What is a Robot?

ESE 6510
Antonio Loquercio

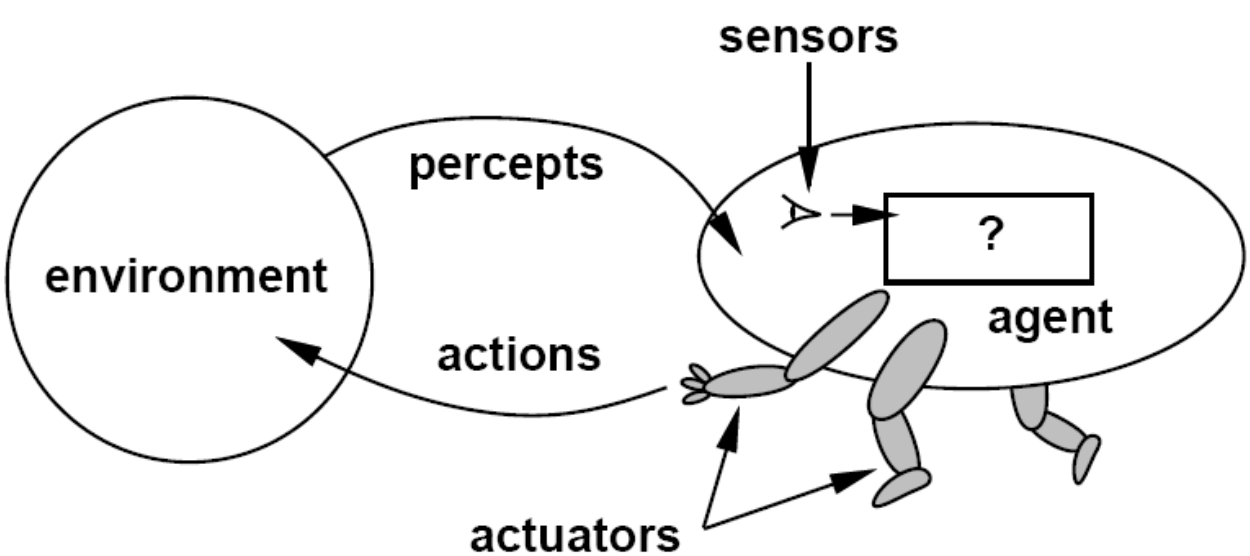
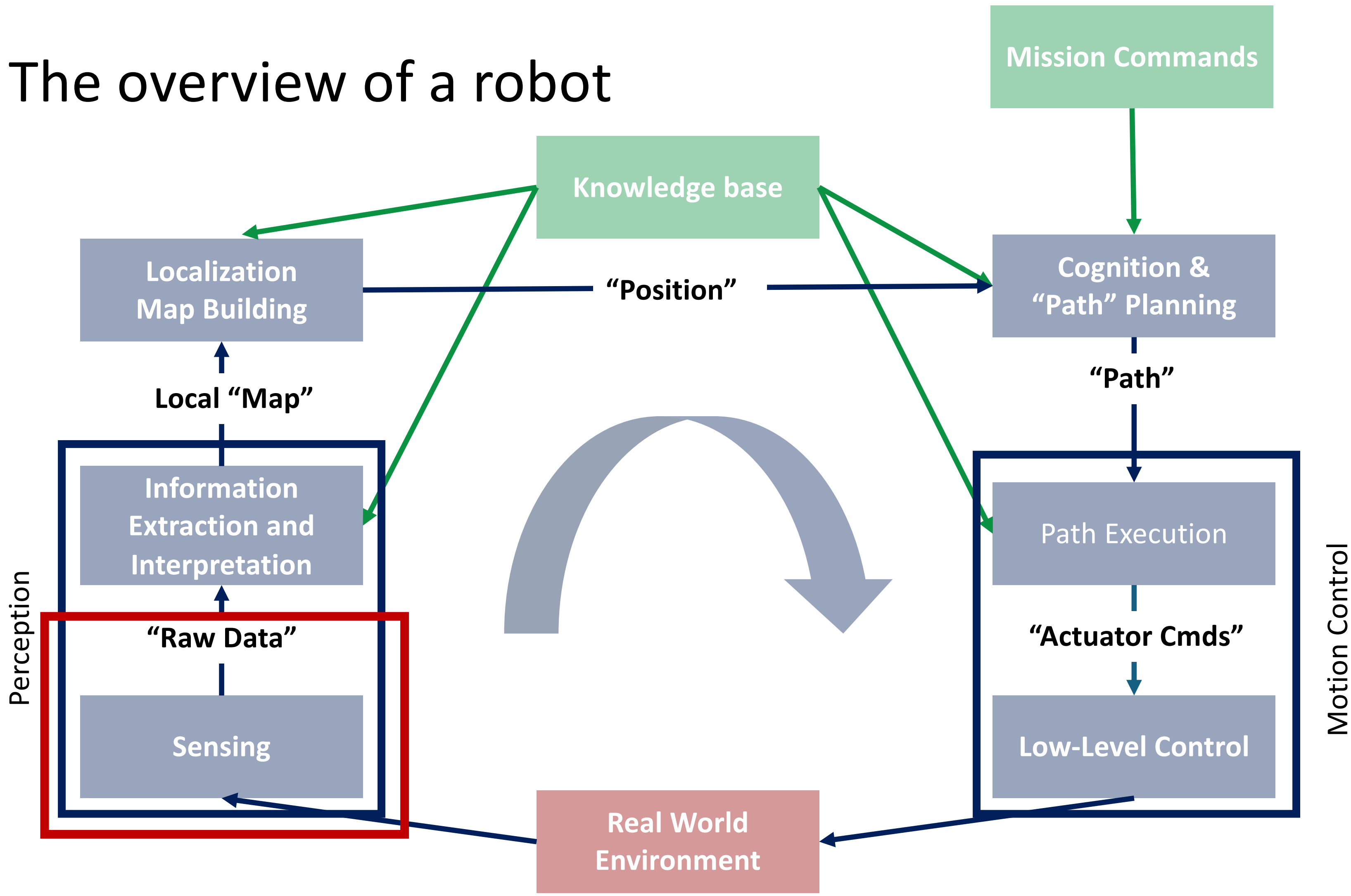


Figure from Russell & Norvig

The overview of a robot



Goal of Today's Lecture: Perception

- Understand general sensor categorizations
- Learn how different sensors work and when they are useful
- Understand that the perfect sensor does not exist, and there is still a lot we can learn from biology.

Sensor Classification

- *Proprioceptive* (Measure values internal to the agent)
  - Motor Speed
  - Wheel Load
  - Joint Angles
  - Battery Voltage
  - …

Aside: History of Proprioception

- Aristotle's “De Anima” listed 5 senses: sight, hearing, touch, taste, and smell.
- Bell (1826) proposed a sixth sense: the “muscular” sense. It was renamed “proprioception” by Sherrington.
- In humans, it refers to the awareness of the position or movement of the body or parts of the body relative to each other.
  - Musculoskeletal system (or kinesthesia): sensors in muscles, tendons, and joints.
  - Vestibular system in the inner ear: semicircular canals

Slide from J. Malik

Demo Time!

Aside: The Importance of Proprioception in Humans

Sensor Classification

- *Passive* (Measure values internal to the robot)
  - Motor Speed
  - Wheel Load
  - Joint Angles
  - Battery Voltage
  - …
- *Exteroceptive* (Measures values about the environment)
  - Distance measurements
  - Video Stream
  - Depth (e.g., LIDAR)
  - Sound
  - …

Sensor Classification (Energy-Based)

- *Passive* (Measure the environment's energy entering the sensor)
    - Temperature Probes
    - Microphones
    - CCD/CMOS Cameras
    - …

- *Active* (Emit energy into the environment)
    - Wheel quadrature encoders
    - Ultrasonic Sensors
    - LIDAR
    - …
    - **Often give superior performance but come with limitations. Examples?**

Characterizing Sensor Performance

Response Ratings

- Sensitivity [dB]. Measures how an incremental change in input changes the output. Complex to estimate.

- Resolution [varies] . Minimum difference between two values that can be detected. Not necessarily equal to the minimum inputs (e.g., digital).

- Linearity []. Measures the behavior of the sensor as the input varies.

- Bandwidth or frequency [Hz]. Measures the speed at which a sensor can provide a stream of readings.

