像的原理与传统 CCD 和 CMOS 相机 成像原理在结构原理上有所差异,传统相机成像是光 线穿过镜头在后续的成像平面上直接成像,一般是 2D 图像,如图 14(a)所示;光场相机成像是在传感器 平面前增加了一个微透镜阵列,将经过主镜头入射的 光线再次穿过每个微透镜,由感光阵列接收,从而获 得光线的方向与位置信息,使成像结果可在后期处 理,达到先拍照,后聚焦的效果,如图 14(b)所示。



图 14 光场相机成像与传统相机成像对比。(a)传统相机成像原理;(b)光场相机结构与成像原理

Fig. 14 Imaging comparison between light field camera and traditional camera. (a) Imaging principle of traditional camera; (b) structure and imaging principle of light field camera

目前已有商业化的光场相机,如 Raytrix 光场 3D 相机<sup>[118]</sup>。如图 15 所示<sup>[119]</sup>,从外观来看,与常规的单 个工业相机无异,但内部结构不同,从图 14(b)可以看 出,它是由微透镜阵列成像实现的,相当于阵列复眼 相机成像<sup>[120-121]</sup>。光场相机的发展历程如图 16 所示。

光场相机的优点是:单个相机可以进行 3D 成 像,横向和深度方向的空间分辨率可以达到 20 μm 到 mm 量级,景深比普通相机大好几倍,比较适合 Eye-in-Hand 系统 3D 测量与引导,但目前精度适中 的商业化光场相机价格昂贵。

3.4.5 小结

立体视觉可分为被动和主动两种形式。被动视 觉成像只依赖相机接收到的由目标场景产生的光辐 射信息,该辐射信息通过 2D 图像像素灰度值进行 度量。被动视觉常用于特定条件下的 3D 测量场 合,如室内、目标场景光辐射动态范围不大和无遮 挡;场景表面非光滑,且纹理清晰,容易通过立体匹 配寻找匹配点;或者像大多数工业零部件,几何规则 明显,控制点比较容易确定等。

主动立体视觉是利用光调制(如编码结构光、激 光调制等)照射目标场景,对目标场景表面的点进行 编码标记,然后对获取的场景图像进行解码,以便可 靠地求得图像之间的匹配点,再通过三角法求解场



图 15 Lytro 光场相机内部结构<sup>[119]</sup>

Fig. 15 Inside structure of the Lytro light field camera<sup>[119]</sup> 景的 3D 结构。前面一节"结构光投影 3D"都可以 认为是主动立体视觉方法。主动立体视觉的优点是 抗干扰性能强、对环境要求不高(如通过带通滤波消 除环境光干扰),3D 测量精度、重复性和可靠性高; 缺点是对于结构复杂的场景容易产生遮挡问题。

立体视觉难点在于视觉匹配和视差求取。视觉 匹配算法研究发展到至今,已经产生了许多种算法, 归纳起来有:基于图像相关和基于特征匹配的方 法<sup>[122]</sup>,以及基于代价函数的方法<sup>[123-125]</sup>等。



图 16 光场相机的发展历程和部分产品结构示例<sup>[121]</sup> Fig. 16 Development trends of light field camera and some products illustration<sup>[121]</sup>

# 4 机器人视觉 3D 成像方法比较

前节内容总结了当前机器人 3D 视觉几种典型的成像方法和特点,限于篇幅问题,细节内容没有展开。本节在前节内容的基础上,对各种 3D 成像方法进行比较,探讨适合机器人 Eye-in-Hand 机器人的成像方法。

#### 4.1 机器人 3D 成像方法图谱

为了比较各种 3D 成像方法的优势与局限性, 现将上述各种方法绘成下列图谱<sup>[125-128]</sup>,如图 17 所 示,并将各类的优缺点标注在右边。从图 17 中可以 看出,扫描和干涉适合于 Eye-in-Hand 系统执行高 精度三维成像测量,不适合机械手臂实时运动引导 和操作(抓取、装配等)。因此,对于 Eye-in-Hand 机 器人系统,优选方案应从 ToF 相机、光场相机、结构 光投影和立体视觉几类中根据具体应用情况进行选 择。

图 18 列出了目前在工业界比较受欢迎的一些 方案,并把前几年比较热门的光场相机从立体视觉 里单独列出。各类 3D 方法的性能比较,详见下节。 4.2 方法比较

1) 类似于 ToF 相机、光场相机这类相机,可以 归类为单相机 3D 成像范围,它们体积小,实时性 好,适合 Eye-in-Hand 系统执行 3D 测量、定位和实 时引导。但是,ToF 相机、光场相机短期内还难以 用来构建普通的 Eye-in-Hand 系统,主要原因如下:

