Robot Vision: Depth sensing

Prof. Friedrich Fraundorfer

SS 2025

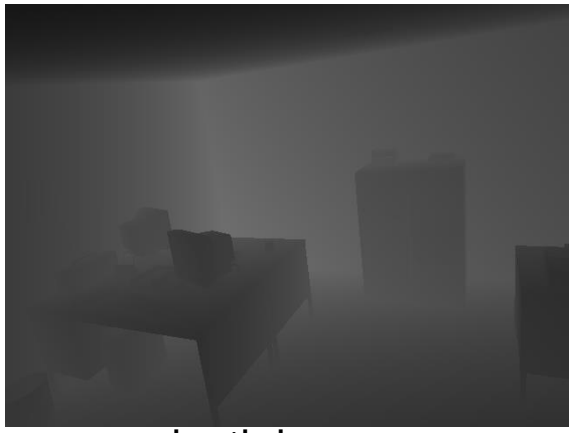
Outline

- Depth cameras
  - Coded light
  - Kinect style depth cameras
  - TOF cameras

Depth cameras - Overview

- Depth cameras or RGBD cameras directly output an RGB image and a depth image
- Principles:
  - Stereo cameras with onboard processing
    - DJI Guidance, Roboception, Perceptin
  - Structured Light
    - Coded light – Projector-camera system
    - Random patterns – Stereo system with active lighting
    - Kinect-style methods – Projector-camera system with fixed random projection
  - TOF cameras – time of flight principle
- Huge importance for mobile robotics

RGB image



depth image

Passive vs. active systems

- Passive systems: Stereo camera systems
  - Have problems with untextured areas (white walls)
  - Have problems with repeated structures (the correspondence problem is challenging)

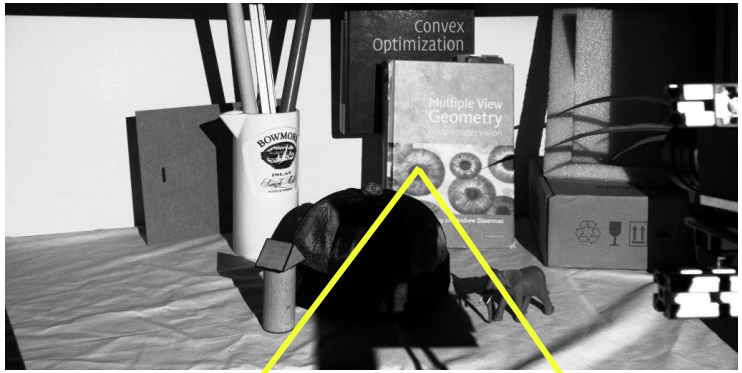
- Active systems: Structured Light, ToF cameras
  - Solve or make the correspondence problem easier by active illumination
  - Works on untextured areas and also in the dark

  - Can disturb the scene if visible light is used
  - Can have problems outdoors when sunlight is stronger than illumination

