Structured light 3D scanning

Morten R. Hannemose, mohan@dtu.dk

May 3, 2024

02504 Computer vision course lectures,
DTU Compute, Kgs. Lyngby 2800, Denmark

This lecture is being livestreamed and recorded

(hopefully)

Two feedback persons

Learning objectives

After this lecture you should be able to:

- explain laser line scanning
- analyse and use Gray code encoding
- analyse and use phase shift encoding

Presentation topics

- Photogrammetry
- Structured light
  - Laser line scanning
  - Encoding surfaces
  - Gray code encoding
  - Phase shift encoding
- Notes on laser/projector calibration
- Exercise: structured light
- Exam information

Photogrammetry

Photogrammetry

Take a lot of pictures of a scene and use SLAM to find camera positions and 3D points. Algorithms for dense estimation of 3D points can be applied to get a full 3D scene.

Photogrammetry

Great:

- Works in daylight
- Handles textured objects

Bad:

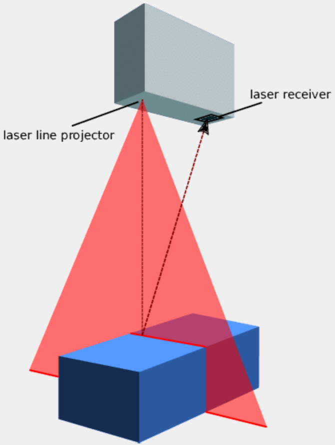
- Requires good illumination
- Requires textured objects

Structured light

Laser line scanning

The first structured light technique.

- Laser projects 3D plane of light
- Camera sees the projected line
- Laser projector and camera are calibrated



Laser line scanning

Example laser line scanner triangulation:

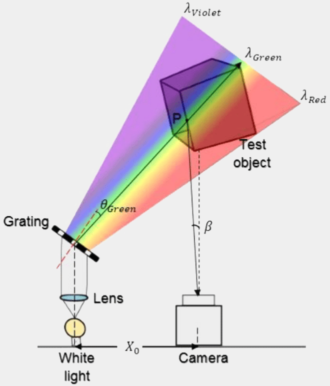
1. detect laser line in image
2. triangulate 3D point as intersection of pixel and the laser plane
3. move laser line and goto 1

Robust method, with a few drawbacks

- Requires laser calibration.
- A slow method; one triangulation line per image

Encoding surfaces

Can we encode a surface?



Yes we can!

Encoding surfaces

Possibilities:

- Continuous encoding
  - Color or intensity gradient
  - Sinusoidal (phase) shifting
- Discrete encoding
  - Binary monochrome
  - ternary RGB encoding
  - quaternary CMYK encoding
- Other encoding schemes

Encoding surfaces

- Colored encodings have problems with colored surfaces
- Intensity codings have problems with textured objects

- Continuous encoding
  - **Sinusoidal (phase) shifting**
- Discrete encoding
  - **Binary monochrome**

Continuous encoding schemes

1. For each pixel in the camera, identify the code/color
2. For each code, identify the corresponding light plane
3. Triangulate using pixel rays and the “laser plane”

Can get a 3D point for each pixel in the camera

Not always robust.

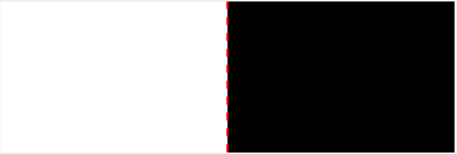
Discrete encoding schemes

1. For each pixel in the camera, identify the code/color
2. For each code border, identify the corresponding plane
3. Triangulate using pixel rays and the “laser plane”

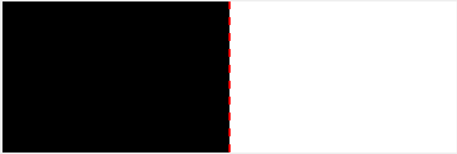
Only 3D points at code-borders, but usually more robust.