Characterizing Sensor Performance

Error Characterization

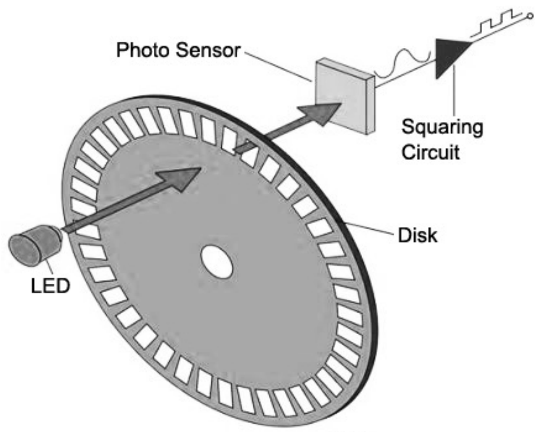
- Sensitivity [varies]. The degree to which an incremental change in the target input changes the output.
- Cross-Sensitivity [varies] . Sensitivity to other environment parameters orthogonal to the target parameters for the sensor.
- Error [varies]. Difference between “true” value and measured value. Generally reported in terms of accuracy. Divides into systematic and random errors.
- Precision []. Measures consistency between the sensor readings.

$$precision = \frac{range}{\sigma}$$

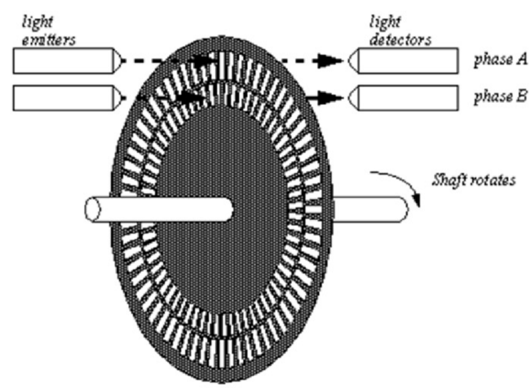
Sensors

Motor/Joint Sensors

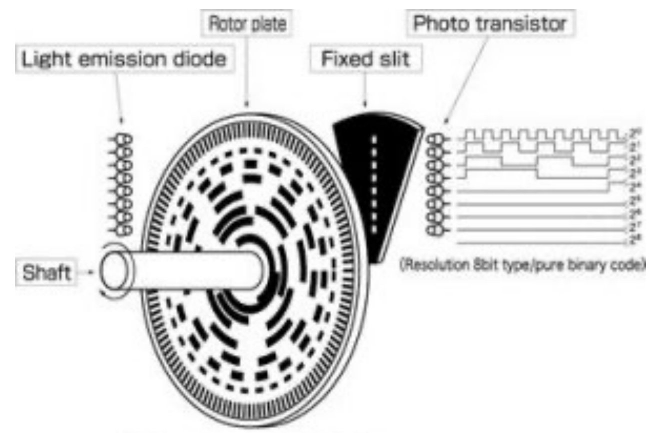
- Optical Encoders



Traditional



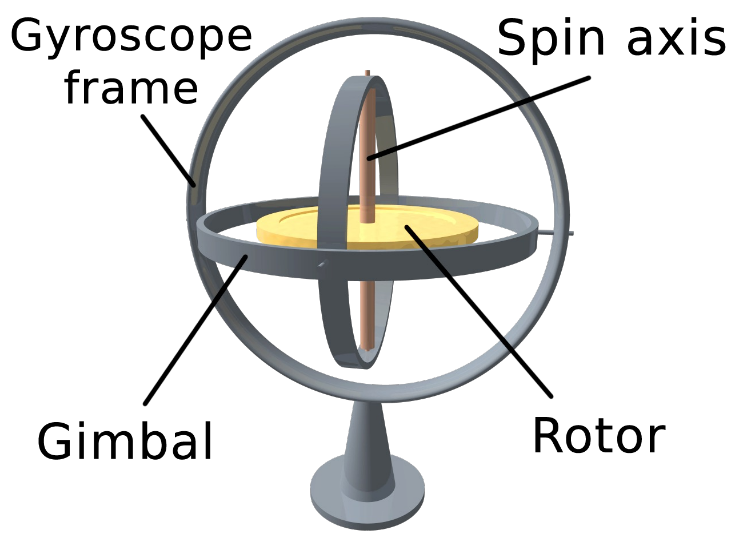
Quadrature



Absolute

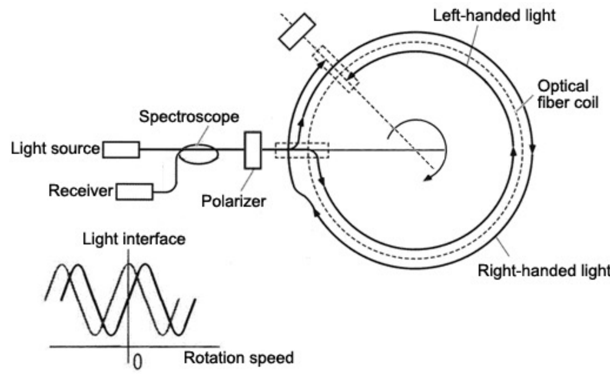
Heading Sensors

- Compasses: The Hall Effect or Flux Gate.
- Gyroscopes:
  - *Mechanical (Used throughout history, named by Leon Foucault). Based on the inertial properties of a fast-spinning rotor. Adding springs allows to measure angular speed.*



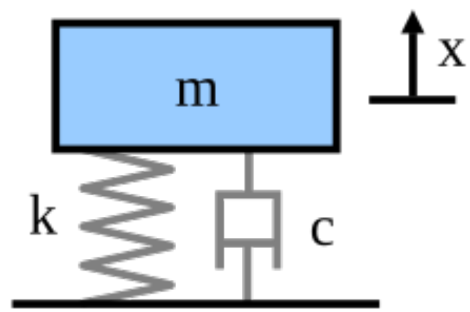
Heading Sensors

- Compasses: The Hall Effect or Flux Gate.
- Gyroscopes:
  - Mechanical *(Used throughout history, named by Leon Foucault). Based on the inertial properties of a fast-spinning rotor. Adding springs allows us to measure angular speed.*
  - Optical. Uses lasers and measurements of the difference in the speed of light between two coils mounted in opposite directions (clockwise and counterclockwise).



Accelerometers

Device measuring all external accelerations (including gravity!) acting upon it.



- Conceptually, a spring-mass-damper system.
- Different types of measurements:
  - Micro Electro-Mechanical Systems
  - Capacitive
  - Piezoelectric
- Two types: Static (0-500Hz), Dynamic (10Hz-50kHz).

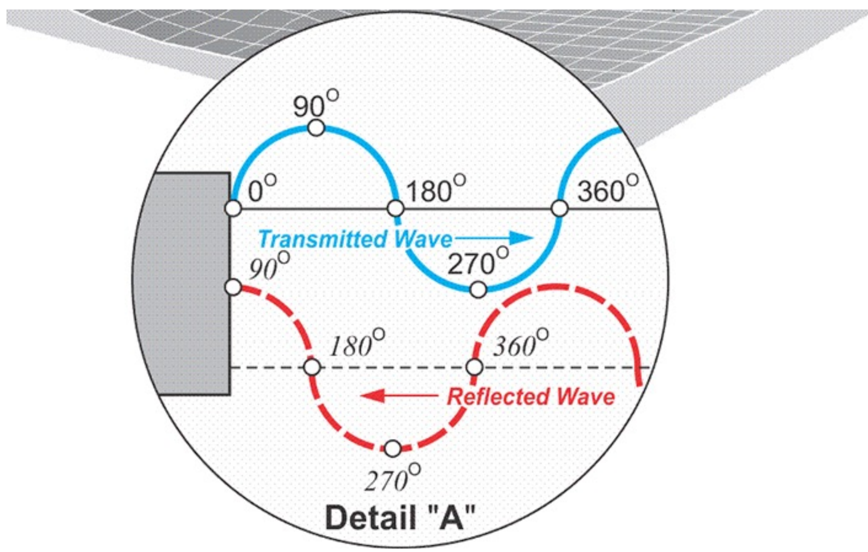
IMUs

- Integrates a gyroscope and an accelerometer.
- Can you give a concrete example of why having a gyro together with an accelerometer is a good idea?
- Large cost range. The more expensive, the less the drift.
- They are surprisingly good


VIO
3D-RONIN
TLIO

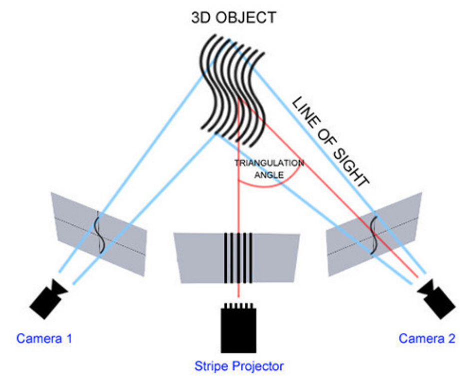
Active Range Sensors

- Time-of-flight:
  - Ultrasonic
  - Laser rangefinders. Generally referred to as lidar (light detection and ranging)
  - Phase shift measurement, approximately $d = (\lambda * N + \frac{\theta}{2\pi})/2$



Active Range Sensors

- Time-of-flight:
  - Ultrasonic
  - Laser rangefinders. Generally referred to as lidar (light detection and ranging)
  - Phase shift measurement, approximately $d = (\lambda * N + \frac{\theta}{2\pi})/2$
  - ToF Cameras. Similar to a 3D rangefinder, but 3D scene is captured all at the same time and there are no moving parts.
  - Triangulation sensors. Use structured light to project a pattern and see how that pattern is changed by the environment.