①ToF相机空间分辨率和 3D 精度低,不适合

高精度测量、定位与引导。

②对于光场相机,目前商业化的工业级产品只 有德国 Raytrix 一家,虽然性能较好,空间分率和精 度适中,但价格太贵,一台几十万元,使用成本太高。

2)结构光投影 3D 系统,精度和成本适中,有相 当好的应用市场前景。它由若干个相机-投影仪组 成,如果把投影仪当作一个逆向的相机,可以认为该 系统是一个双目或多目 3D 三角测量系统。

3)被动立体视觉 3D 成像,目前在工业领域也 得到较好应用,但应用场合有限。因为单目立体视 觉实现有难度,双目和多目立体视觉要求目标物体 纹理或几何特征清晰。

4)结构光投影 3D、双目立体视觉 3D都存在下列缺点:体积较大,容易产生遮挡。因为这几种方法都是基于三角测量原理,要求相机和投影仪之间或双目立体两个相机之间必须间隔一定距离,并且存在一定的夹角θ(通常大于 15°)才能实现测量。如图 19 所示,如果被测物体表面陡峭或有台阶,会引起相机成像遮挡,即相机不能捕捉到这些结构光的照射区域,导致存在不可测量区域,如图 19(a)所示阴影区域。如果减小相机与投影仪(结构光光源)的夹角,虽然在某些程度上可以解决问题,但是却会严重降低系统的测量灵敏度,影响该测量系统的应用。针对上述问题虽然可以增加投影仪或相机覆盖被遮挡的区域,构成投影仪-相机.投影仪系统、相机-投影仪-相机测量系统或者多个相机投影仪系统,如图 19(b)(c)所示,增大可视范围,减小阴影区域,扩大



图 17 机器人 3D 视觉成像方法图谱

Fig. 17 Genealogy of robot 3D vision imaging methods

测量区域,但会增加成像系统的体积,减少在 Eyein-Hand 系统中应用的灵活性。从 Eye-in-Hand 系 统的角度来看,最佳的方案是开发一种成本低廉、精 度适中、被动单目 3D 成像系统。

# 5 结束语

虽然光学 3D 视觉成像测量方法种类繁多,但 能够安装在工业机器人上,组成一种合适的 Eye-in-





Fig. 18 Optimization scheme of 3D vision imaging system for eye-in-hand robots







Hand 系统,对位置变动的目标执行 3D 成像测量、 引导机械手臂准确定位和实施精准操作的方法有 限。因为从工业应用的角度来说,我们更关心的是 3D 视觉传感器的精度、速度、体积与重量。鉴于机 器人末端能够承受的端载荷有限,允许传感器占用 的空间有限,传感器在满足成像精度的条件下,重量 越轻体积越小也就越实用。所以,对于 Eye-in-Hand 机器人系统,最佳 3D 成像方法是采用被动单 目(单相机)3D 成像方法,这样不仅体积小、重量轻, 也解决了双目和多目多视图遮挡难题。但目前还缺 少一种成本低廉、精度适中的单目 3D 成像技术。

从上面分析中可以看出:单目视觉深度感知线 索有一种为焦距差异,利用这种差异有望实现单目 3D成像<sup>[129]</sup>。目前在这方面已经有些相关研究,如 利用 focus/defocus(blurring)成像技术获得深度感知信息<sup>[130-135]</sup>,在显微 3D 测量领域利用显微物镜的小景深实现变焦 3D 测量<sup>[136-137]</sup>,焦深叠加获得显微 3D<sup>[138]</sup>,利用焦深叠加扩展成像景深<sup>[139]</sup>等。基于这些视觉感知原理,也许我们能开发出一种新的机器 人单目视觉三维感知方法。

#### 参考文献

[1] Lu R S, Wu A, Zhang T D, et al. Review on automated optical (visual) inspection and its applications in defect detection [J]. Acta Optica Sinica, 2018, 38(8): 0815002.
卢荣胜, 吴昂, 张腾达, 等. 自动光学(视觉)检测 技术及其在缺陷检测中的应用综述[J]. 光学学报, 2018, 38(8): 0815002.

