



© 2019 PENG LI

机器视觉与图像处理

第11讲 立体视觉（程距法）

李鹏

光电科学与工程学院，玉泉，教三-311

Email: peng_li@zju.edu.cn

HomePage: <http://person.zju.edu.cn/lipeng>

部分资料取自互联网，版权归原作者所有

© 2019 PENG LI

回顾

Inputs: 图像

Outputs: 提取的图像属性

- 第08讲 形态学图像处理
 - 膨胀与腐蚀(基本运算)、开运算与闭运算、击中击不中变换、边界提取、空洞填充
- 第09讲 图像分割（不连续性）
 - 边缘检测（点、线、边缘），一阶梯度、二阶导数；霍夫变换（直线、圆）
- 第10讲 图像分割（相似性）
 - 阈值处理(最大类间方差)、基于区域的处理（生长、分裂与聚合）、运动的应用
- 第11讲 立体视觉



© 2019 PENG LI

3D立体视觉：重构场景的三维几何信息

- 应用价值：包括移动机器人的自主导航系统，航空及遥感测量，工业自动化系统等
- 3D立体视觉技术测量原理：
 - 主动测量(Range data method)
 - 结构光(Structured Light)
 - 光点法
 - 光条法
 - 光面法
 - 飞行时间TOF
 - 脉冲式激光测距
 - 连续波激光测距
 - 非飞行时间——三角测距
 - 被动测量
 - 单目立体视觉
 - 聚焦法
 - 离焦法
 - 双目立体视觉(Binocular Stereo Vision)
 - 多目立体视觉



© 2019 PENG LI

主动测量，程距法 (Range data method)

- 主动测量(Range data method)
 - 结构光(Structured Light)
 - 光点法
 - 光条法
 - 光面法
 - 飞行时间TOF
 - 脉冲式激光测距
 - 连续波激光测距
 - 非飞行时间——三角测距
- 测距器(range finders)提供深度图
- 根据已知的深度图，用数值逼近的方法重建表面信息，根据模型建立场景中的物体描述，实现图像理解功能。
- 这类方法适用于严格控制下的环境(tightly controlled domains)，如工业自动化的应用方面。



© 2019 PENG LI

结构光测距成像

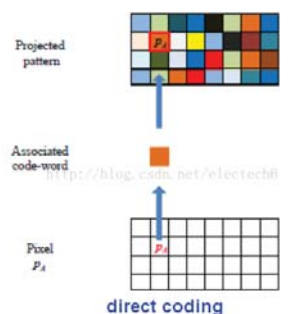
- 简单地来说，结构光的原理就是，将具有一定结构特征的光线投射到被摄物体上，红外相机根据事先标定的结果，还原出拍摄的图像深度。
- 结构光法不依赖于物体本身的颜色和纹理，采用了主动投影已知图案的方法来实现快速鲁棒的匹配特征点，能够达到较高的精度，也大大扩展了适用范围。



© 2019 PENG LI

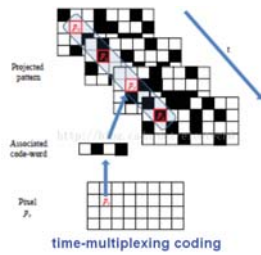
结构光-投射图案的编码方式

- 1、直接编码（direct coding）
- 根据图像灰度或者颜色信息编码，需要很宽的光谱范围。
- 优势：对所有点都进行了编码，理论上可以达到较高的分辨率。
- 缺点：受环境噪音影响较大，测量精度较差。



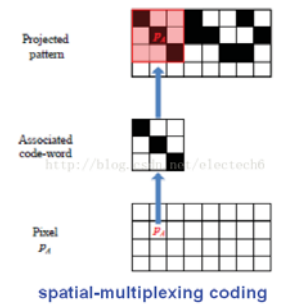
结构光-投射图案的编码方式

- 2、时分复用编码（time multiplexing coding）
- 顾名思义，该技术方案需要投影N个连续序列的不同编码光，接收端根据接收到N个连续的序列图像来每个识别每个编码点。投射的编码光有二进制码（最常用）、N进制码、灰度+相移等方案。
- 该方案的优点：测量精度很高（最高可达微米级）；可得到较高分辨率深度图（因为有大量的3D投影点）；受物体本身颜色影响很小（采用二进制编码）。
- 缺点：比较适合静态场景，不适用于动态场景计算量较大（因为识别一个编码点需要计算连续N次投影）。