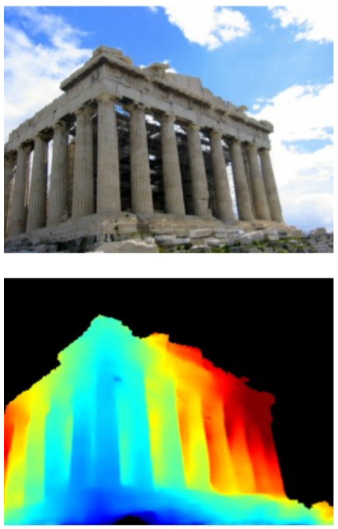


Active Range Sensors

- Time-of-flight:
  - Ultrasonic
  - Laser rangefinders. Generally referred to as lidar (light detection and ranging)
  - Phase shift measurement, approximately $d = (\lambda * N + \frac{\theta}{2\pi})/2$
  - ToF Cameras. Similar to a 3D rangefinder, but 3D scene is captured all at the same time and there are no moving parts.
  - Triangulation sensors. Use structured light to project a pattern and see how that pattern is changed by the environment.
  - Doppler effect sensing for speed measurement.

Vision

- You’ve probably already heard a lot about cameras.
- We can extract a lot of information from it without specific sensors



Parthenon, Athens

Megadepth, Li et al.

Segment Anything, Kirillov et al.

ORB-SLAM 2, Mur-Artal et al.

Pop Quiz!

- What’s the difference between a rolling shutter and a global shutter camera? What are the advantages and limitations of each?

- What is the difference between visual odometry and SLAM?

- Can I estimate the full 6D metric trajectory of a single camera moving around?

- What are good feature extraction methods for images?

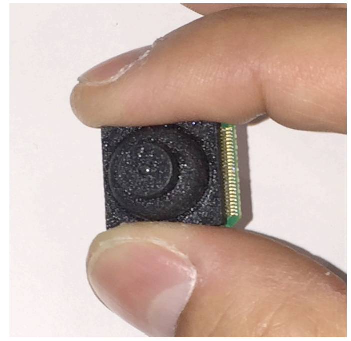
Some Cool Vision Sensors

Wearable Cameras (e.g., Meta Aria Glasses)

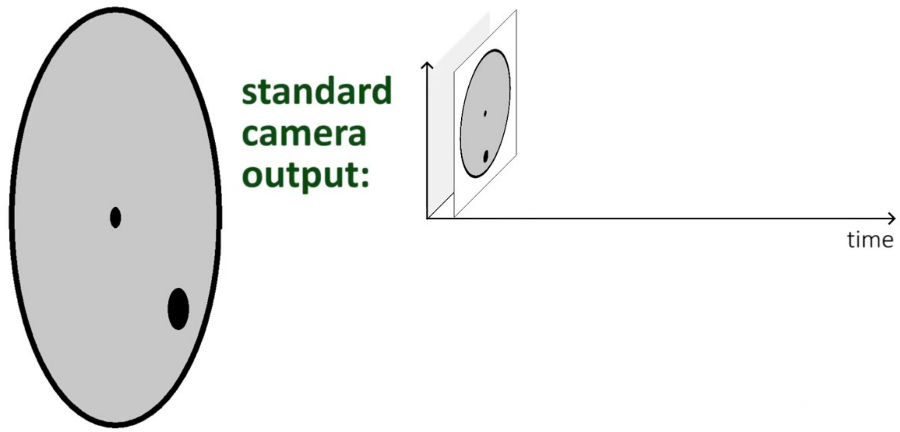
Wearable Cameras (e.g., Meta Aria Glasses)

Event Cameras

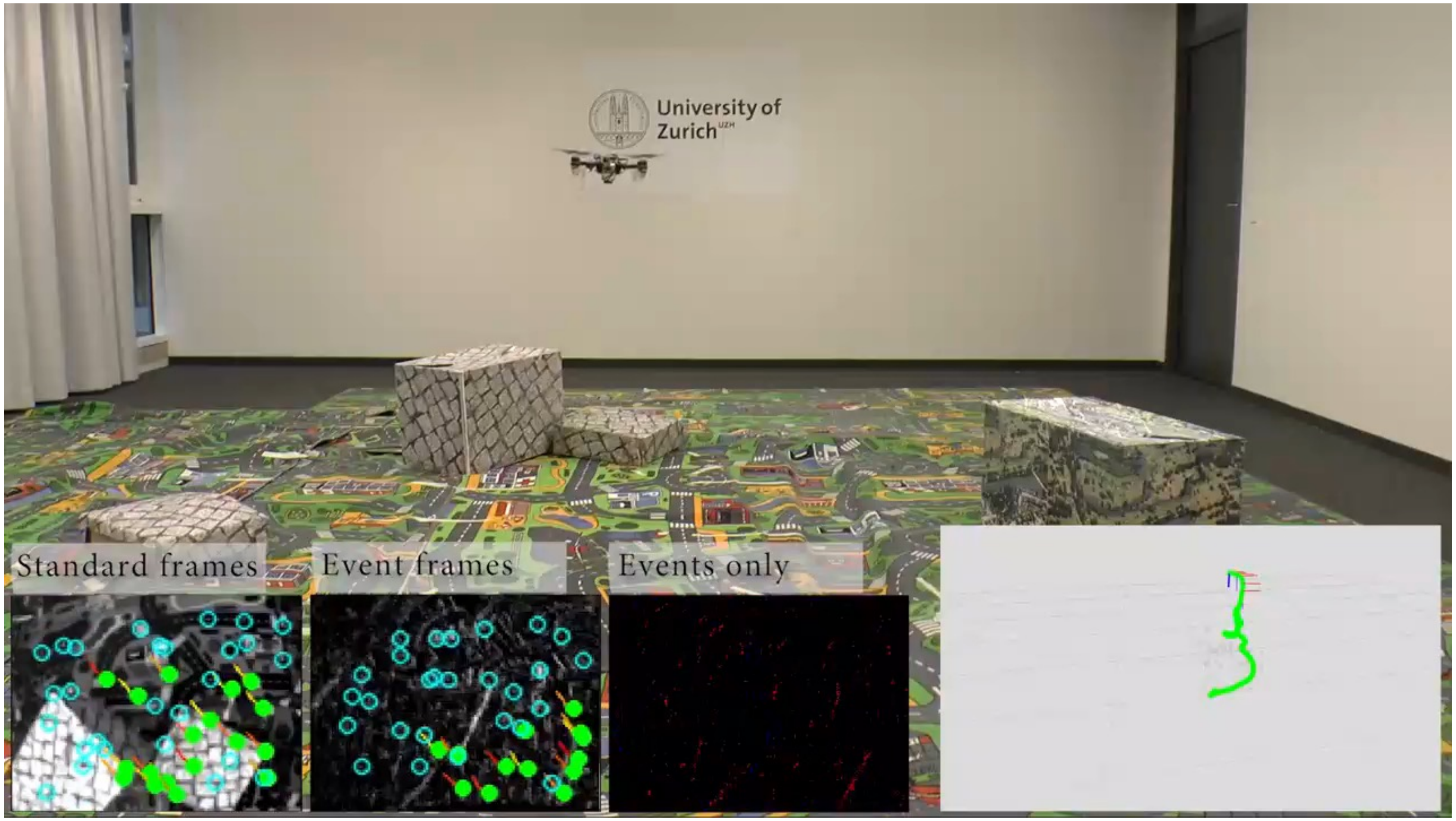
- Novel sensor that measures only **motion in the scene**
- **Low-latency** ($\sim 1\ \mu s$)
- **Ultra-low power** (mean: 1 mW vs. 1 W)
- **High dynamic Range** (140 dB instead of 60 dB)