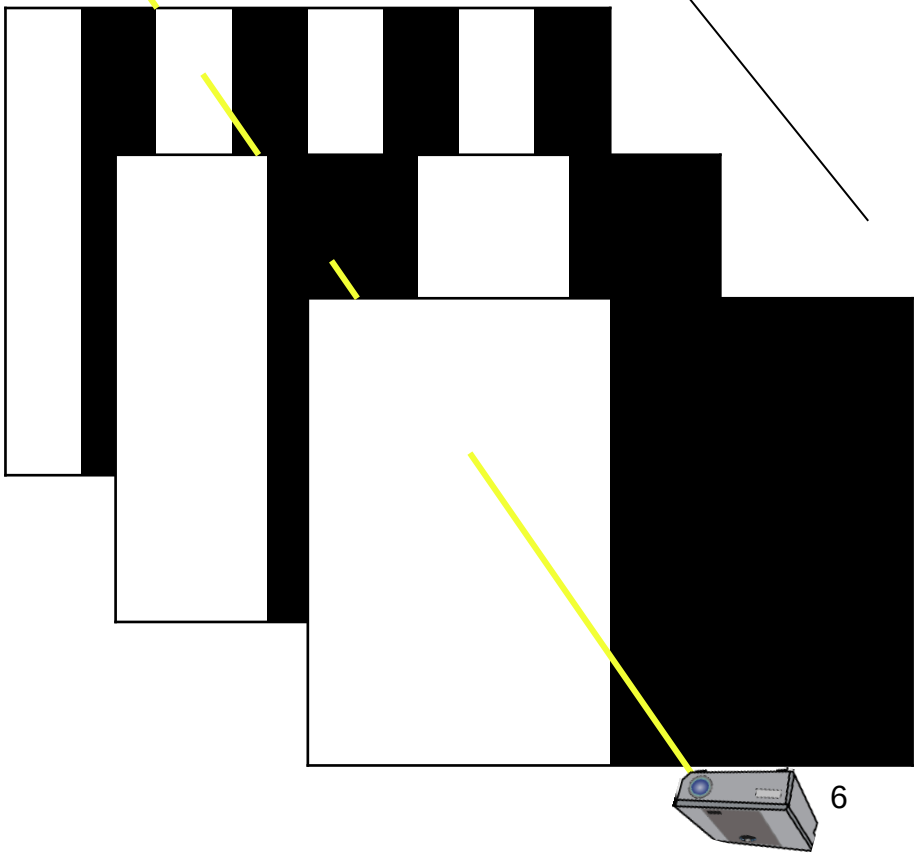
Coded Light

- A projector-camera system where a projector outputs stripe patterns (e.g. binary pattern)
- The pattern solves the correspondence problem in stereo matching
- Projector and camera need to be calibrated
- The stripes are coded and encode directly a unique position of a corresponding pixel in the projector.