结构光-投射图案的编码方式

- 3、空分复用编码（spatial multiplexing coding）
- 根据周围邻域内的一个窗口内所有的点的分布来识别编码。
- 该技术的优势：适用于运动物体。
- 缺点：不连续的物体表面可能产生错误的窗口解码（因为遮挡）。



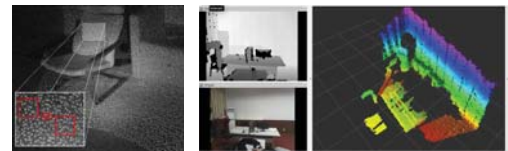
结构光-获取深度图方法

- 业界比较有名的结构光方案就是以色列PrimeSense公司的Light Coding的技术，该方案最早被应用于Microsoft的明星产品Kinect1（Kinect2是基于TOF的技术）上。下面以Kinect1为例，介绍一下其工作原理。



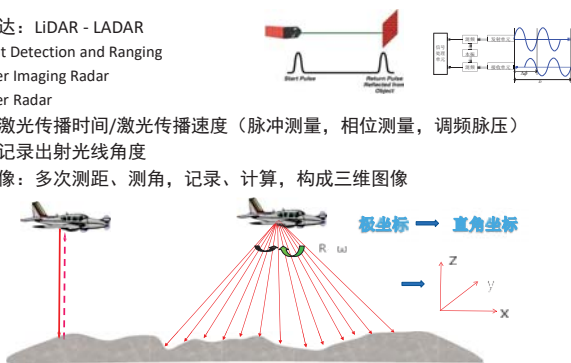
结构光-获取深度图方法

- Kinect1的红外光源投射人眼不可见的伪随机散斑红外光点到物体上，每个伪随机散斑光点和它周围窗口内的点集在空间分布中的每个位置都是唯一且已知的。这是因为Kinect1的存储器中已经预储存了所有的数据。
- 根据红外相机拍摄的图像解码每一个像素对应的空间深度
- RGB相机拍摄正常彩色图像，将RGB图像和解码的深度图像配准，得到一帧点云图



激光雷达技术

- 激光雷达：LiDAR - LADAR
 - Light Detection and Ranging
 - Laser Imaging Radar
 - Laser Radar
- 测距：激光传播时间/激光传播速度（脉冲测量，相位测量，调频脉压）
- 测角：记录出射光线角度
- 三维图像：多次测距、测角，记录、计算，构成三维图像

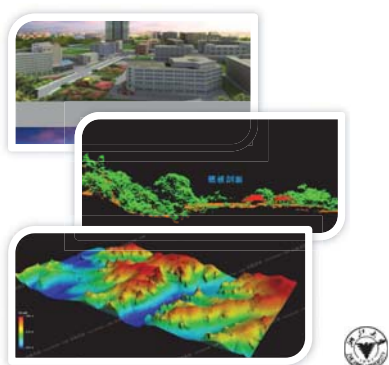


激光雷达发展史

- 1970'：大气探测，激光测距
- 1990'：机载测绘LiDAR
- （高精度IMU、GPS）
 - Optech, Riegl
- 2000'：机载 地面 车载（初创）
 - Riegl, Optech, Leica, FARO
- 2010'：机载 地面 车载（精小）
 - 距离精度 密度 尺寸 & 重量 专用软件及各种辅助设施
 - 专业应用（电力水利 交通 市政...）
 - Leica, Riegl, 北科天绘
 - FARO, Z+F
- 工业自动化LiDAR
 - 安全限界检测、机器人避障
- 导航型LiDAR（2003 ~）
 - DARPA无人驾驶挑战赛
 - 2004年，卡内基梅隆大学
 - 2005年，斯坦福 Sebastian Thrun
 - 2007年，卡耐基梅隆
- 无人驾驶：Google, Baidu
 - GM, Ford, Volvo, Delphi,
- 多线扫描
 - Velodyne, IBEO, 北科天绘
- 单线扫描
 - Hokuyo, SICK, Pioneer, 北科天绘