Mini DVS sensor from IniVation.com

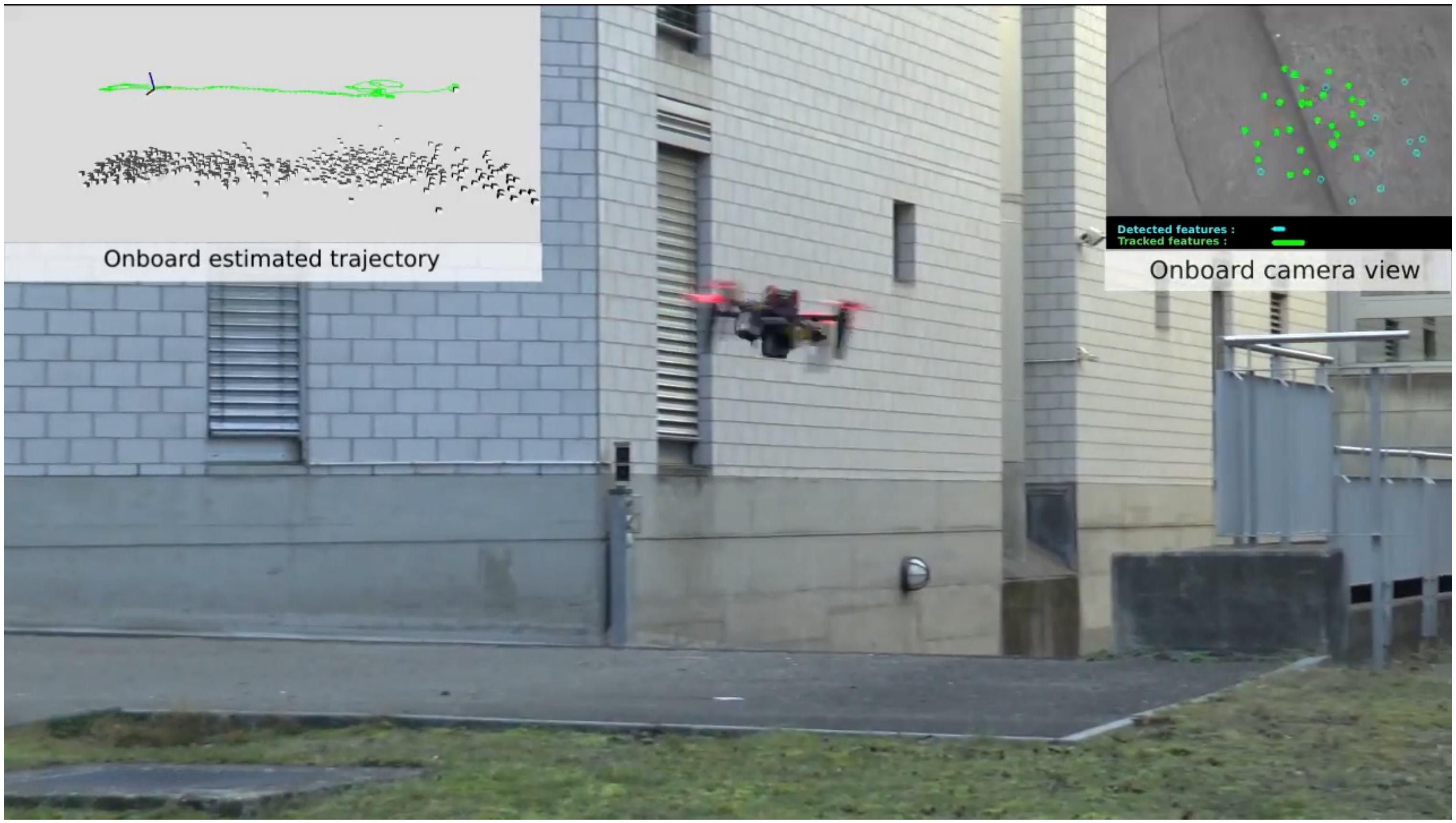


Autonomous Drone Flight in Low Light



Vidal et al., “Ultimate SLAM? Combining Events, Images, and IMU for Robust Visual SLAM in HDR and High Speed Scenarios”, RAL 2017

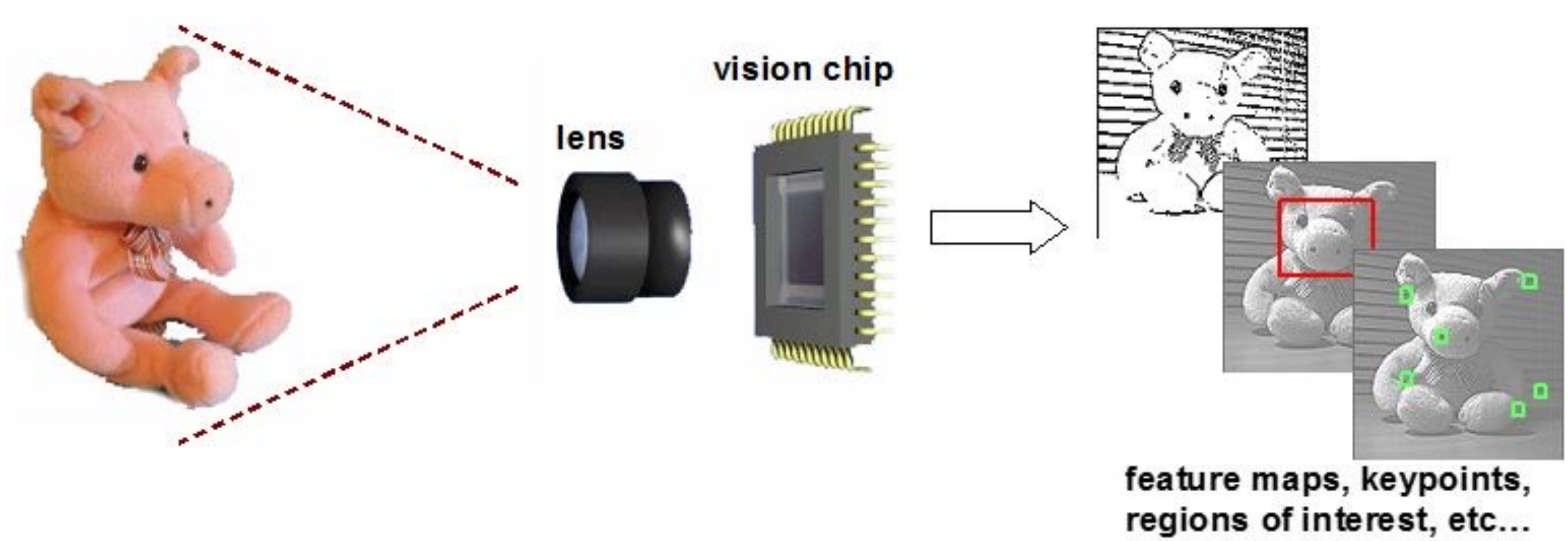
Recovering from Rotor Failure



Sun et al., “Autonomous Quadrotor Flight Despite Rotor Failure With Onboard Vision Sensors: Frames vs. Events“, RAL 2022

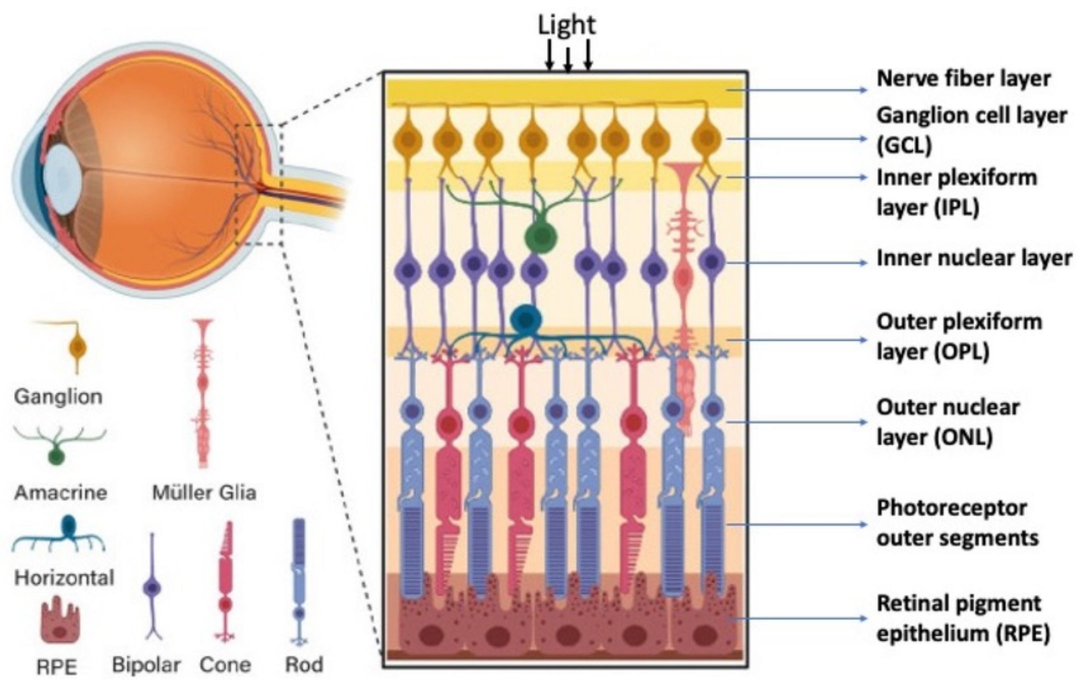
The SCAMP Vision Sensor

Unlike a conventional image sensor, it does not output raw images, but rather the results of **on-sensor computations**. It is fully programmable.



Nothing Beats Human Vision

- One of the sharpest in mammals. (1/6 deg.). Excels in daylight, but not too bad at night either.
- Quite different from a camera.




MATTERHORN
FULL ASCENT

You’re about to see the waiting room of a science laboratory.

See if you can spot the items in the room that are a little “out of place.”

Source: Michael Cohen

How did I not see this?