Binary encoding – single frame

Frame:



Inverted frame:



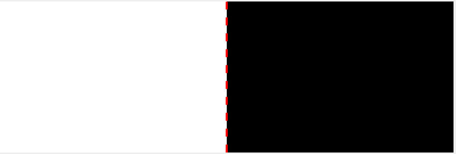
Frame + inverted frame makes us robust towards ambient light and varying object albedo.

Binary test: $\tau(\boldsymbol{p}, \boldsymbol{p}_i) = 1$ if $\boldsymbol{p} > \boldsymbol{p}_i$ else 0.

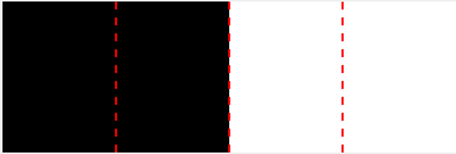
Single test; two regions; one border

Binary encoding – two frames

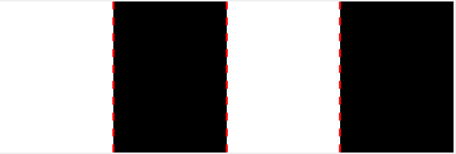
Frame 1:



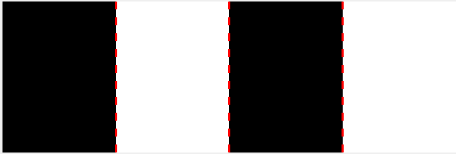
Inverted frame 1:



Frame 2:



Inverted frame 2:



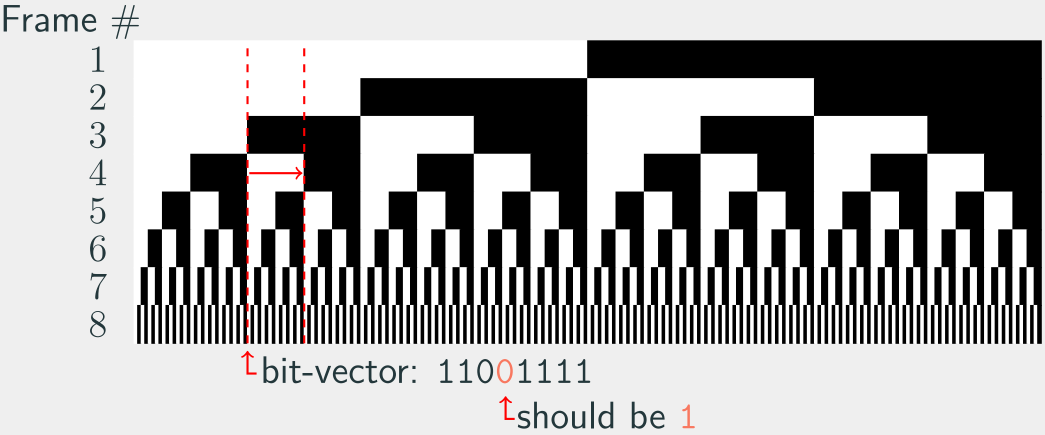
Two tests; four regions; three borders.

