Sensors

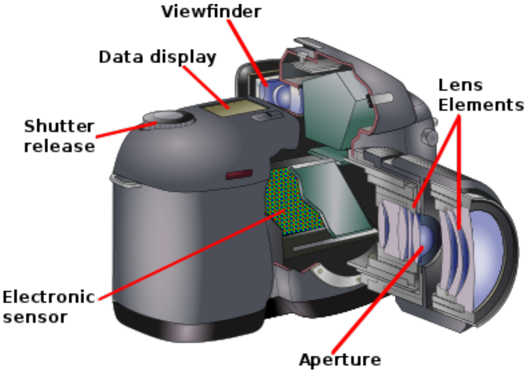
CSE 6367: Computer Vision

Instructor: William J. Beksi

Introduction

- Sensors enable the ability to capture image or range data (or both) of a scene
- Sensor data can be used in numerous applications (robotics/machine vision, factory automation, autonomous vehicles, etc.)
- The predominate sensor is the 2D digital camera, however in recent years low-cost 3D sensors have become increasingly available

The Digital Camera



The Digital Camera

- After starting from one or more light sources, reflecting off one or more surfaces and passing through the camera's optics (lenses), light finally reaches the imaging sensor
- How are the photons arriving at the sensor converted into the digital (R,G,B) values that we observed when looking at a digital image?

The Digital Camera

- Light falling on an imaging sensor is picked up by an **active sensing area**, integrated for the duration of the exposure (expressed as the shutter speed in a fraction of a second, e.g. $\frac{1}{125}, \frac{1}{60}, \frac{1}{30}$), and then passed to a set of **sense amplifiers**
- The two main sensors used in digital still and video cameras today are **charge-coupled device** (CCD) and **complementary metal oxide on silicon** (CMOS)

Charge-Coupled Device (CCD)

- In a CCD, photons are accumulated in each active *well* during exposure time
- Then, during a *transfer* phase, the charges are transferred from well to well ("bucket brigade") until they are deposited at the sense amplifiers
- The sense amplifiers amplify the signal and pass it to an analog-to-digital converter (ADC)

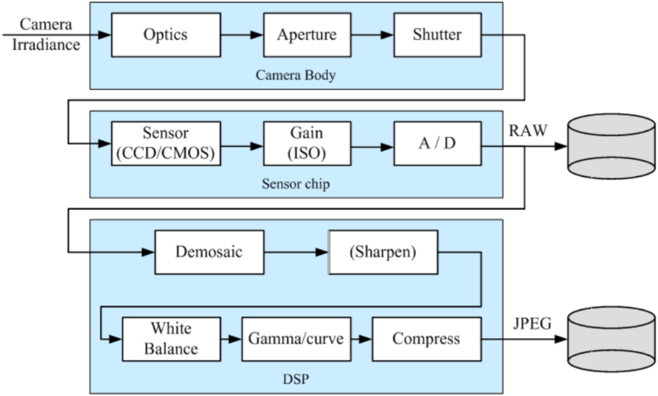
Complementary Metal Oxide on Silicon (CMOS)

- In CMOS, the photons hitting the sensor directly affect the conductivity (or gain) of a photodetector
- The photodetector can be selectively gated to control exposure duration and locally amplified before read out using a multiplexing scheme
- Traditionally, CCD sensors outperformed CMOS in quality sensitive applications while CMOS was better for low-power applications
- Today, CMOS is used in most digital cameras

Performance Factors

- The main factors affecting the performance of a digital image sensor are the shutter speed, sampling pitch, fill factor, chip size, analog gain, sensor noise, and the resolution (and quality) of the ADC
- Many of these performance parameters can be read from the EXIF flags embedded in a digital image while others can be obtained from the camera manufacturers' specification sheets or from camera review or calibration websites