Start Frame

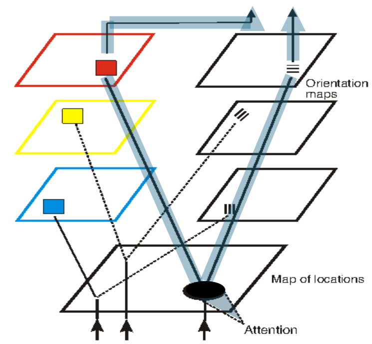
End Frame

Source: Michael Cohen

The Perception Paradox

A popular theory

- Limited capacity =>
  limited access to higher-level processing
- Need to attend to an item to see it fully
  - To bind features into a correct percept
  - To identify it
  - To detect changes to it
  - To remember it
- Without "selecting" an item, you have very limited "preattentive" info
  - Just "maps" of basic features (e.g. orientation, motion, color)



Feature Integration Theory

e.g. Treisman & Gelade, 1980;
Rensink, 2000

**Slide curtesy of
Ruth Rosenholtz**

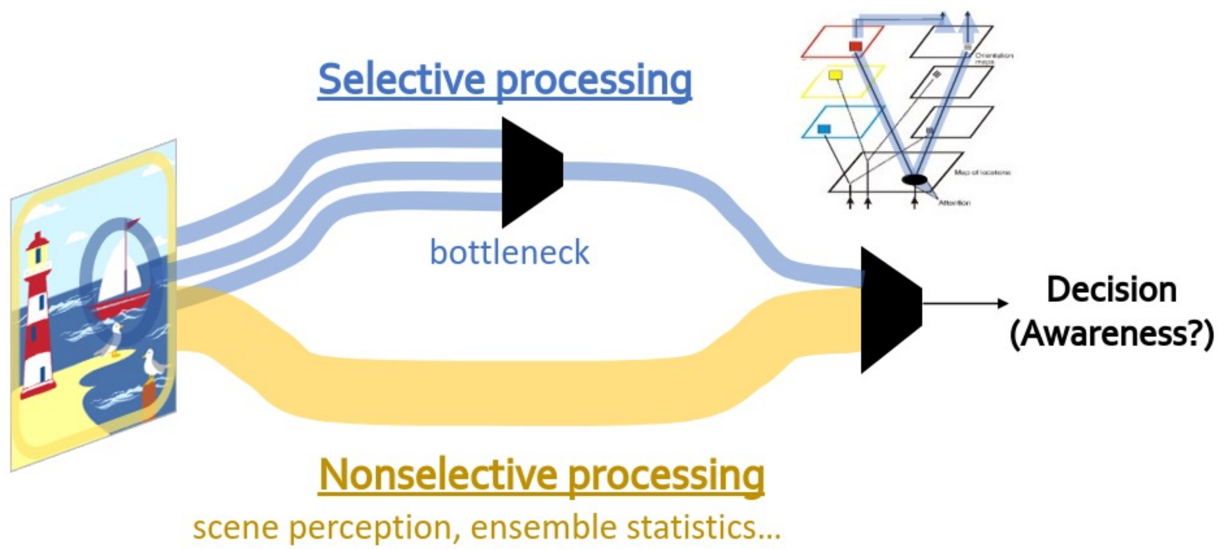
But Vision Feels Different...

First two hours of Matterhorn climbing are done in the dark.

The Attention Vicious Cycle

- Poor vision without selective attention
- Selective attention is limited
- Cannot fill in the details with memory, because that is limited

- *How do we select the attention point?*
- *How does real-world vision work so well?*



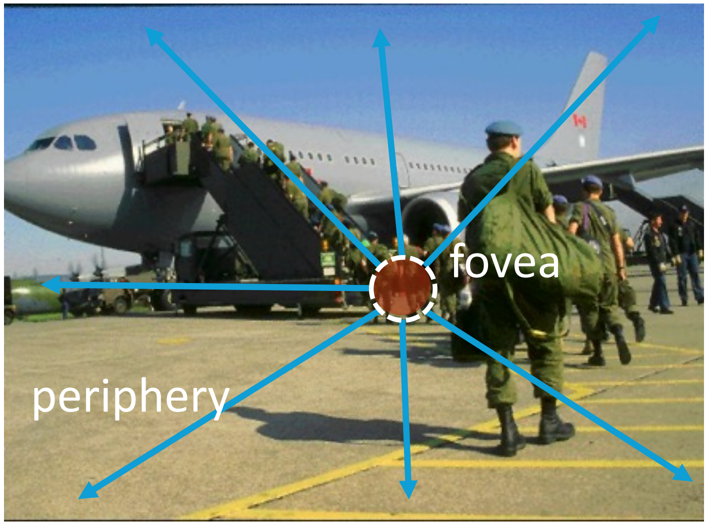
Attention <≠> Fixation

- **Attention**: "Visual attention refers to **the ability to prepare for, select, and maintain awareness of specific locations, objects, or attributes of the visual scene** (or an imagined scene). The focus of visual attention can be redirected to a new target either reflexively or through the purposeful effort of the observer."

- **Fixations:** "Times when **our eyes essentially stop scanning about the scene**, holding the central foveal vision in place so that the visual system can take in detailed information about what is being looked at." Everything else is peripheral vision.

Source: Google

Could peripheral vision be (part of) the explanation?

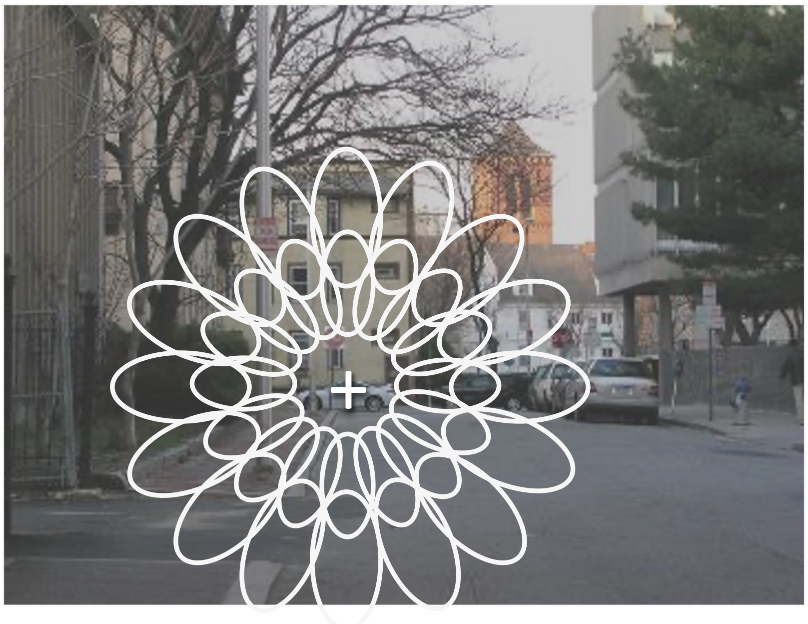


- Peripheral vision poor relative to foveal, especially when there is clutter
- Difficult detecting change where not attending? Or where not fixating?

Slide curtesy of Ruth Rosenholtz

The Texture Tiling Model (TTM) of Peripheral Vision

- Peripheral vision encodes local summary statistics
  - E.G. correlation between vertical stuff and horizontal stuff
- Farther from fixation ->
larger patches ->
more compression/less detail
- Successful at predicting human performance in over 70 visual tasks



Balas, Nakano, & Rosenholtz, 2009; Freeman & Simoncelli, 2011; **Rosenholtz et al., 2012ab**; **Zhang, et al., 2015**; **Chang & Rosenholtz, 2016**; **Ehinger & Rosenholtz, 2016**; **Keshvari & Rosenholtz, 2016**

Loss of resolution in peripheral vision is relatively modest

Original image

Blurred to mimic loss of resolution
(actually a 4x exaggeration)

Anstis, 1998

Slide curtesy of
Ruth Rosenholtz

Loss of resolution in peripheral vision is relatively modest

Original image

Blurred to mimic loss of resolution
(actually a 4x exaggeration)

Anstis, 1998

Slide curtesy of
Ruth Rosenholtz

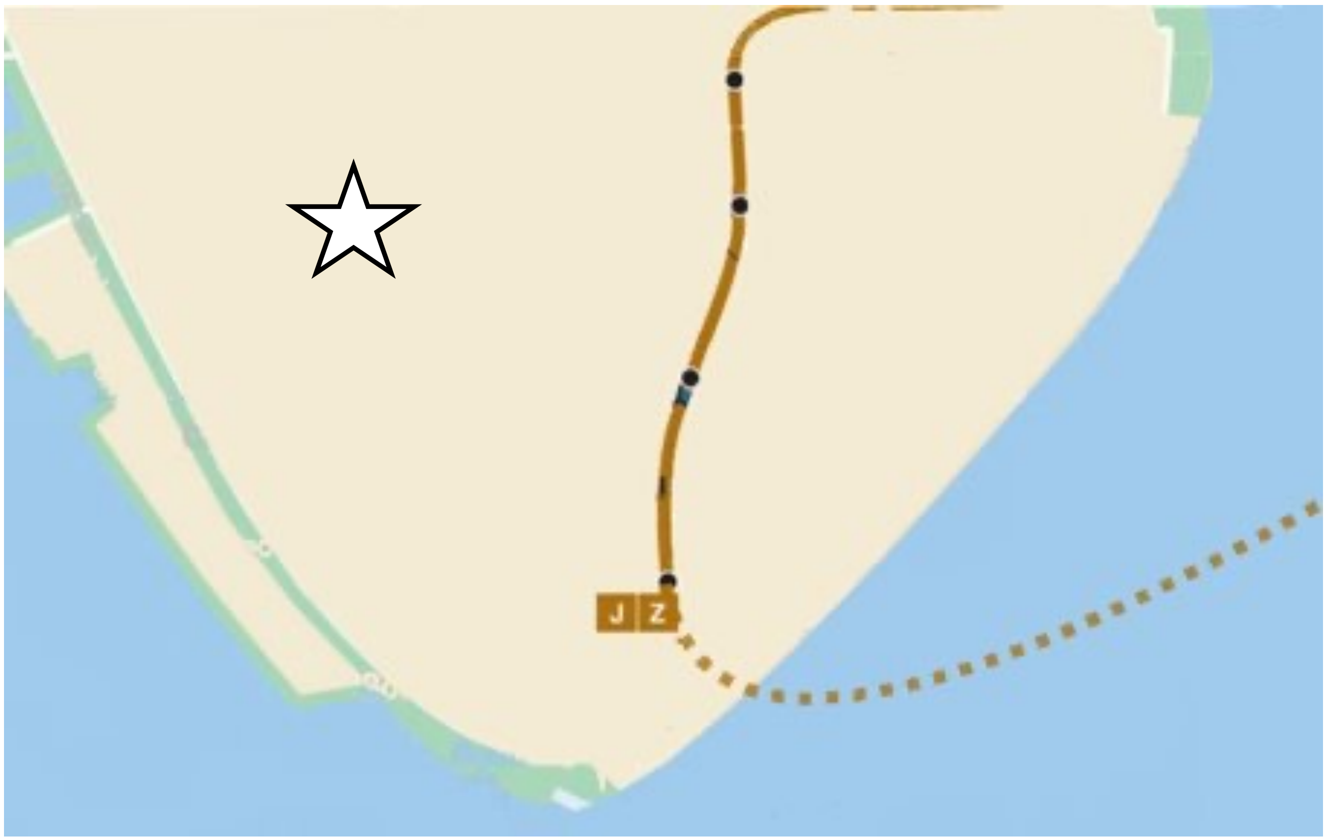
Loss of resolution is not the biggest problem...

