Three-Dimensional Sensors
Lecture 3: Time of Flight Cameras
(Continuous Wave Modulation)

Radu Horaud
INRIA Grenoble Rhone-Alpes, France
Radu.Horaud@inria.fr
http://perception.inrialpes.fr/

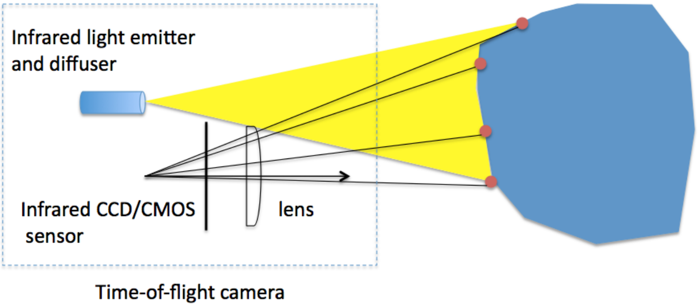
Lecture Outline

- The principle of TOF sensors
- Estimation of phase, amplitude and offset
- Depth estimation
- Camera calibration

Time of Flight Sensors

- Basic principle: measure the absolute time that a light pulse needs to travel from a light emitter to a targeted object and back to a detector.
- The speed of light is of 0.3 meters/nano-second.
- TOF systems use either *pulsed-modulation* or *continuous wave modulation*:
  - **Pulsed-modulation** measures the time-of-flight directly. Allows long-distance measurements. The arrival time must be detected very precisely. Needs very short light pulses with fast rise- and fall-times and with high optical power – lasers or laser diodes.
  - **CW-modulation** measures the phase difference between the sent and received signals. Different shapes of signals are possible, e.g., sinusoidal, square waves, etc. Cross-correlation between the received and sent signals allows phase estimation which is directly related to distance if the modulation frequency is known.

Principle of a TOF Camera



Depth Estimation

- The distance is computed with:

$$d = \frac{1}{2} c\tau$$

- In practice τ cannot be measured directly.
- Continuous wave modulation: The phase difference between the sent and received signals is measured.
- The modulated frequency is in the range 10 to 100 MHz.
- Relationship between phase shift and time of flight (f is the modulation frequency):

$$\phi = 2\pi f\tau$$

Cross-Correlation Between Sent and Received Signals

- Emitted signal: $s(t) = a\cos(2\pi f t)$
- Received signal: $r(t) = A\cos(2\pi f(t-\tau)) + B$
- Cross-correlation between emitted and received signals:

$$C(x) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} r(t)s(t+x)dt$$

- The solution is:

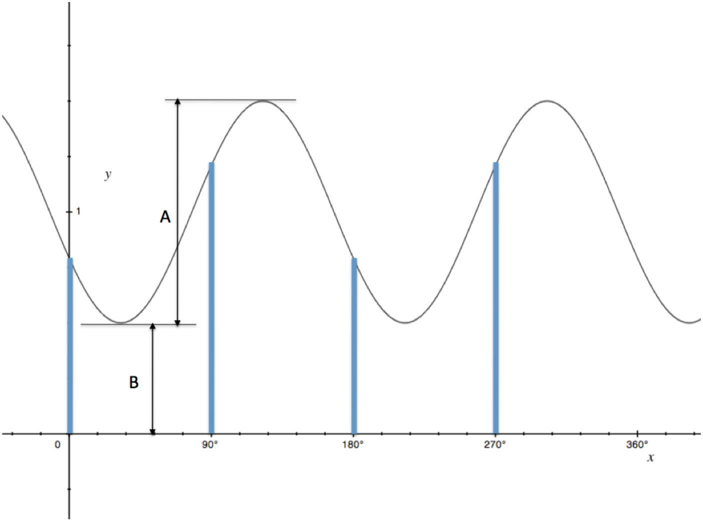
$$C(\psi) = \frac{aA}{2}\cos(\underbrace{2\pi f\tau}_{\phi} + \underbrace{2\pi f x}_{\psi}) + B$$

Demodulation Parameters

- ϕ is the phase and is defined up to 2π, this is called phase wrapping.
- A is the amplitude of the received signal and it depends on the object's reflectivity and of the sensor's sensitivity. The amplitude decreases with $1/d^2$ mainly due to light spreading.
- B is an offset coefficient due to the ambient illumination.

The 4-Bucket Method

- Estimate four values of $C(\psi)$ at $\psi_0 = 0^0$, $\psi_1 = 90^0$, $\psi_2 = 180^0$, $\psi_3 = 270^0$



Phase, Amplitude, and Offset

- From these four values of the cross-correlation signal we obtain the following solutions for the phase, amplitude and offset:

$$
\begin{aligned}
\phi &= 2\pi f \tau = \arctan\left(\frac{C(x_3) - C(x_1)}{C(x_0) - C(x_1)}\right) \\
A &= \frac{1}{2a}\sqrt{(C(x_3) - C(x_1))^2 + (C(x_0) - C(x_1))^2} \\
B &= \frac{1}{4}\left(C(x_0) + C(x_1) + (C(x_2) + C(x_3)\right)
\end{aligned}
$$

Implementation Based on CCD/CMOS Technologies

- The CCD (charge-coupled device) plays several roles:
  - Data acquisition or readout operation: the incoming photons are converted into electron charges.
  - Clocking.
  - Signal Processing (demodulation).
- After the demodulation, the signal $C(\psi)$ is integrated at four equally space intervals, over an equal-length Δt, within one modulation period T.
- These four signal values are stored independently.
- The cycle of integration and storage can be repeated over many periods.
- Example: for $f = 30$MHz and at 30 frames per second (FPS), 10^6 integration periods are possible.