Binary encoding – multiple frames

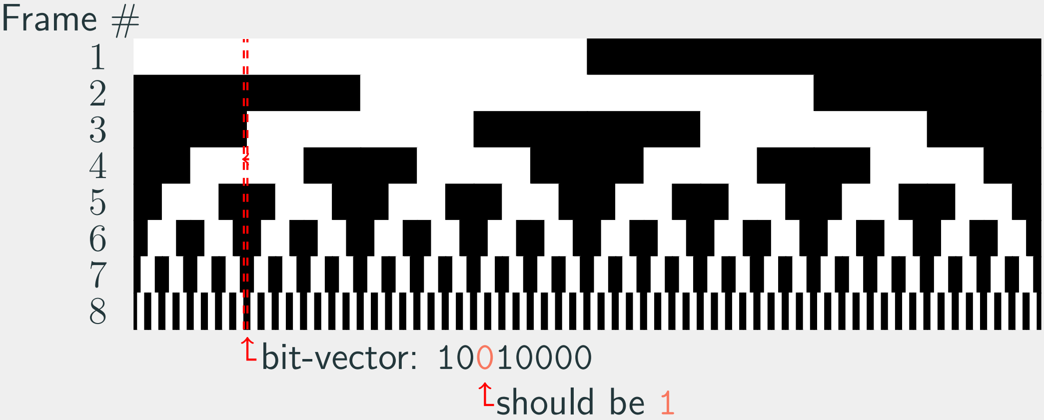
Binary encoding

- For every frame we subdivide the 3D volume.
- For N frames we get 2^N unique regions.
- For N frames we get $2^N - 1$ unique borders.
- If a projector is $W = 1920$ pixels wide, we are limited to $N \leq \log_2(W) \approx 10.9$. That is a max of 10 frames in total or, 20 in total with inverted frames.

Binary encoding – border problems



Gray code encoding

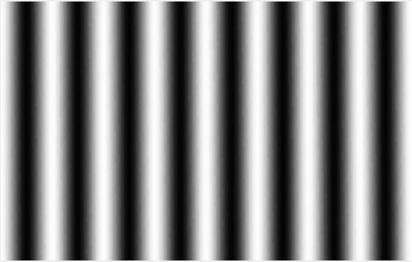


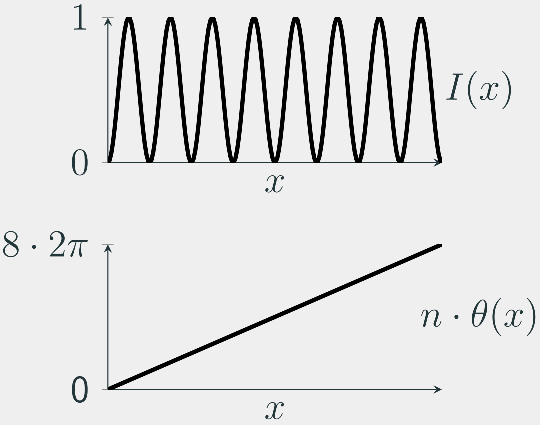
Gray codes

- Only one bit flip at code borders; very robust.
- Uses same number of frames as binary patterns.

Phase shift encoding

Sinusoidal waves



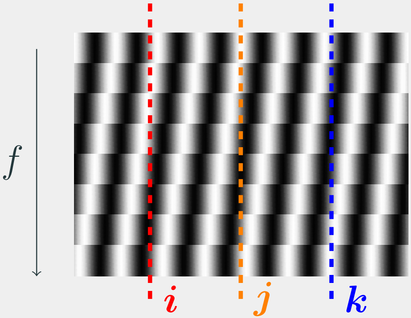


$$I(x, y) = \frac{1}{2} + \frac{1}{2}\cos(n \cdot \theta(x))$$

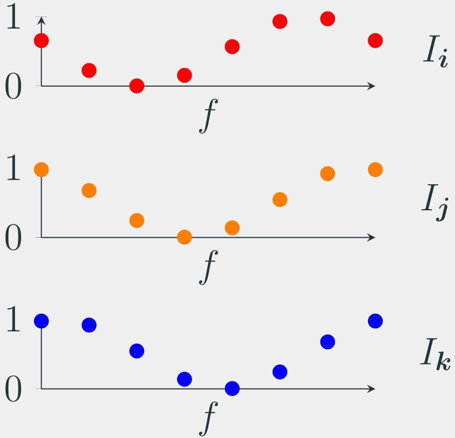
Phase as a unique code

- The pattern is monochrome; good for colored objects.
- The phase is continuous; each point in space has a unique phase-plane.
- Even with a discrete intensity projector, the pattern is approximately continuous.
- $\theta(\cdot)$ is a function of the x-coordinate of the projector
  - From now on θ is the x-coordinate of the projector (from 0 to 2π)