- He Z X, Wu C R, Zhang S Y, et al. Moment-based
   2. 5-D visual servoing for textureless planar part grasping [J]. IEEE Transactions on Industrial Electronics, 2019, 66(10): 7821-7830.
- [3] Malis E, Chaumette F, Boudet S. 2 1/2 D visual servoing [J]. IEEE Transactions on Robotics and Automation, 1999, 15(2): 238-250.
- [4] Kragic D, Christensen H I. Survey on visual servoing for manipulation [R]. Report from Computational Vision and Active Perception Laboratory (CVAP), 2002, 1-59.
- [5] Dong G Q, Zhu Z H. Kinematics-based incremental visual servo for robotic capture of non-cooperative target [J]. Robotics and Autonomous Systems, 2019, 112: 221-228.
- [6] Xu D, Tan M, Li Y. Visual measurement and control for robots: 3rd ed. [M] Beijing: National Defense Industry Press. 2016.
  徐德,谭民,李原. 机器人视觉测量与控制(第3版)[M].北京:国防工业出版社, 2016.
- [7] Hashimoto K. A review on vision-based control of robot manipulators [J]. Advanced Robotics, 2003, 17(10): 969-991.
- [8] Ozato A, Maru N. Position and attitude control of eye-in-hand system by visual servoing using binocular visual space[C] // 2014 World Automation Congress (WAC), August 3-7, 2014. Waikoloa, HI. IEEE, 2014: 7-12.
- [9] Walck G, Drouin M. Progressive 3D reconstruction of unknown objects using one eye-in-hand camera[C] // 2009 IEEE International Conference on Robotics and Biomimetics (ROBIO), Guilin, 2009: 971-976.
- [10] Nobakht H, Liu Y. A hybrid positioning method for eye-in-hand industrial robot by using 3D reconstruction and IBVS [C] // 2015 IEEE International Conference on Robotics and Biomimetics (ROBIO), December 6-9, 2015. Zhuhai. IEEE, 2015: 2609-2614.
- Hentout A, Bouzouia B, Akli I, et al. Multi-agent control architecture of mobile manipulators: extraction of 3D coordinates of object using an eye-in-hand camera [C] // 2009 3rd International Conference on Signals, Circuits and Systems (SCS), November 6-8, 2009. Medenine, Tunisia. IEEE, 2009: 1-6.
- [12] Shaw J, Cheng K Y. Object identification and 3D

position calculation using eye-in-hand single camera for robot gripper [C] // 2016 IEEE International Conference on Industrial Technology (ICIT), March 14-17, 2016. Taipei, Taiwan, China. IEEE, 2016: 1622-1625.

- [13] Luo R C, Chou S C, Yang X Y, et al. Hybrid eyeto-hand and eye-in-hand visual servo system for parallel robot conveyor object tracking and fetching[C] // IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society, October 29-November 1, 2014. Dallas, TX, USA. IEEE, 2014: 2558-2563.
- [14] Flandin G, Chaumette F, Marchand E. Eye-inhand/eye-to-hand cooperation for visual servoing[C] // Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065), San Francisco, CA, USA. IEEE, 2000, 3: 2741-2746.
- [15] Muis A, Ohnishi K. Eye-to-hand approach on eyein-hand configuration within real-time visual servoing[J]. ASME Transactions on Mechatronics, 2005, 10(4): 404-410.
- [16] Lippiello V, Siciliano B, Villani L. Eye-in-hand/ eye-to-hand multi-camera visual servoing [C] // Proceedings of the 44th IEEE Conference on Decision and Control, Seville, Spain. IEEE, 2005: 5354-5359.
- [17] Zhao Y S, Gong L, Huang Y X, et al. A review of key techniques of vision-based control for harvesting robot [ J ]. Computers and Electronics in Agriculture, 2016, 127: 311-323.
- [18] Elarbi-Boudihir M, Al-Shalfan K A. Eye-in-hand/ eye-to-hand configuration for a WMRA control based on visual servoing [C] // 2013 IEEE 11th International Workshop of Electronics, Control, Measurement, Signals and their application to Mechatronics, June 24-26, 2013. Toulouse Cedex 7, France. IEEE, 2013: 1-8.
- [19] Brown G M. Overview of three-dimensional shape measurement using optical methods [J]. Optical Engineering, 2000, 39(1): 10-22.
- [20] Bi Z M, Wang L H. Advances in 3D data acquisition and processing for industrial applications [J]. Robotics and Computer-Integrated Manufacturing, 2010, 26(5): 403-413.
- [21] Aggarwal J K, Xia L. Human activity recognition from 3D data: a review [J]. Pattern Recognition

Letters, 2014, 48: 70-80.