激光成像雷达特点及应用

- 激光雷达特点
 - 测量出三维—直接、准确
 - 全天候、屏蔽背景光干扰
 - 实时数据：三维+灰度
- 激光雷达应用：遥感、测绘
 - 国家基础测绘、国土监测
 - 大地三维测绘、应急救援
 - 潮间地、港口、海岸线测绘
 - 农业、林业、植被勘测
 - 数字城市、城市建筑测绘



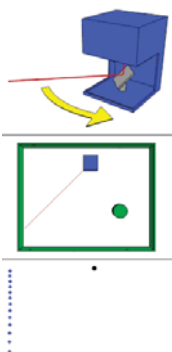
测量型LiDAR与导航型LiDAR比较

主要差异	测量型激光雷达	导航型激光雷达	备注
应用场景	专业测量	移动平台导航避障	
测距范围 (max)	100m - 3500m	6m - 200m	平台、配置产生差异
测距精度	1 mm - 5 mm	5 cm - 10 cm	导航型LiDAR测量精度低
视场范围	60° / 70° / 360°	100° / 270° / 360°	产品细分差异
指向分辨率/精度	0.001° / 0.005°	0.8 - 2° / NA	导航型LiDAR测量精度低
探测成像灰度	> 10 bits	2-4 bits	导航型LiDAR基本不具备目标表面的灰度探测能力
激光探测元素	1/2, 16-32线或Flash LiDAR	1, 2, 4, 8, 16, 32 ...	相对来说导航LiDAR更注重探测单元数量与扫描线数
设备尺寸、重量	4kg - 80 kg	0.6 kg - 10 kg	导航LiDAR以牺牲测距及精度为代价换取尺寸重量降低
作业要求	高精度，离线处理	高可靠，即时处理	
代表厂商	Riegl, Leica, Faro, 北科天绘	Velodyne, IBEO, Quanser, 北科天绘	北科天绘：16线高端导航LiDAR



激光雷达，一种工作在光学波段（近红外）的雷达

- LIDAR，即 Light Detection and Ranging
- LADAR，即 Laser Detection and Ranging，激光探测和测距
- 相比于微波雷达，激光雷达
 - 分辨率高**：频率高2-3个数量级以上，更高的距离分辨率、角分辨率和速度分辨率
 - 多维度信息**：测量过程中可带来距离、角度、反射强度、速度等更丰富的信息，生成目标多维度的图像，协助我们或系统对探测目标拥有更详细的认知
 - 抗干扰强**：激光束发散角非常小（μrad量级），多路径效应小，可实现低空、超低空目标的探测（与微波或者毫米波产生多路径效应）
 - 全天候的工作**：主动探测拥有不依赖于外界光照条件或目标本身辐射的特性，只需通过探测自身发射的激光束的回波信号来获取目标信息，可实现全天候的工作（但是易受大气条件以及工作环境的烟尘影响）

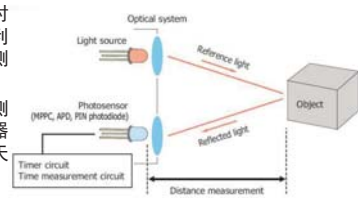


飞行时间测距——脉冲调制（Pulsed Modulation）

- 激光雷达中的激光器发射出一束**超短激光脉冲**
- 激光投射到目标物体上后发生漫反射传感器接收漫反射激光
- 通过激光光束在空中的飞行时间 Δt 准确计算得出目标物体到传感器间的距离

$$D = \frac{c\Delta t}{2}$$

- 光的传播速度 $c = 3 \times 10^8 \text{ m/s}$
- 脉冲激光的发射角小，能量在空间相对集中，瞬时功率极大可达兆瓦以上，利用这些特性可制成各种中远距离激光测距仪、激光雷达等。
- 脉冲式激光测距广泛应用于地形地貌测量、地质勘探、工程施工测量、飞行器高度测量、人造地球卫星相关测距、天体之间距离测量等遥测技术方面。