Image Sensing Pipeline



Shutter Speed

- The **shutter speed** (exposure time) directly controls the amount of light reaching the sensor and thus determines if images are under- over-exposed
- For dynamic scenes, the shutter speed determines the amount of **motion blur** in the resulting picture
- Usually, a high shutter speed (less motion blur) makes subsequent analysis of the image easier

Sampling Pitch

- The **sampling pitch** is the physical spacing between adjacent sensor cells on the imaging chip
- A sensor with a smaller sampling pitch has a higher **sampling density** and hence provides a higher **resolution** (in terms of pixels) for a given active chip area
- However, a smaller pitch means that each sensor has a smaller area and cannot accumulate as many photons; this makes it not as **light sensitive** and more prone to noise

Fill Factor

- The **fill factor** is the active sensing area size as a fraction of the theoretically available sensing area (the product of the horizontal and vertical sampling pitches)
- Higher fill factors are usually preferable as they result in more light capture and less **aliasing**
- However, this must be balanced with the need to place additional electronics between the active sense areas

Chip Size

- Video and point-and-shoot cameras have traditionally used small chip areas ($\frac{1}{4}$-inch to $\frac{1}{2}$-inch sensors), while digital SLR cameras try to come closer to the traditional size of a 35 mm film frame
- When overall device size is not important, having a larger chip size is preferable since each sensor cell can be more photo-sensitive
- However, larger chips are more expensive to produce not only because fewer chips can be packed into each wafer, but also because the probability of a chip defect increases linearly with the chip area

Analog Gain

- Before ADC, the sensed signal is usually boosted by a **sense amplifier**
- In video cameras, the gain on these amplifiers was traditionally controlled by **automatic gain control** (AGC) logic which would adjust these values to obtain a good overall exposure
- In newer digital cameras, the user has some additional control over this gain through the **ISO setting** (typically expressed in ISO standard units such as 100, 200, or 400)

Sensor Noise

- Throughout the whole sensing process, noise is added from various sources: fixed pattern noise, dark current noise, shot noise, amplifier noise, quantization noise, etc.
- The final amount of noise present in the sampled image depends on all of these quantities as well as the incoming light, exposure time, and sensor gain
- It is possible to estimate the **noise level function** (NLF) for a given image which predicts the overall noise variance at a given pixel as a function of its brightness

ADC Resolution

- The final step in the analog processing chain of the imaging sensor is the **analog to digital conversion** (ADC)
- The two quantities of interest are the **resolution** of this process (how many bits it yields) and its **noise level** (how many of these bits are useful in practice)
- For most cameras, the number of bits quoted (e.g. 8 bits for compressed JPEG images) exceeds the actual number of usable bits

Digital Post-Processing

- Once the irradiance values arriving at the sensor have been converted to digital bits, most cameras perform a variety of **digital signal processing** (DSP) operations to enhance the image before compressing and storing the pixel values
- These operations include: color filter array (CFA) demosaicing, white point setting, mapping of the luminance values through a gamma function to increase the perceived dynamic range of the signal, etc.

Rangefinder Sensors

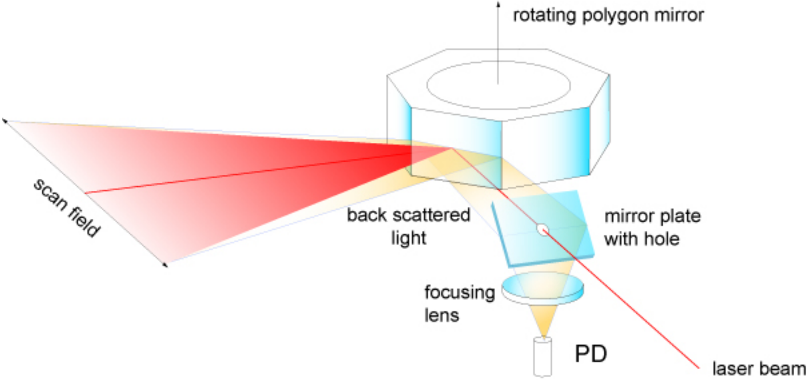
- A **rangefinder** is a device that measures distance from the sensor to a target
- Rangefinding methods use unilateral transmission and passive reflection
- These methods include laser, radar, sonar, light detection and ranging (LIDAR), and ultrasonic