Depth from Phase and Modulation Frequency

- Depth: $d = \frac{1}{2}c\tau = \frac{c}{2f}\frac{\phi}{2\pi}$
- Minimum depth: $d_{\min} = 0 \ \ (\phi = 0)$
- Maximum depth: $d_{\max} = \frac{c}{2f} \ \ (\phi = 2\pi)$.
- Phase wrapping: an inherent 2π phase ambiguity.
- Hence the distance is computed up to a **wrapping** ambiguity:

$$d(i,j) = \left(\frac{\phi(i,j)}{2\pi} + n(i,j) \right) d_{\max}$$

where $n = 0, 1, 2, \ldots$ is the number of wrappings. This can also be written as:

$$d(i,j) = d_{\mathsf{tof}}(i,j) + n(i,j) d_{\max}$$

where d is the *real* depth value and d_{tof} is the *measured* depth value.

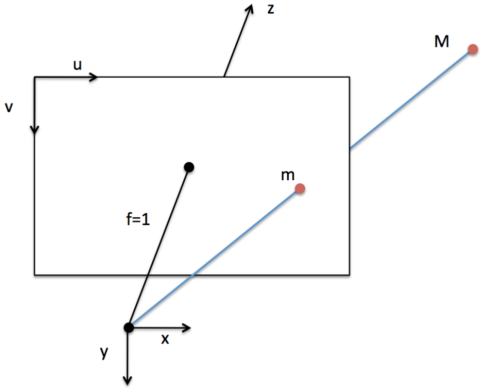
More on Phase Wrapping

- For $f = 30$MHz, the unambiguous range is from $d_{\min} = 0$ to $d_{\max} = 5$ meters.
- Solving for this ambiguity is called **phase unwrapping**.
- The modulation frequency of the SR4000 camera can be selected by the user. The camera can be operated at:
  - $29/30/31$ MHz corresponding to a maximum depth of 5.17m, 5m and 4.84m.
  - $14.5/15/15.5$ MHz corresponding to a maximum depth of 10.34m, 10m and 9.67m.
- The accuracy increases with the modulation frequency.

Observations

- A TOF camera works at a very precise modulation frequency. Consequently, it is possible to simultaneously and synchronously use several TOF cameras by using a different modulation frequency for each one of the cameras, e.g., six cameras in the case of the SR4000.
- A TOF camera needs a long integration time (IT), over several time periods, to increase SNR, hence accuracy. In turn, this introduces "motion blur" in the presence of moving objects. Because of the need of long IT, fast shutter speeds (as done with standard cameras) cannot be envisaged.
- Sources of errors:
  - Demodulation error;
  - Integration error;
  - Temperature error.

Back to Camera Model



- The coordinates of the image point m:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \approx \mathbf{A}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

- The coordinates of the world point M:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = Z \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

- Depth along a line of sight through m: $\frac{\|\boldsymbol{M}\|}{\|\boldsymbol{m}\|} = \frac{Z}{1}$ (1 is the focal length)

From Depth Values to Euclidean Coordinates

- The TOF camera measures the depth d from the 3D point M to the optical center, hence:

$$d = \|\boldsymbol{M}\|$$

.

- Relationship between the *image coordinates* $\boldsymbol{p} = (u\ v\ 1)^\top$ and the *virtual camera coordinates* $\boldsymbol{m} = (x\ y\ 1)^\top$ of the point m:

$$\boldsymbol{m} = \mathbf{A}^{-1}\boldsymbol{p}$$

- Hence:

$$Z = \frac{\|\boldsymbol{M}\|}{\|\boldsymbol{m}\|} = \frac{d}{\|\mathbf{A}^{-1}\boldsymbol{p}\|}$$

- We obtain for the Euclidean coordinates of the measured point M:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{d}{\|\mathbf{A}^{-1}\boldsymbol{p}\|}\mathbf{A}^{-1}\begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

Lens Distortion Model

- For TOF camera calibration it is advised to take into account lens distortion. For that purpose, we introduce $\boldsymbol{m}_l = (x_l \; y_l \; 1)$ *between* the image and camera coordinates: $\boldsymbol{p} = \mathbf{A}\boldsymbol{m}_l$. The lens distortion model is:

$$\left(\begin{array}{c} x_l \\ y_l \end{array} \right) = l_1(r) \left(\begin{array}{c} x \\ y \end{array} \right) + \boldsymbol{l}_2(x, y)$$

with:

$$\left\{ \begin{array}{rcl} l_1(r) & = & 1 + a_1 r^2 + a_2 r^4 \\ \boldsymbol{l}_2(x, y) & = & \left[\begin{array}{ll} 2xy & r^2 + 2x^2 \\ r^2 + 2y^2 & 2xy \end{array} \right] \left(\begin{array}{c} a_3 \\ a_4 \end{array} \right) \\ r^2 & = & x^2 + y^2 \end{array} \right.$$

TOF Camera Calibration

- The CCD image sensor of a TOF camera provides both a depth image and an amplitude+offset image.
- The amplitude+offset image can be used by the OpenCV packages to calibrate the camera, namely its intrinsic and lens-distorsion parameters.
- However, the low-resolution of the TOF images implies some practical considerations.