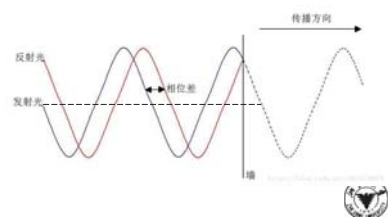


飞行时间测距——连续波调制（Continuous Wave Modulation）

- 相位式测距则是将一调制信号对发射光波的光强进行调制，通过测量相位差来间接测量时间，较直接测量往返时间的处理难度降低了许多。
- 此时距离的计算公式可表示为：

$$D = \frac{\lambda}{2} \times \frac{\varphi}{2\pi}$$

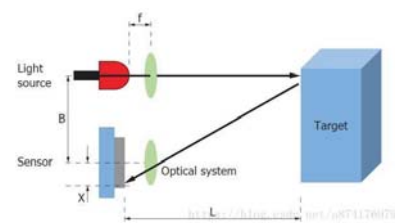
- 相位式激光测距通常适应于**中短距离**的测量，
- 测量精度**可达毫米、微米级，
- 是目前测距精度最高的一种方式
- 大部分中短程高精度测距仪都采用这种工作方式



三角测距法

- 三角测距法即光源、被测物面、光接收系统三点共同构成一个三角形光路
- 由激光器发出的光线，经过汇聚透镜聚焦后入射到被测物体表面上
- 光接收系统接收来自入射点处的散射光，并将其成像在光电位置探测器敏感面上
- 通过光点在成像面上的位移来测量被测物面移动距离的一种测量方法。

$$L = \frac{f(B + X)}{X}$$



三角测距法

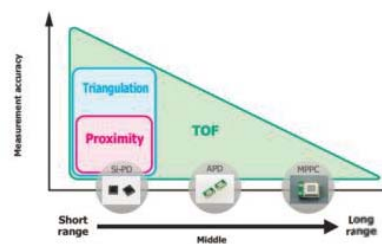
- 激光三角测距法具有结构简单、测试速度快、使用灵活方便等诸多优点，
- 但由于激光三角测距系统中，光接收器件接收的是待测目标面的散射光，所以对器件灵敏度要求很高。
- 另外，如激光亮度高、单色性好、方向性强，在**近距离**的测量中较为容易测量出光斑的位置。
- 因此三角法应用范围主要是**微位移**的测量，测量范围主要在微米、毫米、厘米数量级，
- 广泛应用于物体表面轮廓、宽度、厚度等量值的测量，例如汽车工业中车身模型曲面设计、激光切割、扫地机器人等。



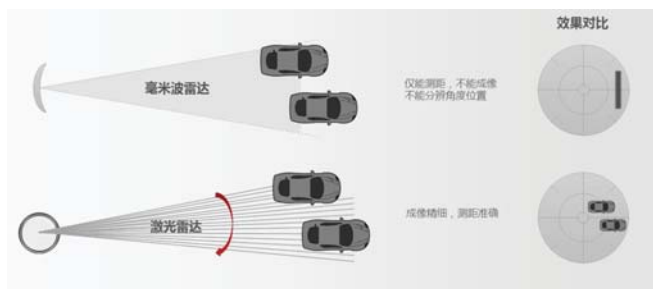
三种方法的量程和精度大致为：

- 相位测距法的测量速度最慢，一般每秒只能测量几个点到几十个点(TOF distance image sensor在几百个点左右)
- 三角测距速度居中，每秒在几百个到几千个点
- 脉冲测距法的测量速度最快，为每秒几万到几十万个点