Visual crowding causes significant loss of information

Slide curtesy of
Ruth Rosenholtz

Loss of resolution is not the biggest problem...

Visual crowding causes significant loss of information

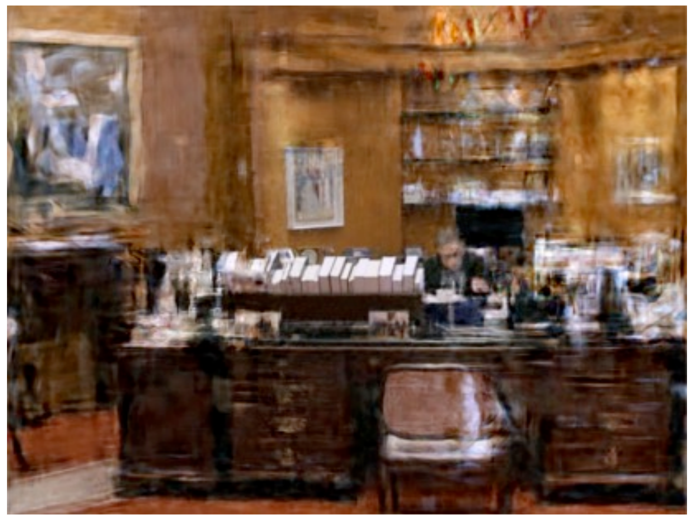


Slide curtesy of Ruth Rosenholtz

Peripheral vision has no chicken-and-egg problem

There's plenty of information to support a whole bunch of tasks, including noticing the painting on the wall to the left, or the person seated at the desk, and to suggest where you should look next.

Attention

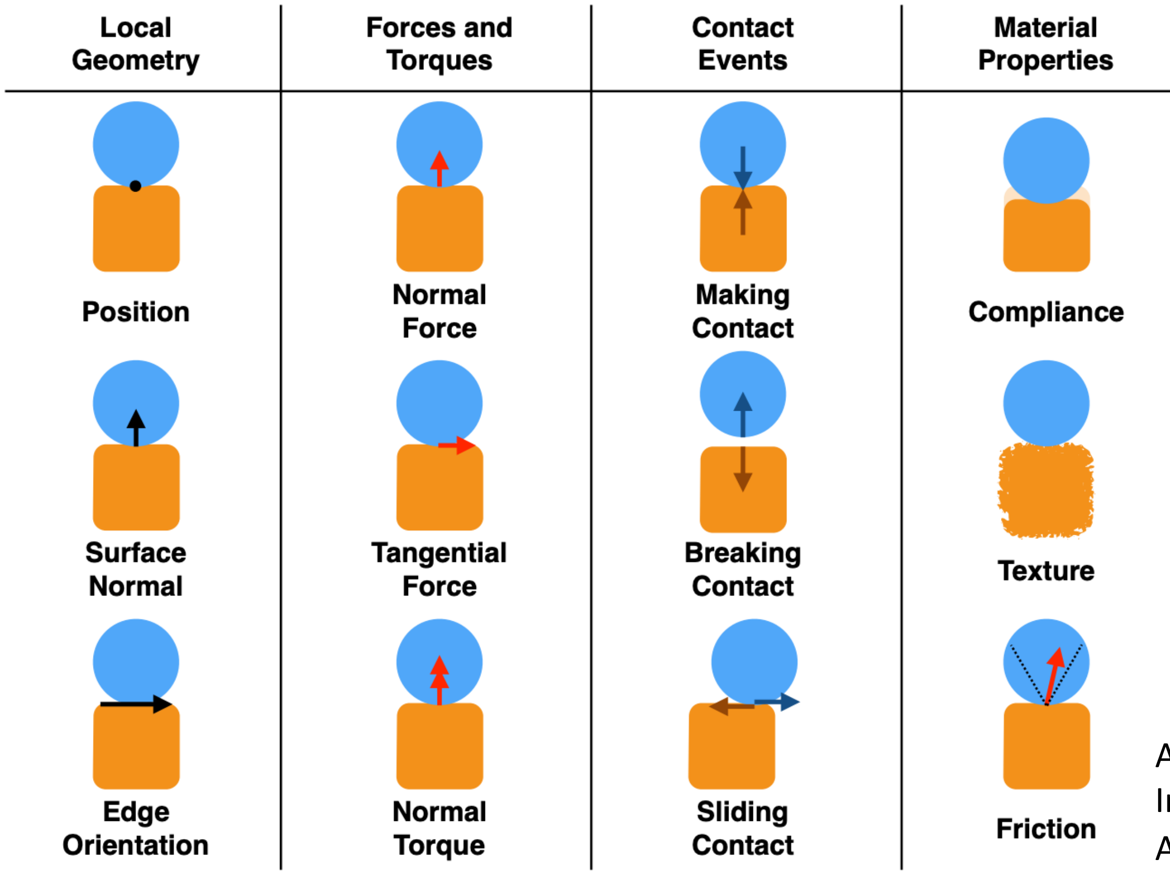


Mongrel

Takeaway: The human eye is an example of vertically integrated software and hardware.

Tactile Sensors

Information that can be extracted from tactile sensors



A Review of Tactile Information: Perception and Action Through Touch, Li et al., IEEE TRO

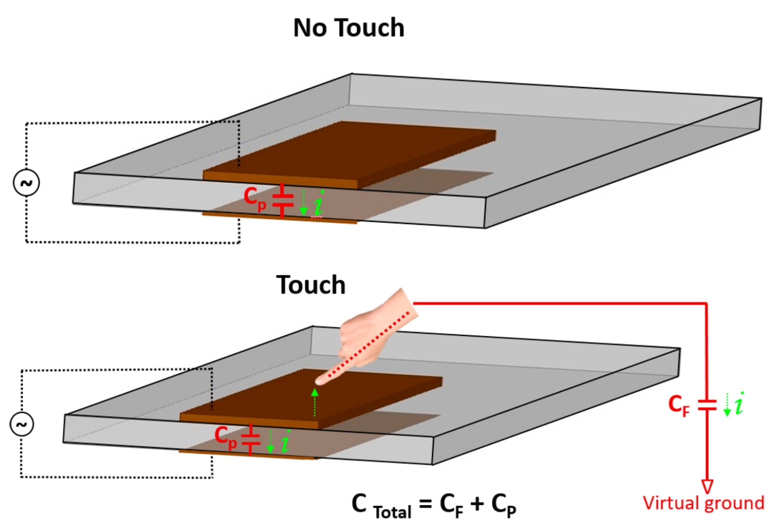
Tactile Sensing Technologies

- Capacitance
- Piezoresistivity
- Optics
- Magnetic
- Acoustic
- Piezoelectric

Capacitive Sensors

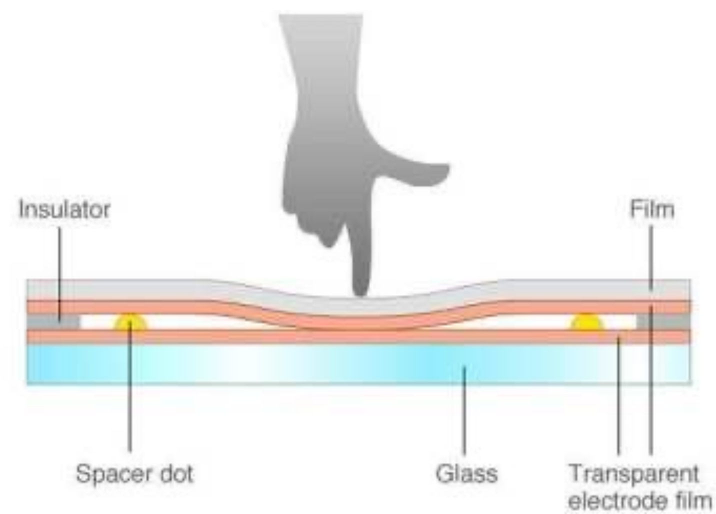
- Measure the variations in capacitance from an applied load over a parallel plate capacitor.