2D Laser Scanner

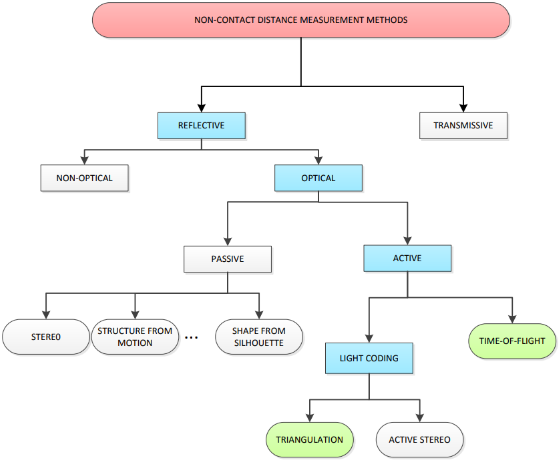
- Laser scanning is the controlled deflection of laser beams (visible or invisible)
- A laser scanner operates on the time of flight principle by sending out a laser pulse in a narrow beam and measuring the time taken by the pulse to be reflected off the surrounding objects and returned to the device

2D Laser Scanner

2D Laser Scanner



Distance Measurement Methods



3D Sensors

- Active 3D sensors can measure depth (distance) to an object by illuminating the object with a light source (e.g. infrared) and measuring the backscattered light
- There are two types of such sensors:
  - **Time-of-flight** (ToF) sensors measure depth by estimating the time delay from light emission to light detection
  - **Structured-light** sensors combine the projection of a light pattern with a standard 2D camera and that measure depth by triangulation

Time-of-Flight Principles

- A ToF sensor operates by measuring the absolute time that a pulse of light needs to travel from a targeted object to a detector
- The speed of light is of 0.3 meters/nanosecond
- ToF sensor systems use either **pulsed-modulation** or **continuous wave modulation**

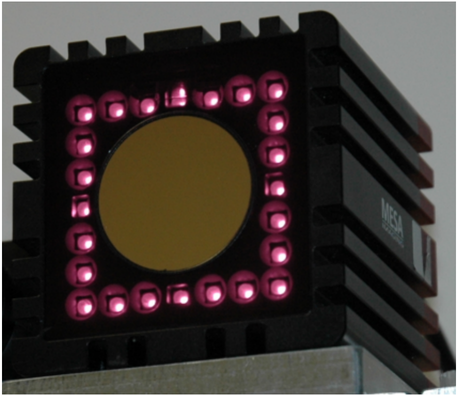
Pulsed-Modulation

- Pulsed-modulation measures the time-of-flight directly and allows for long-distance measurements
- The arrival time must be detected very precisely
- To do this, very short light pulses with fast rise and fall times along with high optical power (lasers or laser diodes) are used

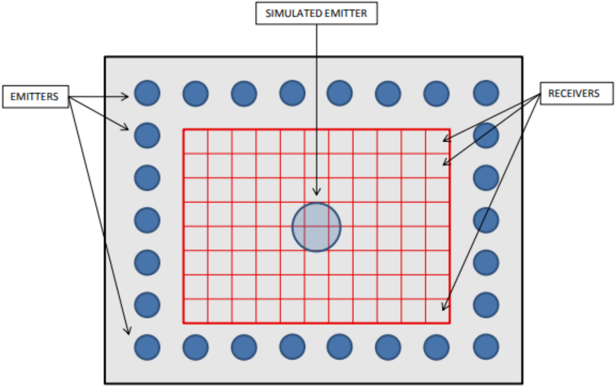
Continuous Wave Modulation

- Continuous wave modulation measures the phase difference between the sent and received signals
- Different shapes of signals are possible e.g. sinusoidal, square waves
- Cross-correlation between the received and sent signals allows phase estimation which is directly related to distance if the modulation frequency is known