Coded light



pattern changed
over time

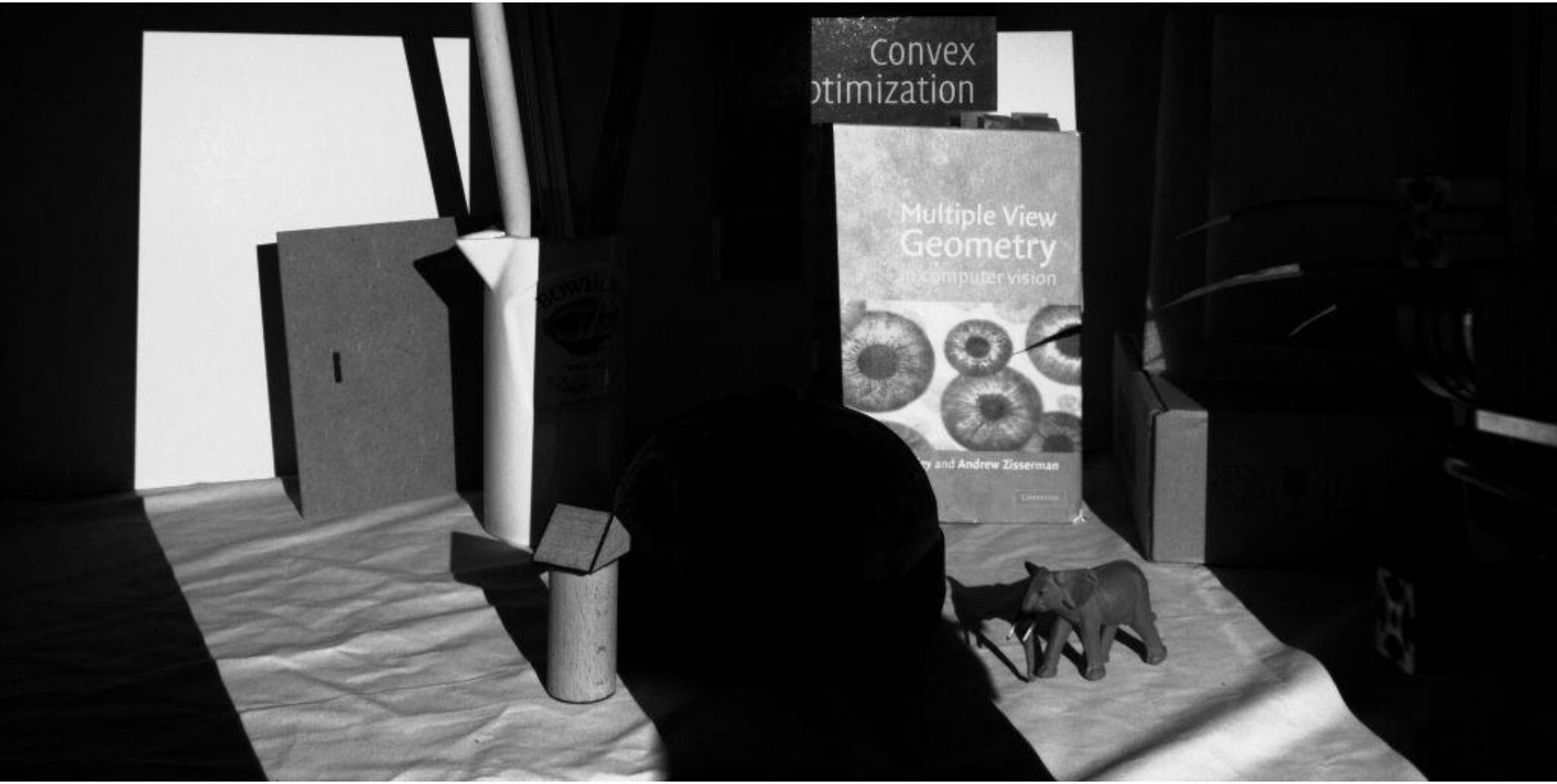


sequence of dark/light pattern defines a code word for a position and thus a unique position in the projector