- Quite popular for high-end smartphones



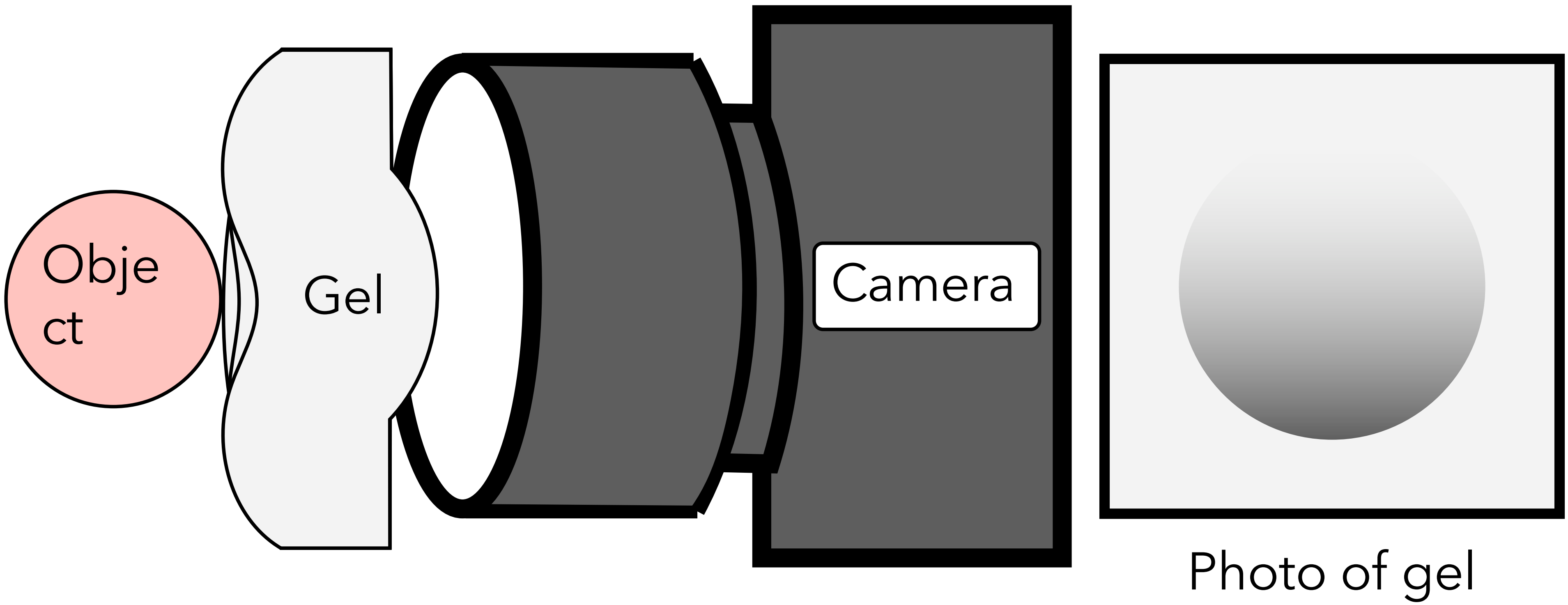
Piezo-Resistive Sensors

- Detect the resistance of a contact when an external force is applied
- These sensors produce a wide dynamic range, good load tolerance, durability, and low-cost fabrication.
- However, they are not very accurate.



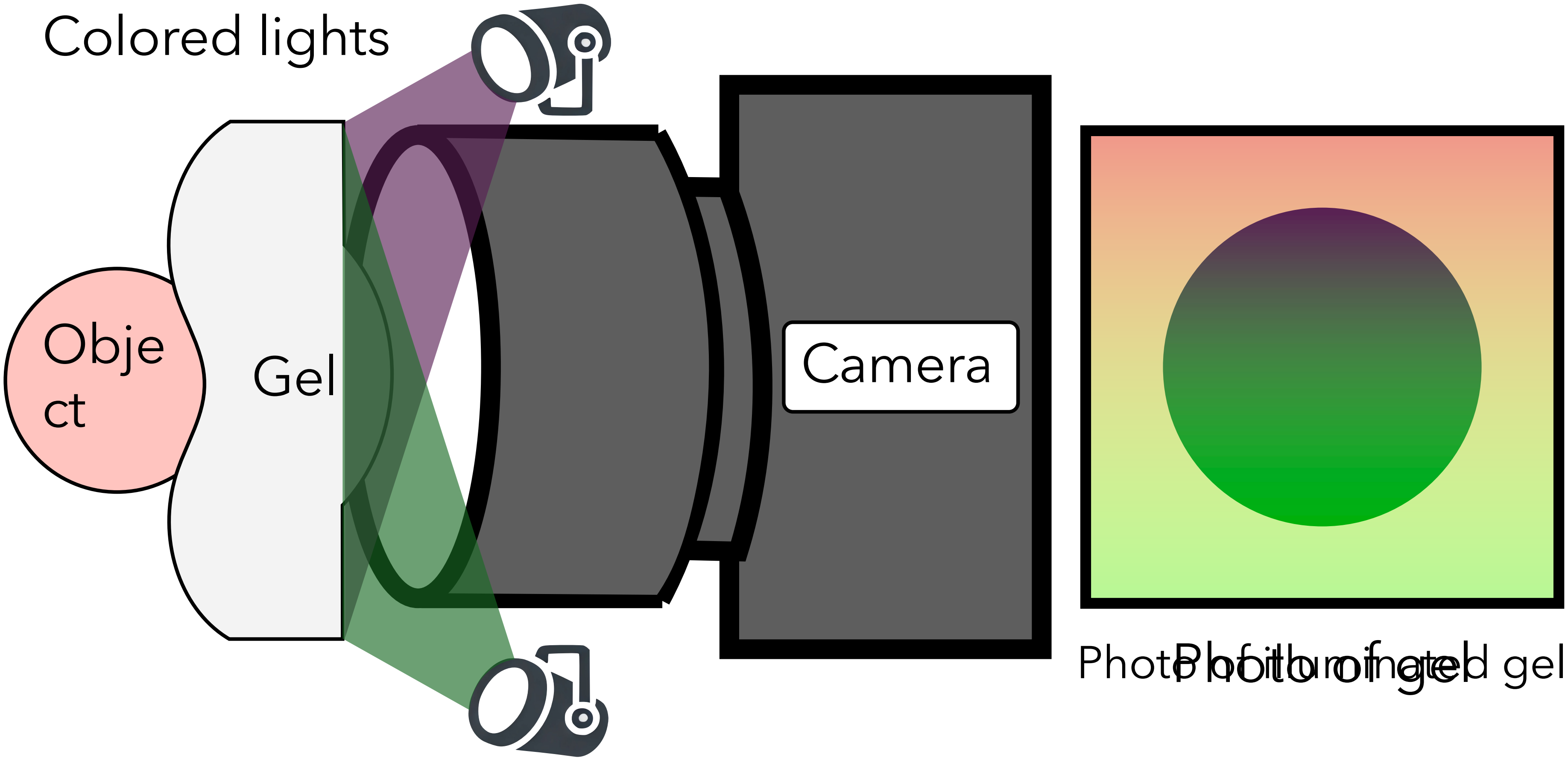
Source: https://www.smbom.com/news/13602

