3D Sensing and Sensor Fusion
http://cg.elte.hu/~sensing

Dmitry Chetverikov, Iván Eichhardt
csetverikov@sztaki.hu, eiiraai@inf.elte.hu

Eötvös Loránd University
Faculty of Informatics

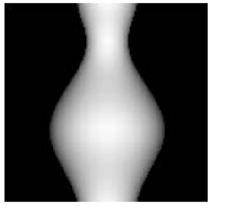
2021 fall (2021-2022-1)

Introduction

Techniques to acquire depth

- Stereo vision
  - calibrated stereo rig
  - multiview
- Shape/structure from X
  - shading, texture, motion
- Photometric stereo
  - multiple light sources
- Laser scanner and structured light
  - light pattern projected onto object
- Direct depth measurement
  - Lidar devices: 3D point clouds
  - Time-of-Flight (ToF) cameras: depth images

Shape from shading



- Shape of object clearly visible from **shading variation**[1]
- Surface normal changes across object
  - → brightness changes as function of angle between surface normal and direction of illumination
- **Shape from shading**: Recover surface from brightness variation

[1] Source: R.Szeliski, Computer Vision

Solution for shape from shading 1/2

- Basic **assumption**: distant light source and observer
  → brightness I is function of surface normal called **reflectance map**
- For diffuse (Lambertian) surface, reflectance map is

$$I(x,y) = R\left(p(x,y), q(x,y)\right) = \max\left(0, \rho \frac{pv_x + qv_y + v_z}{\sqrt{1+p^2+q^2}}\right)$$

  - dot product of surface normal $\mathbf{n} = (p,q,1)/\sqrt{1+p^2+q^2}$ and light source direction $\mathbf{v} = (v_x, v_y, v_z)$
  - ρ: surface reflectance factor (albedo)
  - $(p,q) = (z_x, z_y)$: derivatives of depth map $z(x,y)$

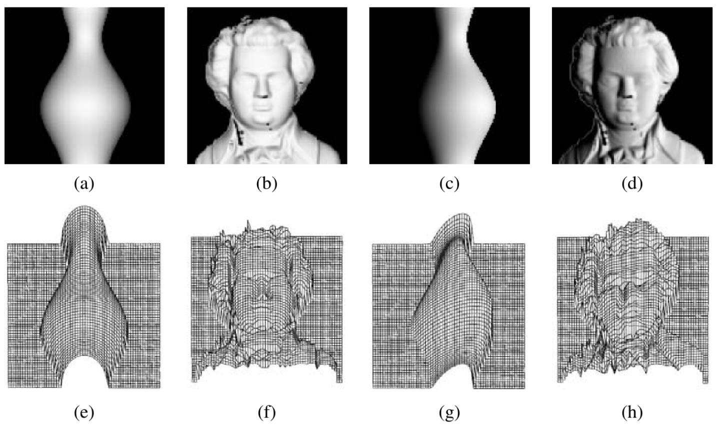
Solution for shape from shading 2/2

- Another **assumption**: Light source direction either **known or can be calibrated**
- More unknowns per pixel (p,q) than measurements (I)
  → additional **constraints** needed
- Possible constraints:
  - **smoothness**
  - **integrability**

→ Cost function minimized by **variational technique**

- Recover orientation field $(p,q) = (z_x, z_y)$
- Integrate orientation field to obtain surface

Example of shape from shading



Synthetic shape from shading. (a-b) front light; (c-d) front right light; (e-f) corresponding reconstructions by shape from shading

Summary of shape from shading

- For smooth, non-textured surfaces with uniform albedo
- Can provide fine details of surface
- Disadvantages:
  - not applicable to textured surfaces
  - assumption of uniform albedo is often not realistic
  - ambiguity (multiple solutions) possible
  - search for cost minimum can end in local minimum
  → less robust
- To cope with multiple albedos and texture, combine (fuse) with other techniques
  - stereo
  - texture-based

Photometric stereo

- Make shape from shading more reliable
- Use multiple light sources selectively turned on/off
  - light sources play role of cameras in standard stereo
  - their directions assumed to be **calibrated and known**
- For each light source, different reflectance map and brightness function
- For diffuse surface, obtain set of linear equations as in shape from shading
  $\rightarrow$ recover surface normal by linear least squares
- Integrate orientations to obtain surface

