Visual Perception Sensors

Depth Determination

Gerrit Glaser

University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics

Technical Aspects of Multimodal Systems

November 13. 2017

Table of Contents

Motivation for Visual Perception in Robotics

- basic question for mobile robotics: Where am I?
- autonomous movement through unknown terrain
  - scan environment for obstacles
  - distances to surroundings

Possible solution

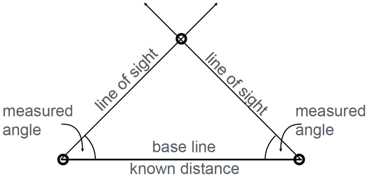
Add visual perception sensors, to allow robots to “see” their environment.

Camera

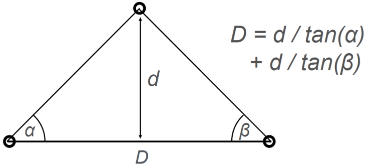
- image as projection of 3D world: leads to loss of depth data
- estimate depths through known size of an object and size of the object in the image.
- error-prone, even in human visual perception
- not applicable outside of known surroundings
- passive approach

Stump in Sequoia National Park. [1, p. 529, fig. 2]

- compute point through known distance and measured angles



Triangulation. [7, p. 19, fig. 1]



Triangulation Calculation. [7, p. 20, fig. 1]

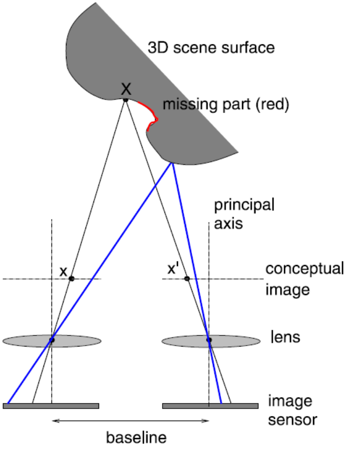
Stereoscopic Cameras

- one camera not sufficient for meaningful depth measurements
- use second camera to recover lost dimension
- triangulate distance
  - known baseline between cameras
  - corresponding points
  - measured angles
- passive approach

Rotated stereo-camera rig and a Kinect. [6, p. 5, fig. 1.2]

Stereoscopic Cameras
Problems

- identification of corresponding points in both images
- occlusion
- computationally expensive
- depends on illumination
- cameras need to be synchronized



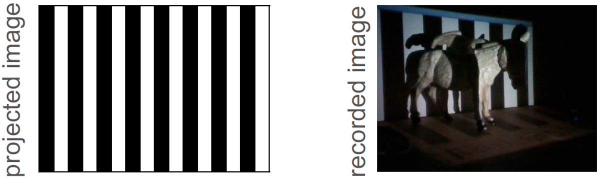
Stereo-Camera example. [5, p. 38, fig. 2.1]

Structured-Light

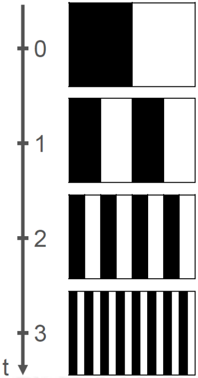
- project additional information on the object to allow recovery of lost depth dimension
- several different approaches
  - time multiplexing
  - spatial multiplexing
  - wavelength multiplexing

Binary Projection

- one camera, one projector
- several passes required
- deformity of lines as measure for depth
- time multiplexing
- active approach



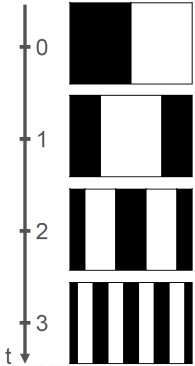
Binary projection. [7, p. 30, fig. 1]



Binary projection at different times t. [7, p. 33, fig. 1]

Binary Projection
Problems

- frames taken at different points in time
  - time multiplexing
  - not applicable for moving objects
- points directly on edges are uncertain
  - soultion: gray code pattern



Gray code projection at different times t. [7, p. 33, fig. 1]

Microsoft Kinect

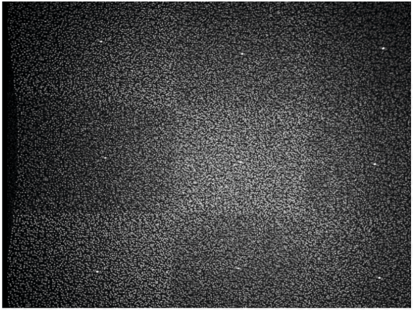
- RGB camera: 30fps @ 640x480px
- spatial multiplexing
- USB 2.0

- depth image: 30fps @ 320x240px
- practical range [4]
  - 0.8-4.5m in default mode
  - 0.4-3m in near mode