- 因此，激光雷达一般采用三角测距法和脉冲测距法。
- 而短距离激光雷达采用三角测距法达到足够的精度
- 长距离激光雷达采用脉冲测距法达到足够的速度。



激光雷达vs.毫米波雷达



- 毫米波雷达的角分辨率较低，无法成像
- 激光雷达的角分辨率能到小数点后两位，够准确扫描物体形成 3D 图像



激光雷达vs.摄像头



激光雷达扫描效果

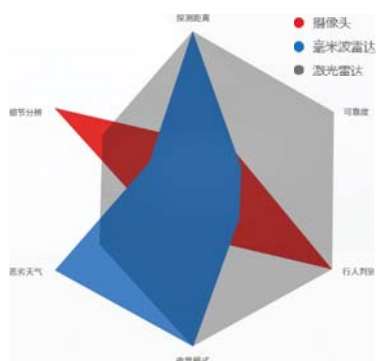


摄像头成像效果

- 摄像头测距，常用双目视觉系统，其准确度主要依赖复杂的算法
- 激光雷达利用飞行时间原理进行测距，百米范围，厘米级测量精度



激光雷达不可替代的作用

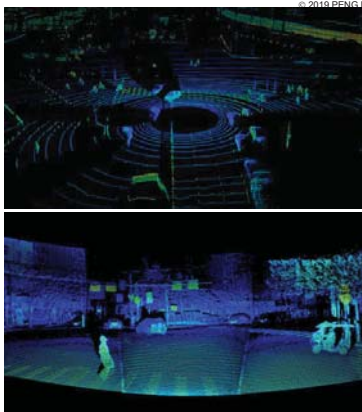


自动驾驶应用

- 自动驾驶中通常有多种雷达方式进行互补：
 - 视觉（单目+结构光、双目、三目）雷达
 - 激光雷达（单线、多线）
 - 微波（毫米波）雷达。
- 在目前的自动驾驶中，激光雷达有**单线**激光雷达和**多线**激光雷达之分
- 单线**激光雷达：一个高重频脉冲激光测距仪 + 一个一维旋转扫描来实现测量。
 - 而它的角分辨率可高于多线激光雷达，所以在行人探测、障碍物探测（小目标探测）以及前方障碍物探测等方面上，比多线激光雷达具有更多优势。
- 多线**激光雷达：多个激光发射器和接收器 + 一个维度上的高度旋转
 - 转速频率可以达到10Hz。
 - 垂直方向的视场角很小，只有20度左右
 - 垂直方向的点分辨率也比较稀疏
 - 主要应用在无人驾驶等实时性要求高、精度要求不高的领域，精度**厘米级**
- 多线方案目前也主要为多路单线集合而成，因而还受到体积和光路的限制



	禾赛Pandar64	禾赛PandarGT
测距原理	飞行时间测量法	飞行时间测量法
扫描原理	机械旋转	二维扫描
线数	64	
探测距离	0.3 m~200 m (10%反射率)	0.5 m~300 m (10%反射率)
测量精度	±5 cm (0.3 m~0.5 m), ±2 cm (0.5 m~200 m)	cm
回波模式	支持单回波和双回波	
扫描频率	10 Hz, 20 Hz	5 Hz, 10 Hz, 20 Hz
垂直视场角	40° (-25° ~15°)	20° (-10° ~+10°, 偏置 ±5° 可调)
垂直角分辨率	0.167° (-6° ~+2°), 最小0.07° (10 Hz)	
水平视场角	360°	60°
水平角分辨率	0.2° (10 Hz), 0.4° (20 Hz)	0.1°
是否完全抗干扰	是	是



无人驾驶应用

- 无人驾驶一般采用脉冲测距：
 - 量程足够远，通常情况下要达到100-120米，
 - 精度则在厘米级
 - 测速要求也高，单个激光发射的速率要达到几万个点每秒，通过多个激光发射器达到实时环境感知的目的
 - 成本会要高得多，开发难度也大
- 扫地机器人一般采用三角测距的方案：
 - 测量距离近（15米内），
 - 精度要求不高，
 - 测量速率低，一般都是几千个点每秒。
 - 技术的门槛不高，硬件成本也低
- 目前国外主流的激光雷达生产厂家，
 - Velodyne采取的是激光发射、接收一起旋转的方式；
 - IBEO采取的是固定激光光源，通过内部玻璃片旋转的方式改变激光光束方向，实现多角度检测的需要；
 - Quanergy采取相控阵技术，内部不存在任何旋转部件。



手机应用，TOF vs 结构光

- 2017年9月13日，苹果推出了基于3D**结构光**技术的iPhone X，实现了3D人脸识别，并以面部识别彻底取代了指纹识别
- 苹果公司目前正在尝试在 iPhone 背面添加一个使用 **TOF**（Time of Flight）技术的 3D 传感器
- OPPO、华为、小米等国内厂商也先后布局了飞行时间技术在手机上的应用，并推出相关产品。
- VIVO更在2017年3月，就开始在内部同时跟进结构光与TOF 这两项3D技术。