Phase shifting



$$I(f,\theta) = \frac{1}{2} + \frac{1}{2}\cos\left(n\cdot\theta + 2\pi\frac{f}{s}\right)$$



When projecting sinusoids and shifting them, each pixel is a sinusoid as a function of the pattern

Phase shifting method

- Phase shift exactly one wavelength in s steps.
- The phase θ corresponds to a unique projector plane.
- We can find $n \cdot \theta$ for a single pixel by fitting a sinusoid to it

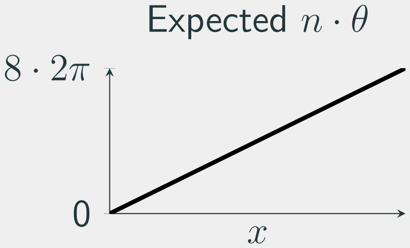
Phase shifting method

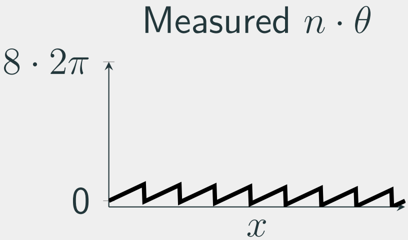
- Phase shift exactly one wavelength in s steps.
- The phase θ corresponds to a unique projector plane.
- We can find $n \cdot \theta$ for a single pixel by fitting a sinusoid to it
- This can be done with least squares (slow) or the fast Fourier transform (FFT) (fast).

$$\mathrm{FFT}_f[I(f,\theta)] = \left\{\frac{s}{2}, \frac{s}{2}e^{i\theta}, 0, \dots, 0\right\}$$

The second element of the FFT is a complex number with $\theta = \mathrm{angle}(\mathrm{FFT}_2)$.

Phase wrapping





The phase $n \cdot \theta = \text{angle}(\text{FFT}_2)$ is wrapped to $0 \leq n \cdot \theta \leq 2\pi$.

We need to unwrap the recovered phase!

Heterodyne principle

Use two sinusoids:

$$\begin{aligned} \theta_1 =& n_1 \cdot \theta \mod 2\pi \quad \text{(primary pattern)} \\ \theta_2 =& n_2 \cdot \theta \mod 2\pi \quad \text{(secondary pattern)} \end{aligned}$$

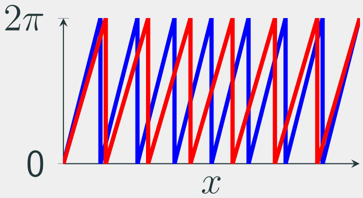
If we choose $n_2 = n_1 + 1$ we can recover θ by subtracting the secondary phase from the primary phase

$$\theta_2 - \theta_1 = n_2 \cdot \theta - n_1 \cdot \theta = \theta \mod 2\pi$$

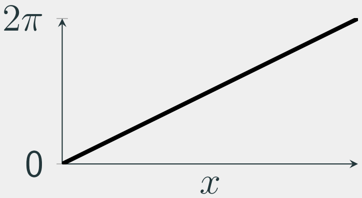
This is the heterodyne principle, and gives us the phase cue (θ_c).

Heterodyne principle

Measured θ_1 and θ_2



$\theta_c = \text{mod}\big(\theta_2(x) - \theta_1(x), 2\pi\big)$



Looks great! 

Questions? Short break

Unwrapping

- If the measurements of θ_1 and θ_2 are noise free, the phase cue θ_c is exactly equal to θ.
- All measurements have noise, but in this case we can improve the handling of noise substantially.

Unwrapping

To make the system more robust to noise in the measurements we can compute θ using the phase cue and the primary phase θ_1.