- [22] Maier-Hein L, Mountney P, Bartoli A, et al. Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery [J]. Medical Image Analysis, 2013, 17(8): 974-996.
- [23] Schöning J, Heidemann G. Taxonomy of 3D sensors - a survey of state-of-the-art consumer 3Dreconstruction sensors and their field of applications[C] // Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, February 27-29, 2016. Rome, Italy. SCITEPRESS-Science and and Technology Publications, 2016.
- [24] Remondino F, Editors D S. TOF range-imaging cameras[M]. London: Springer, 2013.
- [25] Besl P J. Active, optical range imaging sensors [J].
   Machine Vision and Applications, 1988, 1(2): 127-152.
- [26] Lange R, Seitz P. Seeing distances-a fast time-offlight 3D camera[J]. Sensor Review, 2000, 20(3): 212-217.
- [27] PMD Technologies Website [EB/OL]. [2020-01-20]. www. pmdtec. com.
- [28] SoftKinetic Website [EB/OL]. [2020-01-20]. www.softkinetic.com.
- [29] Reiser U, Kubacki J. Using a 3D time-of-flight range camera for visual tracking [J]. IFAC Proceedings Volumes, 2007, 40(15): 355-360.
- [30] Ollikkala A V H, Makynen A J. Range imaging using a time-of-flight 3D camera and a cooperative object [C] // 2009 IEEE Intrumentation and Measurement Technology Conference, May 5-7, 2009. Singapore. Singapore. IEEE, 2009: 817-821.
- [31] Alenyà G, Foix S, Torras C. ToF cameras for active vision in robotics [J]. Sensors and Actuators A: Physical, 2014, 218: 10-22.
- [32] Piatti D, Rinaudo F. SR-4000 and CamCube3. 0 time of flight (ToF) cameras: tests and comparison
   [J]. Remote Sensing, 2012, 4(4): 1069-1089.
- [33] Hebert M. Active and passive range sensing for robotics[C] // Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065), San Francisco, CA, USA. IEEE, 2000, 1: 102-110.
- [34] Chougule V N, Gosavi H S, Dharwadkar M M, et al. Review of different 3D scanners and scanning techniques [C] // IOSR Journal of Engineering

(IOSRJEN), 7th National conference on Recent Developments In Mechanical Engineering RDME-2018, 2018: 41-44.

- [35] de Sousa G B, Olabi A, Palos J, et al. 3D metrology using a collaborative robot with a laser triangulation sensor [J]. Procedia Manufacturing, 2017, 11: 132-140.
- [36] Ettl S. Introductory review on 'Flying Triangulation ': a motion-robust optical 3D measurement principle[J]. Contemporary Physics, 2015, 56(2): 144-158.
- [37] Ettl S, Arold O, Yang Z, et al. Flying triangulation: an optical 3D sensor for the motionrobust acquisition of complex objects [J]. Applied Optics, 2012, 51(2): 281-289.
- [38] Arold O, Ettl S, Willomitzer F, et al. Hand-Guided 3D Surface Acquisition by Combining Simple Light Sectioning with Real-Time Algorithms [EB/OL]. [2020-01-20]. https: // arxiv. org/abs/1401. 1946.
- [39] Hillenbrand M, Mitschunas B, Brill F A G, et al. Spectral characteristics of chromatic confocal imaging systems [J]. Applied Optics, 2014, 53: 7634-7642.
- [40] Hillenbrand M, Lorenz L, Kleindienst R, et al. Spectrally multiplexed chromatic confocal multipoint sensing[J]. Optics Letters, 2013, 38(22): 4694-4697.
- [41] Hillenbrand M, Weiss R, Endrödy C, et al. Chromatic confocal matrix sensor with actuated pinhole arrays[J]. Applied Optics, 2015, 54(15): 4927-4936.
- [42] Beyerer J, Puente F, Frese L C. Chapter 7. Methods of image acquisition, in book: machine vision -- automated visual inspection: theory, practice and applications [M]. London: Springer, 2016, 307-308.
- [43] Taphanel M, Zink R, Langle T, et al. Multiplex acquisition approach for high speed 3d measurements with a chromatic confocal microscope[J]. Proceedings of SPIE, 2015, 9525: 95250Y.
- [44] Taphanel M, Beyerer J. Fast 3D in-line sensor for specular and diffuse surfaces combining the chromatic confocal and triangulation principle[C] // 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings, May 13-16, 2012. Graz, Austria. IEEE, 2012:

1072 - 1077.