TOF 和 结构光

10倍 3倍

TOF 3D 传感器

结构光

vivo NEX 双屏版



Michelson干涉仪

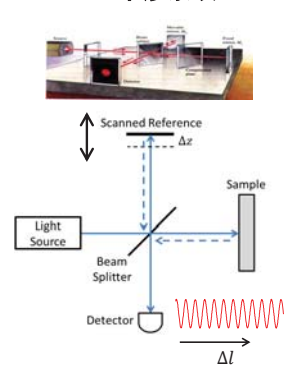
Albert Michelson

1877年，测量光速
1887年，改进光学干涉仪
1907年，荣获Nobel物理学奖

1887年 迈克尔逊-莫雷实验：
光的传播介质是“以太”，
地球绕太阳运动：每秒30公里
迎面有每秒30公里的“以太风”
实验表明“以太是随着地球运动”



Michelson干涉条纹



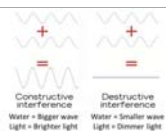
光程差:
 $\Delta l = 2\Delta z$

相位差:
 $\Delta \phi = k\Delta l = 2\pi\Delta l/\lambda$

探测光强:
 $I_d = I_{01} + I_{02} + 2\sqrt{I_{01}I_{02}}\cos(\Delta \phi)$

相长干涉:
 $\Delta \phi = 2\pi\Delta l/\lambda = 2m\pi$
 $\Delta l = m\lambda, (m=0,1,2,\dots)$

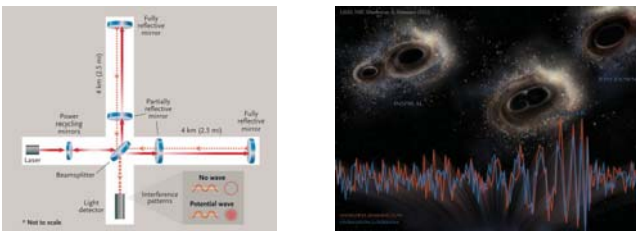
相消干涉:
 $\Delta \phi = 2\pi\Delta l/\lambda = (2m+1)\pi$
 $\Delta l = (m+1/2)\lambda, (m=0,1,2,\dots)$



近年，哪个Nobel奖与Michelson干涉仪相关？

2017年诺贝尔物理学奖：

测量引力波：Michelson干涉仪测应变Strain 约 10^{-18} 米的臂长变化

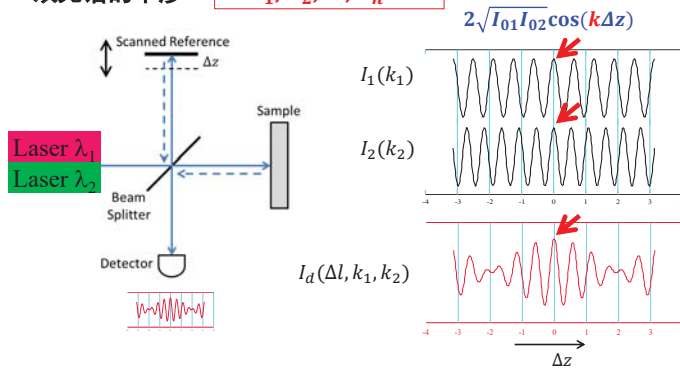


光程变化 $\ll \lambda$ ；难点在于：减小噪声干扰，提高探测灵敏度

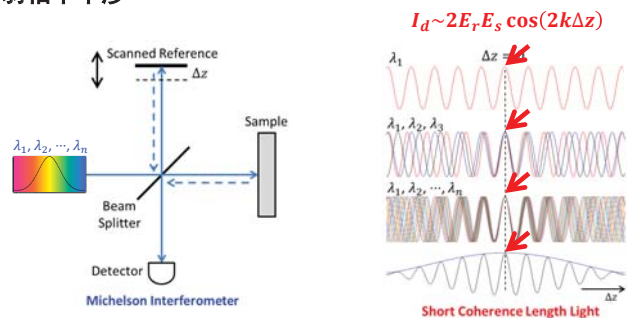
如何测量毫米量级 ($\gg \lambda$) 的空间位置变化？提高量程

双光谱的干涉

$\lambda_1, \lambda_2, \dots, \lambda_n$?