The order counts how many times θ_1 has wrapped around

$$o_1 = \left\lfloor \frac{n_1 \cdot \theta_c - \theta_1}{2\pi} \right\rceil$$

The rounding $\lfloor \cdot \rceil$ makes it robust to noise. We can now estimate θ

$$\theta_{\text{est}} = \frac{2\pi o_1 + \theta_1}{n_1} \mod 2\pi$$

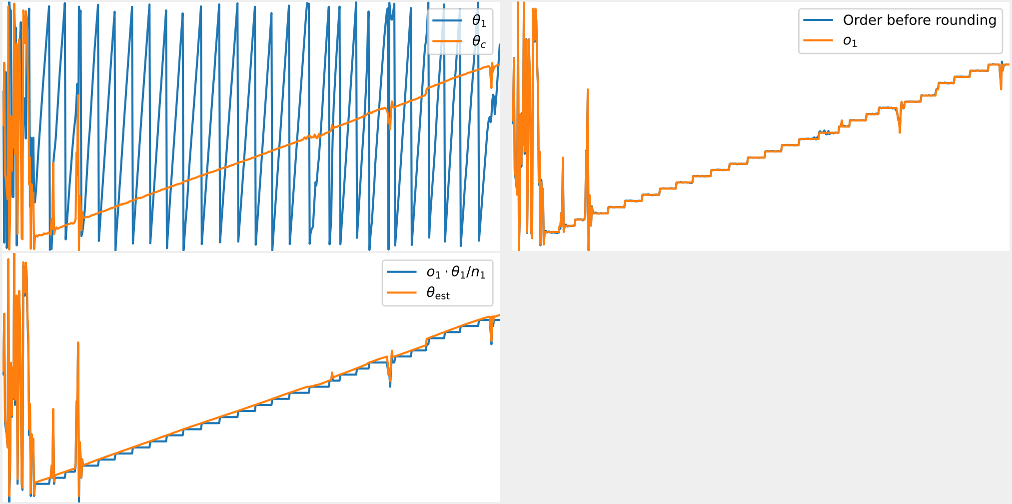
Unwrapping and phase cue – Motivation

- Why not just use the phase cue?
- $\theta_1 = \theta_1^{\text{true}} + \epsilon_1$
- $\theta_2 = \theta_2^{\text{true}} + \epsilon_2$
- $\theta_c = \theta^{\text{true}} - \epsilon_1 + \epsilon_2$

Unwrapping and phase cue – Motivation

- Why not just use the phase cue?
- $\theta_1 = \theta_1^{\text{true}} + \epsilon_1$
- $\theta_2 = \theta_2^{\text{true}} + \epsilon_2$
- $\theta_c = \theta^{\text{true}} - \epsilon_1 + \epsilon_2$
- The error of the phase cue is linear in the errors of θ_1 and θ_2
- θ has an error of $\frac{\epsilon_1}{n_1}$, which is much lower

Unwrapping – examples from today's exercise.



Examples from row 400 of camera 0.

Phase shift encoding

1. Project two sets of sinusoidals.
2. For each pixel, find the wrapped phases θ_1 and θ_2.
3. Calculate the phase cue $\theta_c = \theta_2 - \theta_1 \mod 2\pi$.
4. Calculate the primary order $o_1 = \left\lfloor \frac{n_1 \cdot \theta_c - \theta_1}{2\pi} \right\rceil$
5. Calculate the theta $\theta_{\text{est}} = \frac{2\pi o_1 + \theta_1}{n_1}$
6. For each phase, find the corresponding projector plane.
7. Triangulate pixel rays and the projector plane.

Phase shift encoding

- Very robust since the FFT is also a low-pass filter.
- Potentially one triangulation per camera pixel.
- Easy to make more precise; more projected frames gives a higher precision.

Notes on laser/projector calibration

Laser/projector calibration

- Is needed for many structured light methods.
- Requires correspondences between the laser/projector and the real world.
- Often we use a calibrated camera to relate projected lines/pixels to world coordinates.
- Not ideal for projectors, as they usually have poor quality lenses.