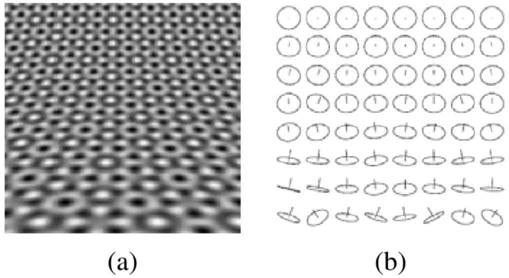
Properties of photometric stereo

- More robust than shape from shading
  - but multiple solutions still possible
- Good normal vectors
- Fine surface details
  - good local geometry
- Less precise positions
  - global geometry not that good
- Combine with methods providing good global geometry

Shape from texture

- Texture variation on smooth, well-textured surface
- Extract repeated patterns or measure local frequences
- Obtain local affine deformations
- Local affine deformation $\longrightarrow$ local surface orientation
- Integrate orientations to obtain surface

Example of shape from texture



Synthetic shape from texture. (a) regular texture wrapped onto curved surface; (b) corresponding surface normal estimates

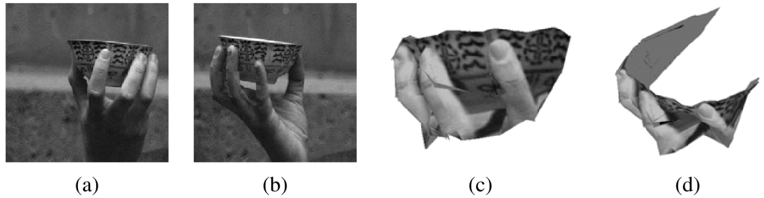
Disadvantages of shape from texture:

- Not applicable to textureless surfaces
- Less robust

Structure from Motion (SfM)

- Relative motion of scene, or object, and camera
- Simultaneous estimation of
  - 3D geometry (structure) and
  - camera pose (motion)
- For different camera models
- For rigid and non-rigid objects
- Based on feature tracking $\longrightarrow$ **sparse 3D model**
- Uses **factorization** of measurement matrix into
  - motion matrix and
  - structure matrix

Example of reconstruction by SfM



3D teacup model reconstructed from video sequence. (a) first frame; (b) last frame; (c) side view of 3D model; (d) top view of 3D model

Summary of Structure from Motion

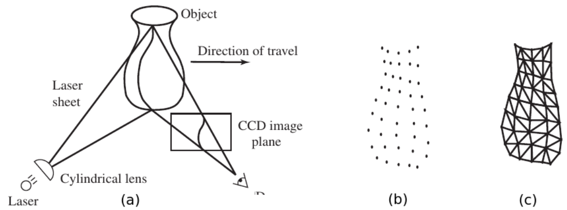
Used for **simultaneous localization and mapping** (SLAM)

- In robotics navigation
- Estimate robot position relative to environment
- Build sparse 3D model of environment to avoid obstacles

Limitations of Structure from Motion:

- Not applicable to textureless surfaces
- Often, sparse 3D model only
  - in principle, dense optical flow can also be used
- Sensitive to missing data and outliers

Range data scanning



(a) laser stripe (sheet) imaged by camera; (b) resulting 3D point set; (c) triangulated mesh

- Laser light stripe sweeps plane of light across object
  - curve on surface is observed from known viewpoint
  - stripe deforms according to object shape
  - **optical triangulation** used to estimate 3D locations
  - **calibrated setup**: 3D plane & camera pose known

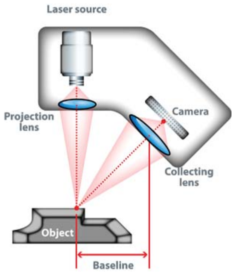
Advantages of laser scanners

- Devices available for different applications
  - range of object size
  - type of surface
  - precision
  - price
- Reliable, stable
- Quite accurate
- Used in everyday practice
  - → many years of experience
  - → development and improvement