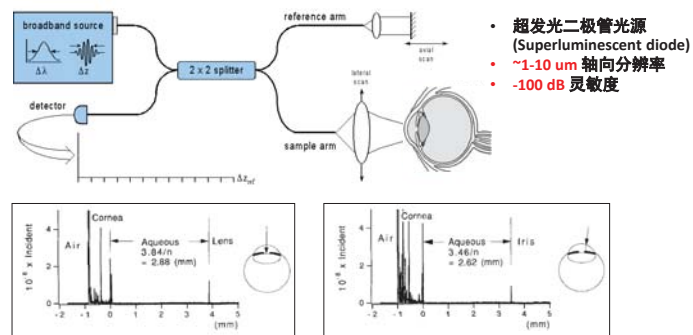


弱相干干涉



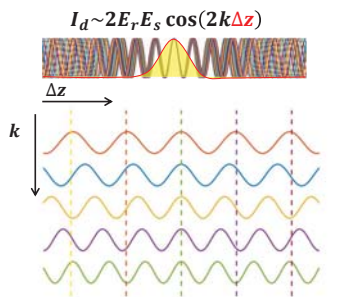
当且仅当光程差为0时，才能产生干涉包络

弱相干测距 (1D)



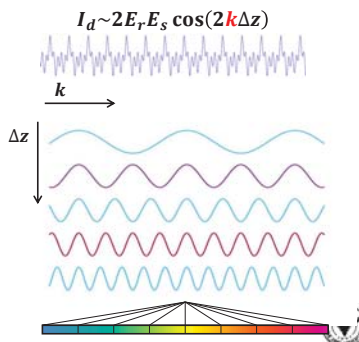
D. Huang, et al. Micron-resolution ranging of cornea anterior chamber by optical reflectometry, Laser Surg Med. (1991)

时域探测Time domain: 用时间编码空间（深度）

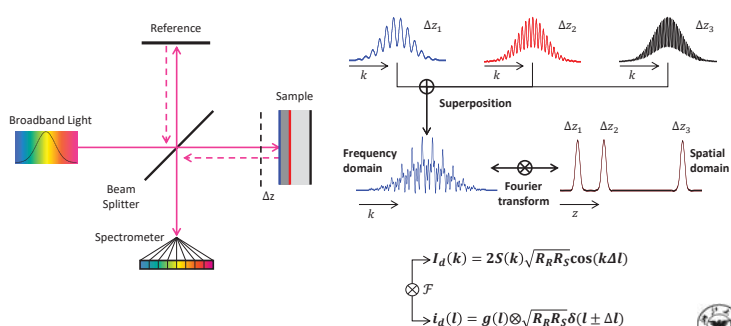


在不同的 Δz ，光条纹 k 有不同的变化频率

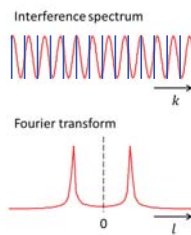
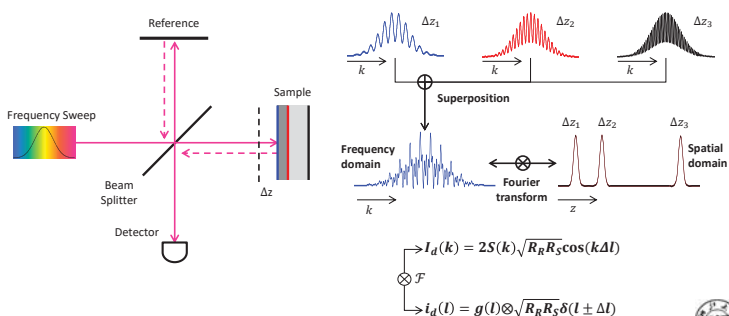
如何记录光谱干涉条纹？用频率编码空间（深度）



谱域探测Spectral domain: 用频率编码空间（深度）



扫频探测 Swept source: 用频率编码空间 (深度)



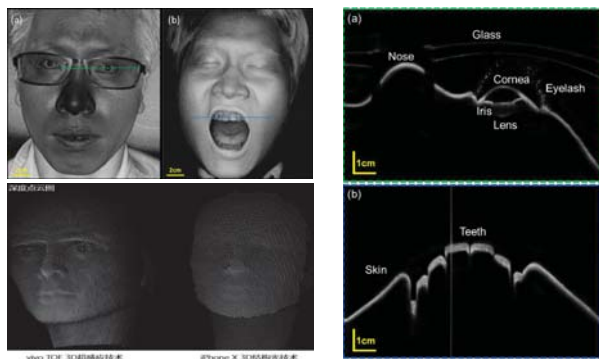
$$\text{分辨率: } \delta z = 0.44 \frac{\lambda_0^2}{\Delta \lambda}$$

如何实现:
厘米分辨率、百米量程?