Coded light

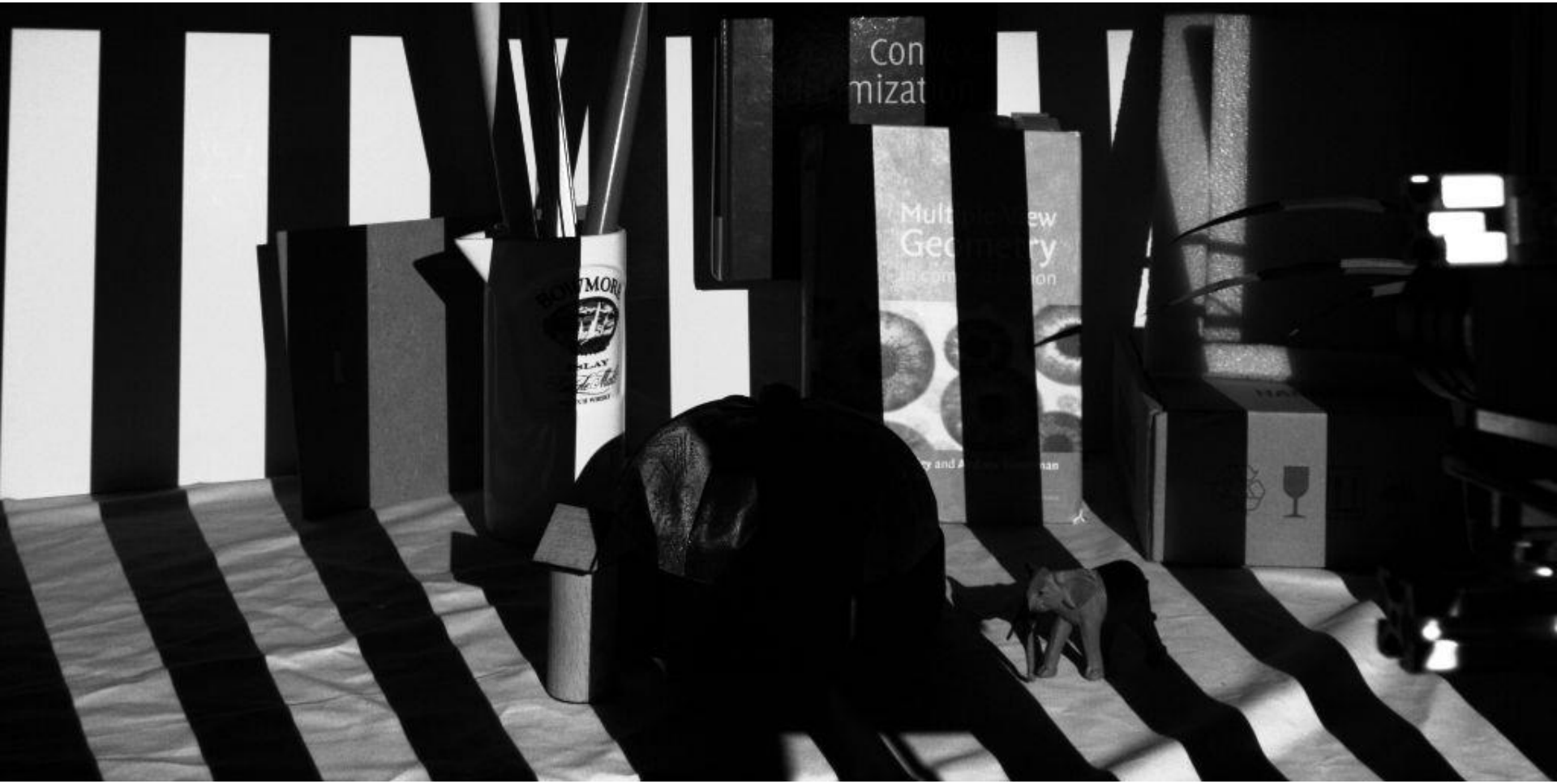

Convex
Optimization
Multiple View
Geometry
computer vision
and Andrew Zisserman

Coded light



Coded light

Coded light



Coded light

Coded light

Random patterns

- Stereo camera system cannot measure depth in textureless/homogeneous areas
- Solution: Project random pattern as texture to ease stereo matching
- Typically, this is done in infrared spectrum such that it is not visible for users
- Such a system works in the dark as well

- Example: Intel Realsense
- Standard stereo system (2 calibrated cameras)
- 1 IR projector for random dot pattern
- Works outdoors as well, however then the pattern is not visible due to strong sunlight (then it just works like standard stereo matching)

Kinect style method

- Kinect is a projector-camera system with onboard depth processing
- Projects a **known** static IR-dot pattern
- Depth is computed from a combination of depth from stereo and depth from focus
- The system also contains an RGB camera
- Sensors is often called a RGBD sensor

The elements of a depth sensor

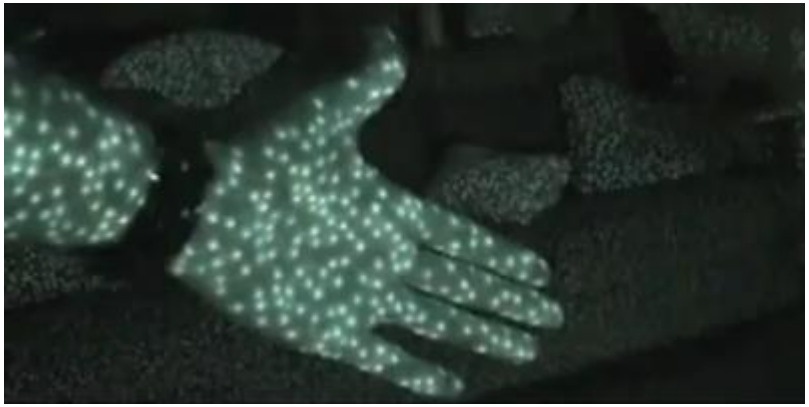


image of IR pattern

Time-of-flight cameras

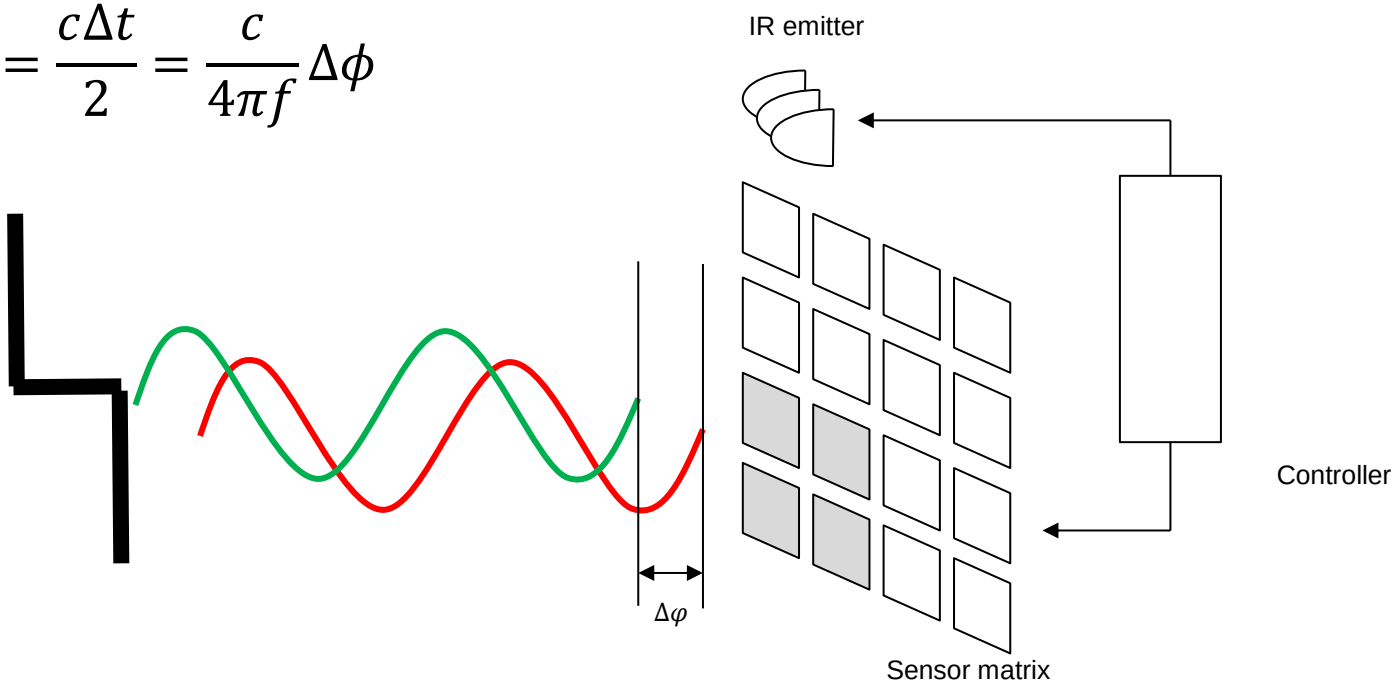
- Does not work with the stereo (triangulation) principle but with time-of-flight principle
- Principle:
  - Sends out NIR light (no spatial coding)
  - Sensor array measures response
  - Distance is measured by measuring time between emitting and receiving the light (pulsed or continuous wave method)