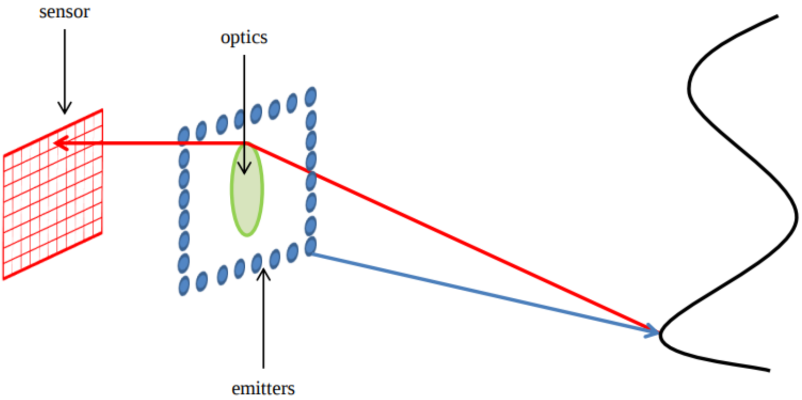
Time-of-Flight Sensor



Time-of-Flight Sensor



Time-of-Flight Sensor



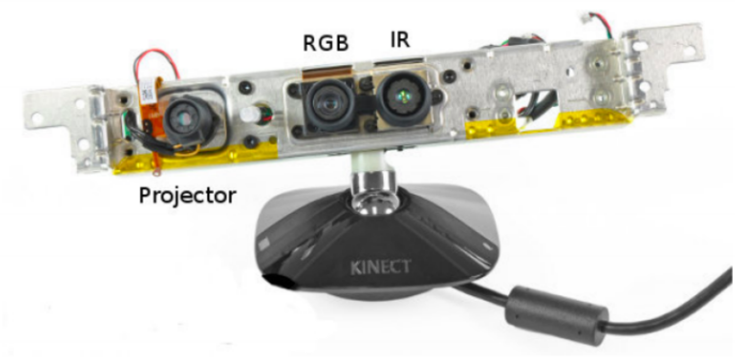
Time-of-Flight Sensor

Structured-Light Principles

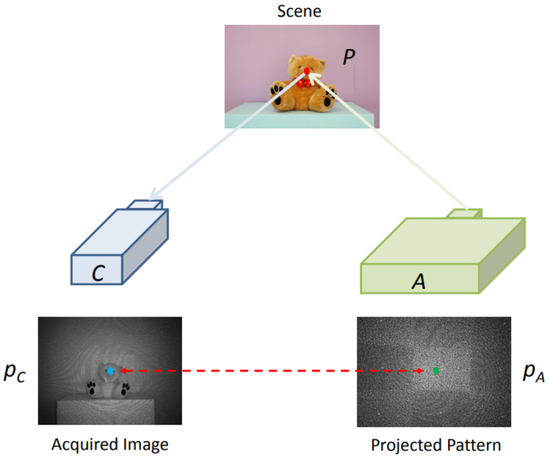
- A structured-light sensor measures the 3D shape of an object using projected light patterns and a camera system
- Structured light is the process of projecting a known pattern (often grids or horizontal bars) on to a scene
- By measuring the deformation of the pattern upon striking the surface of an object the depth information can be calculated

Structured-Light Sensor

Structured-Light Sensor



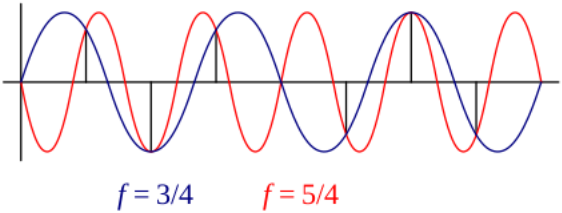
Matricial Active Triangulation



Sampling and Aliasing

- What happens when a field of light impinging on the image sensor falls onto the active sense areas of the imaging chip?
- The photons arriving at each active cell are integrated and then digitized
- However, if the fill factor on the chip is small and the signal is not otherwise band-limited, then visually unpleasing **aliasing** can occur