- [45] Lin P C, Sun P C, Zhu L J, et al. Single-shot depth-section imaging through chromatic slit-scan confocal microscopy[J]. Applied Optics, 1998, 37 (28): 6764-6770.
- [46] Chun B S, Kim K, Gweon D. Three-dimensional surface profile measurement using a beam scanning chromatic confocal microscope [J]. The Review of Scientific Instruments, 2009, 80(7): 073706.
- [47] Zhong K, Li Z W, Zhou X H, et al. Enhanced phase measurement profilometry for industrial 3D inspection automation[J]. The International Journal of Advanced Manufacturing Technology, 2015, 76 (9/10/11/12): 1563-1574.
- [48] Servin M, Padilla M, Garnica G, et al. Profilometry of three-dimensional discontinuous solids by combining two-steps temporal phase unwrapping, co-phased profilometry and phaseshifting interferometry [J]. Optics and Lasers in Engineering, 2016, 87: 75-82.
- [49] Servin M, Garnica G, Padilla J M. 360-Degree Profilometry of Discontinuous Solids Co-Phasing 2-Projectors and 1-Camera [EB/OL]. [2020-01-20]. https: // arxiv. org/abs/1408. 6463.
- [50] Servin M, Garnica G, Estrada J C, et al. Coherent digital demodulation of single-camera N-projections for 3D-object shape measurement: Co-phased profilometry[J]. Optics Express, 2013, 21(21): 24873.
- [51] Application Report, Getting Started with TI DLP Display Technology [EB/OL]. [2020-01-20]. http: // www. ti. com. cn/cn/lit/an/dlpa059c/ dlpa059c. pdf.
- [52] Sun W S, Chiang Y C, Tsuei C H. Optical design for the DLP pocket projector using LED light source[J]. Physics Procedia, 2011, 19: 301-307.
- [53] Ishiyama H, Terabayashi K, Umeda K. A 100 Hz real-time sensing system of textured range images[C] // International Symposium on Optomechatronic Technologies. IEEE, 2011.
- [54] van der Jeught S, Dirckx J J J. Real-time structured light profilometry: a review [J]. Optics and Lasers in Engineering, 2016, 87: 18-31.
- [55] Geng J. Structured-light 3D surface imaging: a tutorial [J]. Advances in Optics and Photonics, 2011, 3(2): 128-160.
- [56] Fernandez S. One-shot pattern projection for dense and accurate 3D acquisition in structured light[D].

Girona : University of Girona, 2012.

- [57] Salvi J, Fernandez S, Pribanic T, et al. A state of the art in structured light patterns for surface profilometry [J]. Pattern Recognition, 2010, 43 (8): 2666-2680.
- [58] Lin H B, Nie L, Song Z. A single-shot structured light means by encoding both color and geometrical features[J]. Pattern Recognition, 2016, 54: 178-189.
- [59] Chen L C, Nguyen X L. Dynamic 3D surface profilometry using a novel colour pattern encoded with a multiple triangular model[J]. Measurement Science and Technology, 2010, 21(5): 054009.
- [60] Zhang Z H. Review of single-shot 3D shape measurement by phase calculation-based fringe projection techniques [J]. Optics and Lasers in Engineering, 2012, 50(8): 1097-1106.
- [61] Tang S M, Zhang X, Song Z, et al. Robust pattern decoding in shape-coded structured light[J]. Optics and Lasers in Engineering, 2017, 96: 50-62.
- [62] Khoshelham K, Elberink S O. Accuracy and resolution of kinect depth data for indoor mapping applications[J]. Sensors, 2012, 12(2): 1437-1454.
- [63] Lin C, Li Y, Xu G, et al. Optimizing ZNCC calculation in binocular stereo matching [J]. Signal Processing: Image Communication, 2017, 52: 64-73.
- [64] Birchfield S, Tomasi C. A pixel dissimilarity measure that is insensitive to image sampling [J].
   IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(4): 401-406.
- [65] Mei X, Sun X, Zhou M C, et al. On building an accurate stereo matching system on graphics hardware[C] // 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), November 6-13, 2011. Barcelona, Spain. IEEE, 2011: 467-474.
- [66] Hirschmuller H. Accurate and efficient stereo processing by semi-global matching and mutual information [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA. IEEE, 2005, 2: 807-814.
- [67] Hirschmuller H. Stereo processing by semiglobal matching and mutual information [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 328-341.
- [68] Hong L, Chen G. Segment-based stereo matching

using graph cuts[C] // Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA. IEEE, 2004.