- Typically do not provide synchronized color image but a reflectance image
- Example: PMD Flex (224 x 171px resolution), Creative TOF sensor

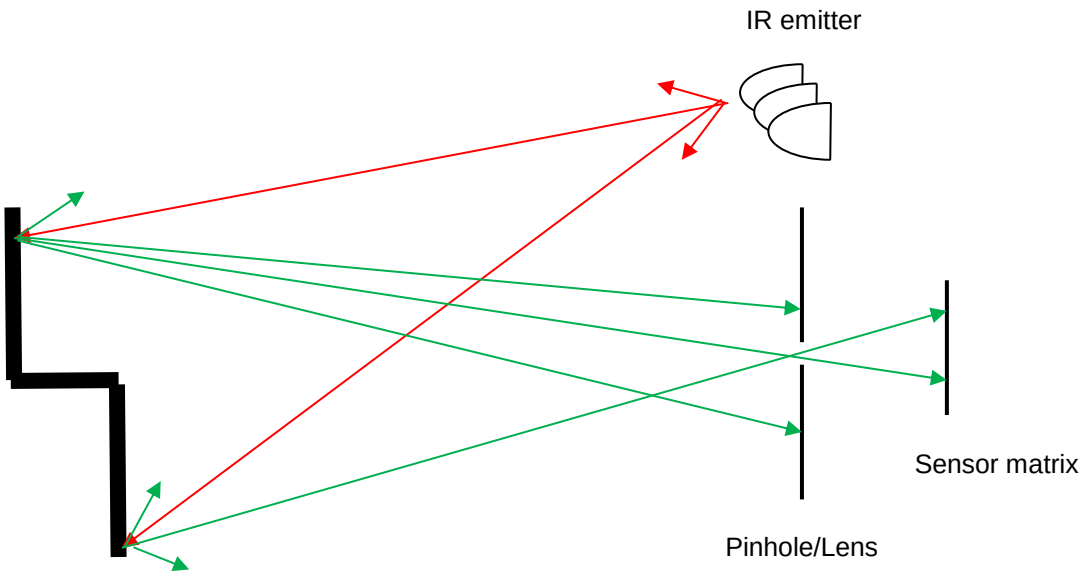
Time-of-flight cameras – Principle

- Depth can be computed from time or phase difference
- Continuous wave method:
  - Camera emits NIR light where amplitude is a sine wave.
  - Phase shift is measured between emitted and received light
  - Phase shift can be converted into distance

$$d = \frac{c\Delta t}{2} = \frac{c}{4\pi f}\Delta\phi$$



Time-of-flight cameras – Principle



Time-of-flight cameras – Multi-Path Interference (MPI)

- ToF cameras have an error source called Multi-Path Interference (MPI)