Microsoft Kinect. [3, p. 2, fig. 1-1]

Microsoft Kinect
IR Laser Emitter

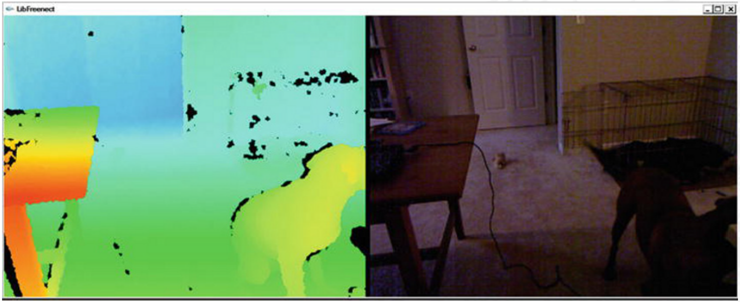
- projection
  - pseudo random noise-like pattern
  - 830nm wavelength
- laser
  - heated/cooled to maintain wavelength
  - 70mW output power
  - eye safety through scattering



Projected IR pattern. [3, p. 12, fig. 2-2]

Microsoft Kinect
Depth Image

- IR camera image compared to known pattern
  - disturbances can be used to calculate distances
- distances visualized as depth images
  - red areas: close
  - blue areas: further away
  - black areas: no depth information available



Depth image and corresponding RGB image. [3, p. 9, fig. 1-3]

- overexposure of IR camera
  - by sunlight (only usable indoors)
  - by reflecting surfaces
- only close range distances
  - limited by laser output
- translucent objects not measurable
- latency of ~100ms [4]
- active approach, not easy to scale-out
  - interferences with projected patterns

Triangulation Approaches Conclusion

Stereo Cameras

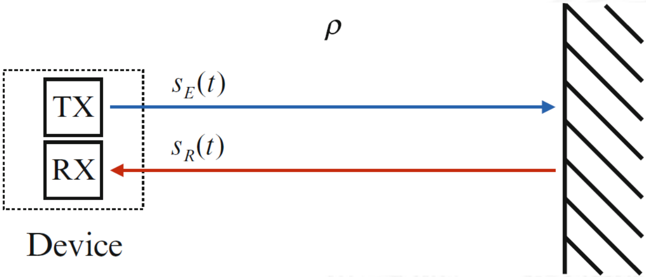
- good to calculate depths for distinct markers
  - otherwise computationally expensive
- works indoors and outdoors
- completely passive, scaling out is possible without problems

Triangulation Approaches Conclusion
Structured-Light Cameras

- all approaches
  - trouble measuring reflecting or transparent objects
- time multiplexing
  - depth calculation for whole field of vision
  - only for stationary objects
- spatial multiplexing (Kinect)
  - computation done by hardware
  - pretty complete depth map
    - occluded areas
    - too close or too far points
- wavelength multiplexing
  - depth calculation with one photo
  - low spatial resolution achievable

Time of Flight Approaches

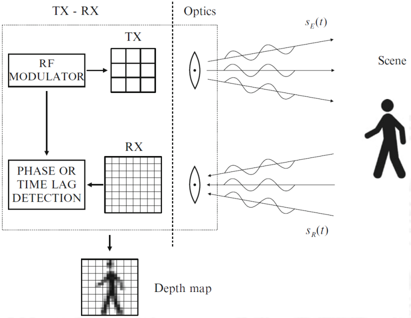
- actively send out a signal
- measure time until reflection returns
- Light: $P = \frac{299.792.458 \frac{m}{s} * t}{2}$



Simple ToF measurement. [8, p. 28, fig. 1.14]

Depth Camera

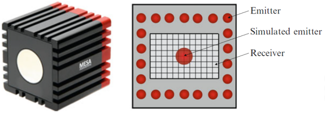
- active approach
- TX: illuminates whole scene with array of IR emitters
- RX: ToF-receiver grid
- commonly used: sinus modulation for emitted light
- measure point in time when emitted signal returns
- calculate distance through ToF



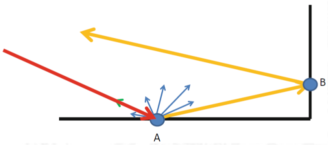
MESA Imaging SR4000, IR emitters. [8, p. 32, fig. 1.16]

Depth Camera
Problems

- hardware restrictions
  - IR-emitter and ToF-receievers in different position
  - simulate central emitter to avoid occlusion effects
- falsification of measurements through multi path hopping
  - point B will measure a combination of two distances
- accurate time measurement required



Pattern of IR emitters to avoid occlusion. [8, p. 34, fig. 1.17]