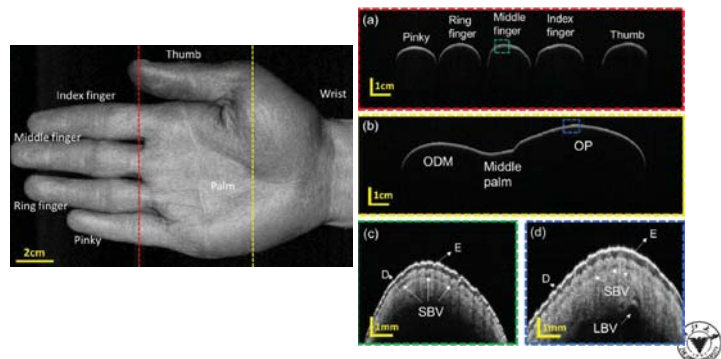
$$\text{量程: } \pm z_{\max} = \pm \frac{\lambda_0^2}{4 \cdot \delta_s \lambda}$$



光学相干成像, 3D人脸, ~10 cm量程



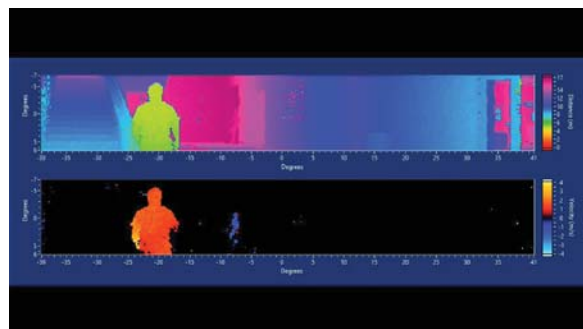
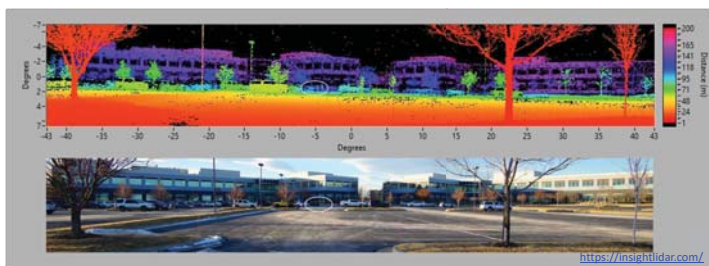
光学相干成像, 3D手掌, ~10 cm量程



光学相干成像雷达 Coherence LiDAR, ~200m量程

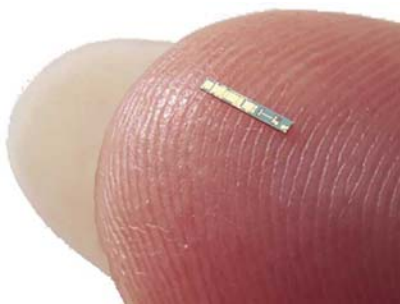
主要优势:

- 灵敏度提高10-100X
- 程长: 200 m+ 以及低反射率目标
- 分辨率提高40X
- 抗干扰提高10,000X (太阳光、雷达光)
- 每个像素自带速度信息
- 全半导体结构



The OCT technology of the future, 小型化

Swept lasers are being manufactured from wafer-scale semiconductors, dramatically reducing costs while improving performance. A single wafer carries over 2,000 lasers. Courtesy of Insight Photonic Solutions.



讨论:

1

- 面向无人驾驶, 可以采用哪些光学测距方案?
- 这些技术目前存在哪些局限性?
- 可能的解决方案?

2

- 假设现在我们利用kinect扫描一个房间内部, 得到了深度图像的视频。
- 由于kinect的视野有限, 如何从视频中恢复出房间的三维结构?
- 思考仅仅利用kient重建三维结构是否可行?