- [69] Middlebury Stereo Evaluation Version 3[EB/OL].
  [2020-01-20]. http: // vision. middlebury. edu/ stereo/eval3/.
- [70] Salvi J, Pagès J, Batlle J. Pattern codification strategies in structured light systems [J]. Pattern Recognition, 2004, 37(4): 827-849.
- [71] Ayubi G A, di Martino J M, Flores J L, et al. Binary coded linear fringes for three-dimensional shape profiling [J]. Optical Engineering, 2012, 51 (10): 103601.
- [72] Zhang Q C, Su X Y, Xiang L Q, et al. 3D shape measurement based on complementary Gray-code light[J]. Optics and Lasers in Engineering, 2012, 50(4): 574-579.
- [73] Zuo C, Feng S J, Huang L, et al. Phase shifting algorithms for fringe projection profilometry: a review[J]. Optics and Lasers in Engineering, 2018, 109: 23-59.
- [74] Mao C L, Lu R S, Dong J T, et al. Overview of the 3D profilometry of phase shifting fringe projection
   [J]. Acta Metrologica Sinica, 2018, 39(5): 628-640.

毛翠丽,卢荣胜,董敬涛,等.相移条纹投影三维 形貌测量技术综述[J].计量学报,2018,39(5): 628-640.

- [75] Zuo C, Huang L, Zhang M L, et al. Temporal phase unwrapping algorithms for fringe projection profilometry: a comparative review[J]. Optics and Lasers in Engineering, 2016, 85: 84-103.
- [76] Talebi R, Johnson1J, Abdel-Dayem A, et al. Multiwavelength vs. binary code pattern unwrapping in fringe projection method [J]. Journal of Communication and Computer 2014(3): 291-304.
- [77] Posdamer J L, Altschuler M D. Surface measurement by space-encoded projected beam systems [J]. Computer Graphics and Image Processing, 1982, 18(1): 1-17.
- Moreno D, Taubin G. Simple, accurate, and robust projector-camera calibration [C] // 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, October 13-15, 2012. Zurich, Switzerland. IEEE, 2012: 464-471.
- [79] Mao C L, Lu R S, Liu Z J. A multi-frequency

inverse-phase error compensation method for projector nonlinear in 3D shape measurement [J]. Optics Communications, 2018, 419: 75-82.

- [80] Zheng D L, Da F P, Qian K M, et al. Phaseshifting profilometry combined with Gray-code patterns projection: unwrapping error removal by an adaptive median filter[J]. Optics Express, 2017, 25 (5): 4700-4713.
- [81] Wu Z J, Guo W B, Zhang Q C. High-speed threedimensional shape measurement based on shifting Gray-code light[J]. Optics Express, 2019, 27(16): 22631-22644.
- [82] Sansoni G, Carocci M, Rodella R. Threedimensional vision based on a combination of Graycode and phase-shift light projection: analysis and compensation of the systematic errors [J]. Applied Optics, 1999, 38(31): 6565-6573.
- [83] Balzer J, Werling S. Principles of shape from specular reflection [J]. Measurement, 2010, 43 (10): 1305-1317.
- [84] Häusler G, Knauer M C, Faber C, et al. Deflectometry: 3D-metrology from nanometer to meter [M] // Fringe 2009. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009: 1-6.
- [85] Zhang H W, Ji L S, Liu S G, et al. Threedimensional shape measurement of a highly reflected, specular surface with structured light method[J]. Applied Optics, 2012, 51(31): 7724-7732.
- [86] Zhang Z H, Wang Y M, Huang S J, et al. Threedimensional shape measurements of specular objects using phase-measuring deflectometry [J]. Sensors, 2017, 17(12): 2835.
- [87] Liu M M, Hartley R, Salzmann M. Mirror surface reconstruction from a single image [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(4): 760-773.
- [88] Savarese S, Chen M, Perona P. Local shape from mirror reflections [J]. International Journal of Computer Vision, 2005, 64(1): 31-67.
- [89] Knauer M C, Kaminski J, Hausler G. Phase measuring deflectometry: a new approach to measure specular free-form surfaces [C] // Optical Metrology in Production Engineering. International Society for Optics and Photonics, 2004.
- [90] Wang Z, Inokuchi S. Determining shape of specular surfaces[C] // The 8th Scandinavian Conference on Image Analysis, Tromso, Norway, May 25-28,

1993:1190-1187.