Aliasing of a 1D Signal



- The blue sine wave at $f = 3/4$ and the red sine wave at $f = 5/4$ have the same digital samples when sampled at $f = 2$, i.e. they are **aliased**
- Why is this a bad effect?

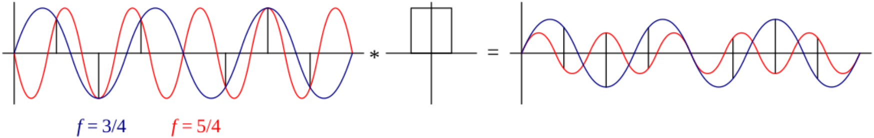
Minimum Sampling Rate

- Shannon's sampling theorem shows that the minimum sampling rate required to reconstruct a signal from its samples must be at least twice the highest frequency

$$f_s \geq 2f_{max}$$

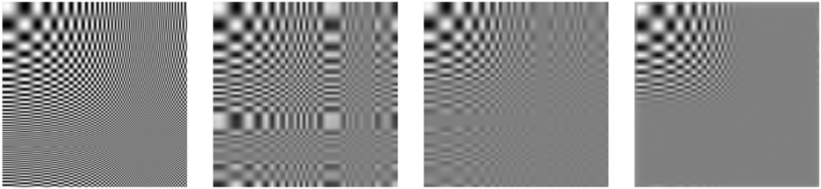
- The maximum frequency in a signal is known as the **Nyquist frequency** and the inverse of the minimum sampling frequency, $r_s = 1/f_s$, is known as the **Nyquist rate**

Filtering of a 1D Signal



- Even after convolution with a 100% fill factor box filter the two signals, while no longer of the same magnitude, are still aliased in the sense that the sampled red signal looks like an inverted lower magnitude version of the blue signal

Aliasing of a 2D Signal



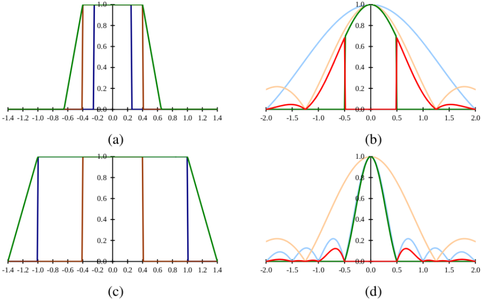
(a) (b) (c) (d)

- (a) original full-resolution image; (b) downsampled 4 x with a 25% fill factor box filter; (c) downsampled 4 x with a 100% fill factor box filter; (d) downsampled 4 x with a high-quality 9-tap filter

Predicting the Amount of Aliasing

- The best way to predict the amount of aliasing an imaging system will produce is to estimate the **point spread function** (PSF) which represents the response of a particular pixel sensor to an ideal point light source
- If we know the blur function of the lens and the fill factor (sensor area shape and spacing) for the imaging chip, then we can convolve these to obtain the PSF

Sample Point Spread Functions



(a) (b) (c) (d)

- The diameter of the blur disc (blue) in (a) is equal to half the pixel spacing while the diameter in (c) is twice the pixel spacing
- The horizontal fill factor of the sensing chip is 80% (brown)
- The convolution of these two kernels gives the PSF (green)
- The Fourier response of the PSF is shown in (b) and (d)

Summary

- The most widely used sensor in computer vision is the digital camera
- 3D sensors are becoming increasingly popular as a low-cost alternative to complicated stereo vision systems
- Aliasing effects both the quality of the image and the ability to reconstruct the original signal