However, do we need to calibrate the laser/projector?

“Passive” structured light

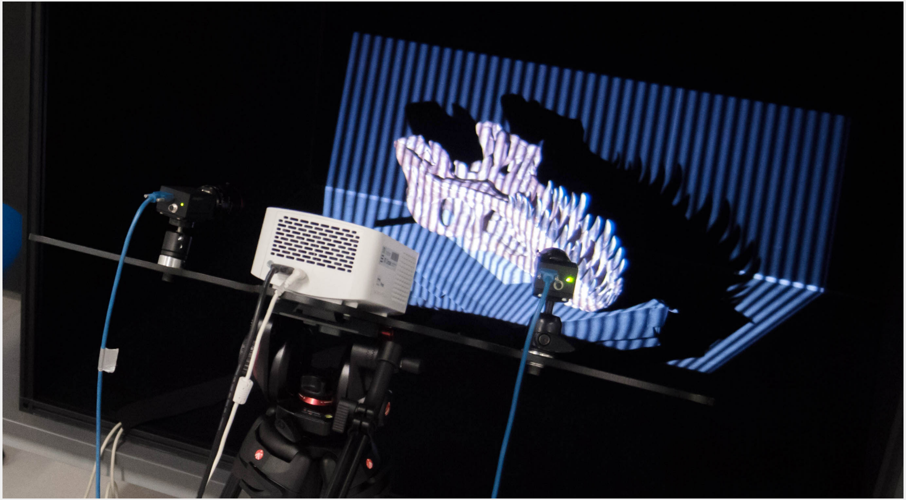
- Use a stereo camera set up and a laser/projector.
- The laser/projector encodes the object surface, but is not calibrated.
  - Accurately calibrating a projector is harder than calibrating a camera
- Instead of laser/projector plane triangulation, we use epipolar lines in the cameras.
- Requires that the projected planes are not parallel with the epipolar planes.

Exercise: structured light

Structured light exercise

- Scan using phase shifting
- You will perform:
  - Image rectification
  - Phase decoding (and unwrapping)
  - Masking
  - Matching
  - Triangulation

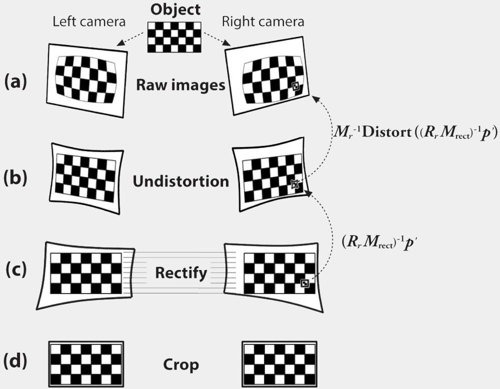
Casper the baby T-Rex



Casper the baby T-Rex

Rectified cameras

Cameras are (virtually) made parallel to each other.



Masking

We need to only match phases, which is efficient and robust.

With rectified images, epipolar lines are the corresponding rows.

Live demo[ish]

Further courses

- Are you interested in working more with images?
- 02506 Advanced Image Analysis (Spring)
- 02516 Introduction to Deep Learning in Computer Vision (January, moving to Fall)
- 02501 Advanced Deep Learning in Computer Vision (Spring)

Learning objectives

After this lecture you should be able to:

- explain laser line scanning
- analyse and use Gray code encoding
- analyse and use phase shift encoding

Exam information

Exam information

Exam time and place: `www.eksamensplan.dtu.dk`

- Old exam has been uploaded (same as quizzes)
- Multiple choice exam
  - No negative points for wrong answers
- Questions are a mix of understanding and calculation in Python.
  - No internet access during exam
  - You can download the documentation for OpenCV to your computer.

Quiz awards

Exercise time!