- [91] Ren H Y, Gao F, Jiang X Q. Iterative optimization calibration method for stereo deflectometry [J].
   Optics Express, 2015, 23(17): 22060-22068.
- [92] Xu Y J, Gao F, Zhang Z H, et al. A holistic calibration method with iterative distortion compensation for stereo deflectometry [J]. Optics and Lasers in Engineering, 2018, 106: 111-118.
- [93] Nguyen H, Nguyen D, Wang Z Y, et al. Realtime, high-accuracy 3D imaging and shape measurement[J]. Applied Optics, 2015, 54(1): A9-A17.
- [94] Zhang S. Recent progresses on real-time 3D shape measurement using digital fringe projection techniques[J]. Optics and Lasers in Engineering, 2010, 48(2): 149-158.
- [95] Howard I P, Rogers B J. Development and pathology of binocular vision[M] // Binocular Vision and Stereopsis. Oxford: Oxford University Press, 1996: 603-644.
- [96] Trucco E, Verri A. Chapter 7 Stereopsis [M] // Introductory techniques for 3D computer vision, Prentice Hall PTR Upper Saddle River, NJ, USA, 1998:139-177.
- [97] Ge G T, Cheng Z Q, Ke P P, et al. Depth map extracting based on geometric perspective: an applicable 2D to 3D conversion technology [C] // 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), October 14-16, 2017. Shanghai. IEEE, 2017: 1-5.
- [98] Watt S J, Akeley K, Ernst M O, et al. Focus cues affect perceived depth[J]. Journal of Vision, 2005, 5(10): 834-862.
- [99] Billiot B, Cointault F, Journaux L, et al. 3D image acquisition system based on shape from focus technique[J]. Sensors, 2013, 13(4): 5040-5053.
- [100] Schechner Y Y, Kiryati N. Depth from defocus vs. stereo: how different really are they? [C] // Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No. 98EX170), Brisbane, Qld., Australia. IEEE Comput. Soc, 1998.
- [101] Furukawa Y, Hernández C. Multi-view stereo: a tutorial[J]. Foundations and Trends O in Computer Graphics and Vision, 2015, 9(1/2): 1-148.
- [102] Moons T. 3D reconstruction from multiple images part 1: principles[J]. Foundations and Trends © in

Computer Graphics and Vision, 2008, 4(4): 287-404.

- [103] Wang G H, Cheng J. Three-dimensional reconstruction of hybrid surfaces using perspective shape from shading [J]. Optik, 2016, 127 (19): 7740-7751.
- [104] Durou J D, Falcone M, Sagona M. Numerical methods for shape-from-shading: a new survey with benchmarks [J]. Computer Vision and Image Understanding, 2008, 109(1): 22-43.
- [105] Kim H G R, Angelaki D E, DeAngelis G C. The neural basis of depth perception from motion parallax[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2016, 371(1697): 20150256.
- [106] Kellnhofer P, Didyk P, Ritschel T, et al. Motion parallax in stereo 3D [J]. ACM Transactions on Graphics, 2016, 35(6): 1-12.
- [107] Reskó B, Herbay D, Krasznail P, et al. 3D image sensor based on parallax motion [J]. Acta Polytechnica Hungarica, 2007, 4(4): 37-53.
- [108] Zhou F Q, Wang Y X, Chai X H, et al. Review on precise measurement technology based on mirror binocular vision [J]. Acta Optica Sinica, 2018, 38 (8): 0815003.
  周富强, 王晔昕, 柴兴华, 等. 镜像双目视觉精密测量技术综述[J]. 光学学报, 2018, 38(8): 0815003.
- [109] Solgi M, Weng J Y. WWN-8: incremental online stereo with shape-from-X using life-long big data from multiple modalities [J]. Procedia Computer Science, 2015, 53: 316-326.
- [110] Blake R, Wilson H. Binocular vision [J]. Vision Research, 2011, 51(7): 754-770.
- [111] Qian N. Binocular disparity and the perception of depth[J]. Neuron, 1997, 18(3): 359-368.
- [112] Bagga P J. Real time depth computation using stereo imaging [J]. Journal of Electrical and Electronic Engineering, 2013, 1(2): 51-54.
- [113] Mattoccia S, Updates S M, Outline S M. Stereo Vision: Algorithms and Applications [EB/OL]. [2020-01-20]. http: // vision. deis. unibo. it/~ smatt/Seminars/StereoVision. pdf.
- [114] Moulon P, Monasse P, Marlet R. Adaptive structure from motion with a contrario model estimation [M] // Computer Vision-ACCV 2012. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013: 257-270.