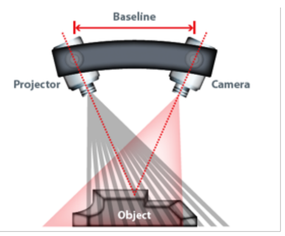
Limitations of laser scanners

- Application often limited to indoor conditions
  - needs controlled illumination conditions
  - outdoor light can interfere measurement
- Limited range and object size
  - long-range scanners for large objects very expensive
- Relatively slow scanning
  - faster operation using structured light patterns
- Not applicable to specular or transparent surfaces
  - applicability to strongly textured surfaces also limited

Laser scanner vs structured light



laser scanner

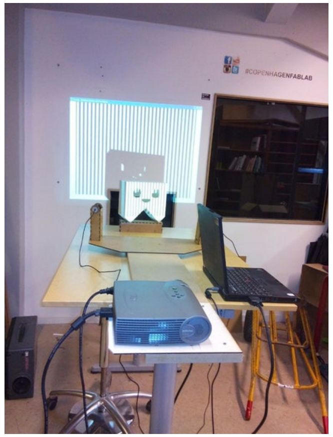


structured light[2]

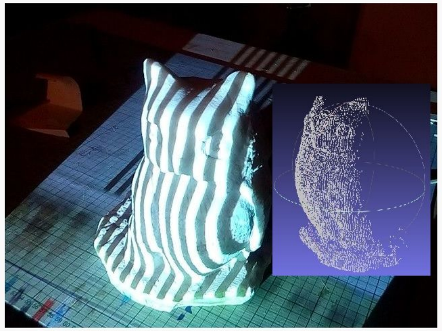
- Laser light **stripe** scans surface of object
- Projector casts structured light **pattern** covering object
  - different patterns used to texture object surface
  - → parallel stripes, checkerboard grid, random encoding, etc.
  - applies optical triangulation to pattern

[2]Source: http://3dscanningservices.net

Example of 3D object reconstruction by structured light



projecting light pattern



object and its 3D reconstruction

Laser scanner: pros and cons

Pros:

- Available in many forms:
  - area scanner, handheld, portable arm
- Often, more portable
- Less sensitive to surface finish (condition)
- Less sensitive to ambient light

Cons:

- Generally, less accurate
- Generally, lower resolution
- Higher noise

Structured light: pros and cons

Pros

- Usually, more accurate
- Often, higher resolution
- Lower noise

Cons

- Limited to area scanner type
- Generally, not as small/portable
- More sensitive to surface finish
- May require specific lighting

Lidar devices

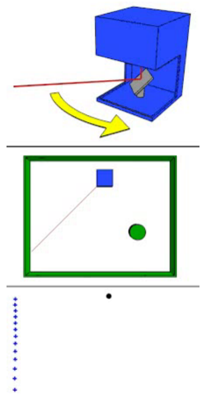
short-range Lidar

long-range lidar

- LIght Detection And Ranging[3]
- Measure distance to target
  - scan and illuminate target with laser light
  - measure reflected light with sensor
  - → differences in laser return times and wavelengths

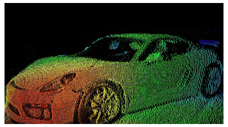
[3] Source: Wikipedia

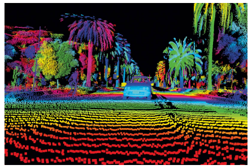
Illustration of Lidar scanning

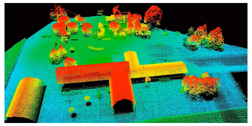


- Top: Basic lidar system has laser range finder reflected by rotating mirror
- Middle: Laser is scanned around scene being digitized, in 1D or 2D
- Bottom: Gathering distance measurements at specified angle intervals

Examples of Lidar outputs and applications


short-range


long-range


aerial

- High-resolution 3D **point clouds**
  - in illustrations, coloured for better visibility
- Short range
  - mobile robotics, (autonomous) cars, transport
- Long range
  - 3D models of buildings, streets, squares
- Aerial and spaceflight
  - surveing large territories, spacecrafts, atmospheric studies