Multipath phenomenon. [8, p. 104, fig. 3.16]

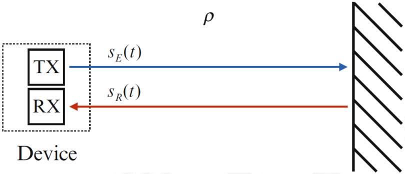
Kinect V2

- depth image: 50fps @ 512x424px
- range 0.5-8m [4]
- latency of ~50ms [4]
- square wave modulation
- differential pixel array
  - switches with square wave
  - save returned light
  - difference used to compute distances
- high volume of data, requires USB 3.0

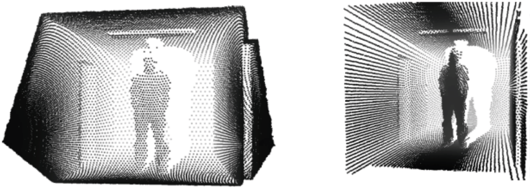
Kinect V2. [4, p. 6, fig. 1-5]

LIDAR

- **L**i**ght** **D**etection **A**nd **R**anging
- sends out single laser beam
- ToF to calculate distance
- single point sampling
  - mirrors rotate laser beam to scan line of points
  - additional rotation possible to scan area instead of line



Simple ToF measurement. [8, p. 28, fig. 1.14]



Point clouds created by rotated line scanners. [2, p. 46, fig. 2.21]

LIDAR
Problems

- loss of spatial resolution with increased measurement distance
- transparent objects can not be measured
- mechanical moving parts

- high laser outputs possible
  - high measurement range
  - sunlight can be compensated
- high sampling rates possible
- dynamic measurement range
  - short and long distances can be measured together

Conclusion
Required Ambient Lighting

- structured light approaches require dark surroundings
  - often used for optical measurements and inspection in industrial robotics
  - very precise measurements
- LIDAR can be built for outdoor usage
- other active approaches falsified/annulled by direct sunlight

Conclusion

Computational Costs

- active approaches
  - distance calculation mostly handled by hardware
- stereoscopic cameras
  - expensive: calculate matching points in both images

Conclusion
Moving Objects

- depth cameras, spatial multiplexing structured light
  - well suited
  - record whole scene at single point in time
- binary projection
  - not usable, time encoding through different frames
- LIDAR
  - suitability depends on sampling rate and object movement

	Triangulation			Time-of-Flight	
	Stereo Camera	Binary Projection	Kinect	Kinect V2	LIDAR
outdoor usability	✓	✗	✗	✗	(✓)
complete depth map	(✗)	✓	✓	✓	(✓)
passive	✓	✗	✗	✗	✗
scale out	✓	(✗)	(✗)	(✓)	(✓)
moving parts	✗	✗	✗	✗	✓
“cheap”	(✓)	✗	✓	✓	✗

[1] Prof. Dr. Leonie Dreschler-Fischer. Interactive visual computing lecture ws2012. Handouts 2012.

[2] Joachim Hertzberg, Kai Lingemann, and Andreas Nüchter. *Mobile Roboter*. Springer, Berlin, 2009. ISBN 978-3-642-01726-1.

[3] Jeff Kramer, Matt Parker, Daniel Herrera C., Nicolas Burrus, and Florian Echtler. *Hacking the Kinect*. Apress, New York, 2012. ISBN 978-1-4302-3868-3.

[4] Mansib Rahman. *Beginning Microsoft Kinect for Windows SDK 2.0*. Apress, Berkeley, CA, 2017. ISBN 978-1-4842-2316-1. Motion and Depth Sensing for Natural User Interfaces.

[5] Stephen Se and Nick Pears. *Passive 3D Imaging*, pages 35–94. Springer London, London, 2012. ISBN 978-1-4471-4063-4. DOI:10.1007/978-1-4471-4063-4_2. URL https://doi.org/10.1007/978-1-4471-4063-4_2.

[6] Jan Smisek, Michal Jancosek, and Tomas Pajdla. *3D with Kinect*, pages 3–25. Springer London, London, 2013. ISBN 978-1-4471-4640-7. DOI:10.1007/978-1-4471-4640-7_1. URL https://doi.org/10.1007/978-1-4471-4640-7_1.

[7] Prof. Dr. Frank Steinicke. User interface & software & technology lecture ws2014. Chapter 5: 3D Capturing Systems (Slides by Susanne Schmidt), 2014.

[8] Pietro Zanuttigh, Giulio Marin, Carlo Dal Mutto, Fabio Dominio, Ludovico Minto, and Guido Maria Cortelazzo. *Time-of-flight and structured light depth cameras : technology and applications*. Springer, Switzerland, 2016. ISBN 978-3-319-30973-6.