- [115] Seitz S M, Curless B, Diebel J, et al. A comparison and evaluation of multi-view stereo reconstruction algorithms [C] // 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1 (CVPR'06), New York, NY, USA. IEEE, 2006, 1: 519-528.
- [116] Hartley R, Zisserman A. Multiple view geometry in computer vision [ M ]. 2<sup>nd</sup> ed. Cambridge: Cambridge University Press, 2004.
- [117] Szeliski R. Computer vision: algorithms and applications[M]. London: Springer, 2011.
- [118] Raytrix. 3D Light Field Camera Technology for Professional Applications [EB/OL]. [2020-01-20]. https: // www. raytrix. de/wp-content/uploads/ delightful-downloads/2016/07/R42. pdf.
- [119] Lytro Light Field Camera[EB/OL]. [2020-01-20]. https: // priceonomics. com/lytro-light-fieldcamera/.
- [120] Ng R. Digital light field photography[D]. Stanford: Stanford University, 2006.
- Zhu H, Wang Q, Yu J Y. Light field imaging: models, calibrations, reconstructions, and applications [ J ]. Frontiers of Information Technology & Electronic Engineering, 2017, 18 (9): 1236-1249.
- [122] Abdelhamid M. Extracting depth information from stereo vision system, using a correlation and a feature based methods [D]. Clemson: Clemson University, 2011.
- [123] Scharstein D, Szeliski R, Zabih R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [C] // Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001), Kauai, HI, USA. IEEE Comput. Soc, 2002, 47 (1/2/3): 7-42.
- [124] Hirschmuller H, Scharstein D. Evaluation of stereo matching costs on images with radiometric differences [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(9): 1582-1599.
- [125] Boykov Y, Veksler O. Graph cuts in vision and graphics: theories and applications [M] // Handbook of Mathematical Models in Computer Vision. New York: Springer-Verlag, : 79-96.
- [126] Beraldin J A, Blais F, Cournoyer L, et al. Active 3D sensing [J]. NRC Publications Archive (NPArC), 2010: 1-21.

- [127] Sansoni G, Trebeschi M, Docchio F. State-of-theart and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation[J]. Sensors, 2009, 9(1): 568-601.
- [128] Zhang S. High-speed 3D shape measurement with structured light methods: a review [J]. Optics and Lasers in Engineering, 2018, 106: 119-131.
- [129] Read J C A. Visual perception: understanding visual cues to depth [J]. Current Biology, 2012, 22(5): R163-R165.
- [130] Mather G, Smith D R. R. Depth cue integration: stereopsis and image blur [J]. Vision Research, 2000, 40: 3501-3506.
- [131] Held R T, Cooper E A, Banks M S. Blur and disparity are complementary cues to depth [J]. Current Biology, 2012, 22(5): 426-431.
- [132] Trouvé-Peloux P, Champagnat F, Le Besnerais G, et al. Theoretical performance model for single image depth from defocus[J]. Journal of the Optical Society of America A, 2014, 31(12): 2650-2662.
- Jin H L, Favaro P. A variational approach to shape from defocus [M] // Computer Vision-ECCV 2002.
   Berlin, Heidelberg: Springer Berlin Heidelberg, 2002: 18-30.
- [134] Kumar H, Yadav A S, Gupta S, et al. Depth map estimation using defocus and motion cues[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 29(5): 1365-1379.
- [135] Martišek D. Fast shape-from-focus method for 3D object reconstruction[J]. Optik, 2018, 169: 16-26.
- [136] Danzl R, Helmli F, Scherer S. Focus variation-a robust technology for high resolution optical 3D surface metrology[J]. Strojniški Vestnik-Journal of Mechanical Engineering, 2011, 2011(3): 245-256.
- [137] Helmli F, Danzl R, Prantl M, et al. Ultra high speed 3D measurement with the focus variation method [M] // Fringe 2013. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014: 617-622.
- [138] Wlodek J, Gofron K J, Cai Y. Achieving 3D imaging through focus stacking[C] // Proceedings of the 13th International Conference on Synchrotron Radiation Instrumentation-SRI2018, AIP Conference Proceedings, 2019, 2054: 050001.
- [139] Qian Q, Guntur B K. Extending depth of field and dynamic range from differently focused and exposed images [J]. Multidimensional Systems and Signal Processing, 2016, 27(2): 493-509.