Advantages of Lidar devices

- Lidar can use different types of light to image objects
  - ultraviolet, visible, or near infrared
- Can target wide range of materials
  - non-metallic objects, rocks, rain, vegetation
  - chemical compounds, aerosols, clouds
- Can operate at different ranges and lighting conditions
- Numerous applications

Disdvantages of Lidar devices 1/2

- Contains moving parts
  - → sensitive to outdoor conditions
  - → may need frequent re-calibration
- Resolutions in vertical and horizontal directions differ
  - vertical angular resolution is significantly lower
- May yield point clouds with varying point density
  - when size of object is comparable to distance to object
  - → problem for point cloud registration

Disdvantages of Lidar devices 2/2

- 3D points are obtained by scanning
  - → in different time
  - → distortion when sensor or scene are moving
- For **static scene**, distortion can be compensated
  - if sensor motion is known
- For **dynamic scene**, more difficult or even impossbile
- Potential solution: **Solid State Lidar**
  - based on silicon chip → no moving parts
  - sends out bursts of photons in specific patterns and phases → create directional emission
  - **under development and testing**

Operation of ToF cameras

- We only consider widely used **lock-in** ToF cameras
- Principle of operation
  - emit infrared light
  - measure time-of-flight to observed object
- ToF camera delivers
  - depth images at video frame rates
  - registered reflectance images of same size
  - reliability values of depth measurements

Features of ToF cameras

- Main features
  - small, compact
  - low weight
  - low consumption
  - no moving parts (contrary to Lidar)
- Most exploited features
  - ability to operate without moving parts
  - providing depth maps at high frame rates
  - → greatly simplifies foreground-background separation

Disadvantages of ToF cameras

- Low resolution
  - highest current resolution: QVGA (320×240)
  - target of future development: VGA (640×480)
  - → limited by chip size
- Significant acquisition noise
  - low quality depth images
  - quality gradually improving
  - → limited by small active illumination energy
- Problems with outdoor use
  - bright lighting can increase ambient light noise in ToF
  - → if ambient light contains same wavelength as camera light

Applications of ToF cameras

- Mobile robot vision
  - navigation
  - 3D pose estimation and mapping
- 3D reconstruction of objects and environments
- Computer graphics and 3D television (3DTV)
- Recognition and tracking of people and parts of body
  - face, hands

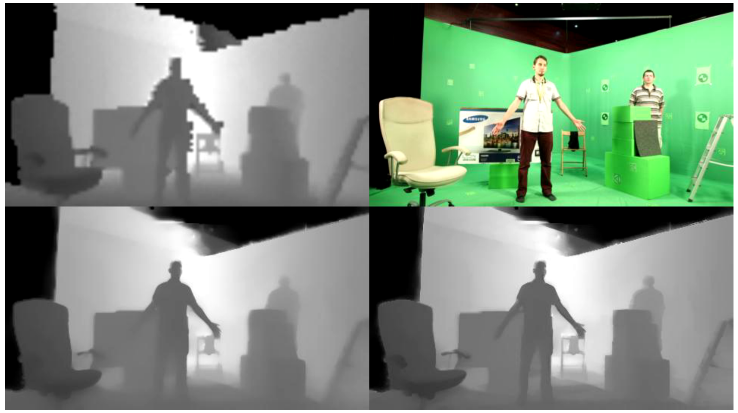
Example of depth image captured in studio

input data

resized depth

- Part of data lost due to very low resolution
- Some shapes, e.g., heads and hands, distorted
- Contours of depth and colour images may not coincide
- Erroneous and missing measurements
  - especially, at depth edges

Video of depth upsampling in studio



Upsampling algorithms with optical flow.

Comparison of four depth acqusition techniques

feature	stereo	phot. stereo	struct. light	ToF camera
correspond.	yes	no	yes	no
extrinsic calib.	yes	yes	yes	no
active illum.	no	yes	yes	yes
weak texture	weak	good	good	good
strong texture	good	medium	medium	medium
low light	weak	good	good	good
bright light	good	weak	med./weak	medium
outdoor	yes	no	no	yes?
dynamic	yes	no	yes	yes
resolution	cam.dep.	cam.dep.	cam.dep.	low
accuracy	mm to cm	mm	μm to cm	mm to cm