Optical Sensors



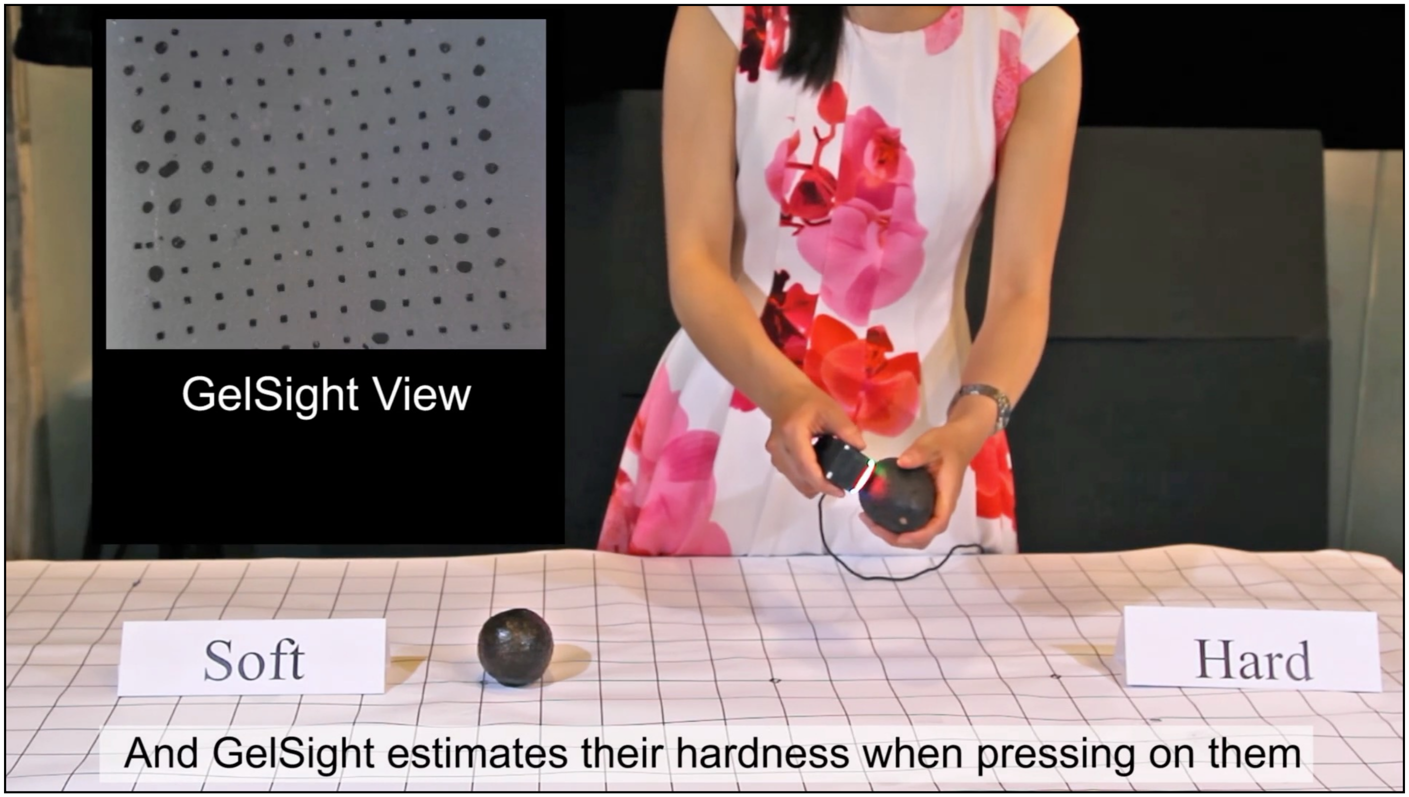
e.g., GelSight [Johnson and Adelson, 2009], [Yuan et al., 2017]

Optical Sensors



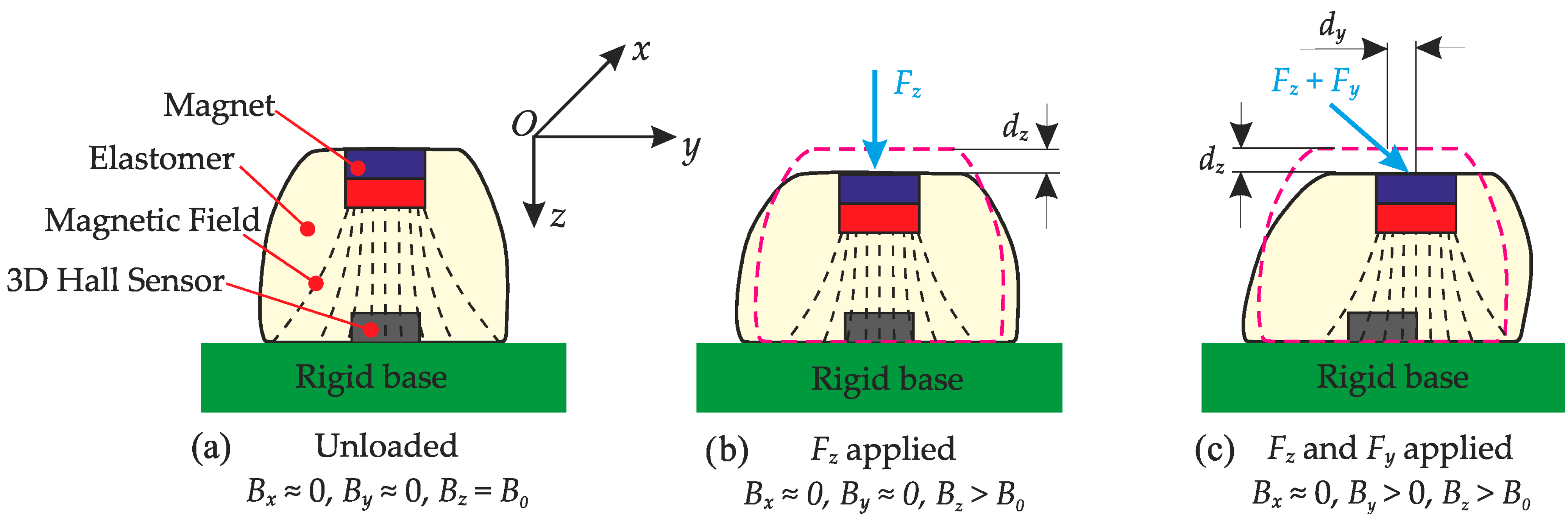
e.g., GelSight [Johnson and Adelson, 2009], [Yuan et al., 2017]

Optical Sensors



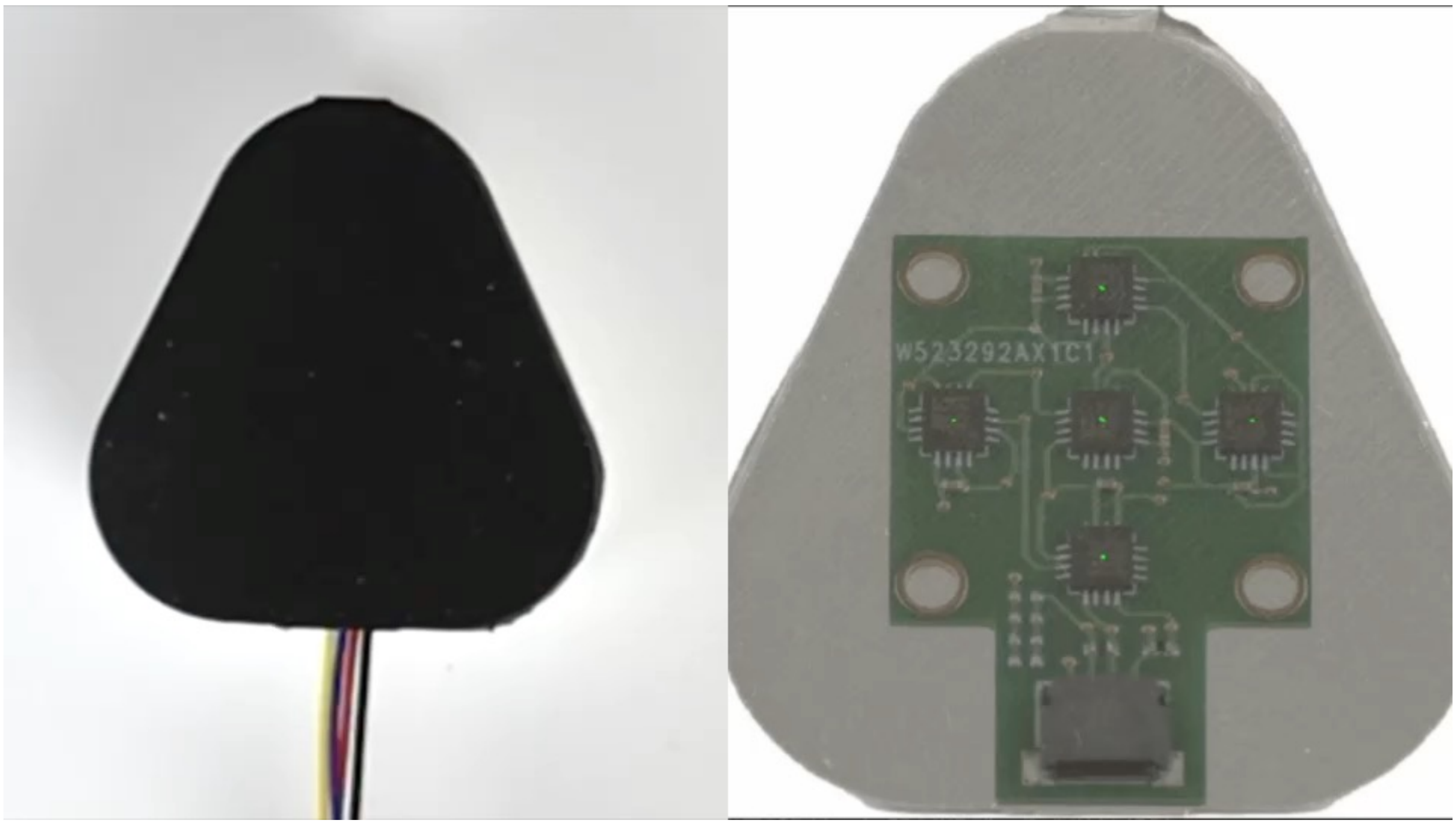
GelSight [Johnson and Adelson, 2009], [Yuan et al., 2017]
Example from [Yuan, Zhu, Owens, Srinivasan, Adelson, "Shape-independent Hardness Estimation…", ICRA 2017]

Magnetic Sensors



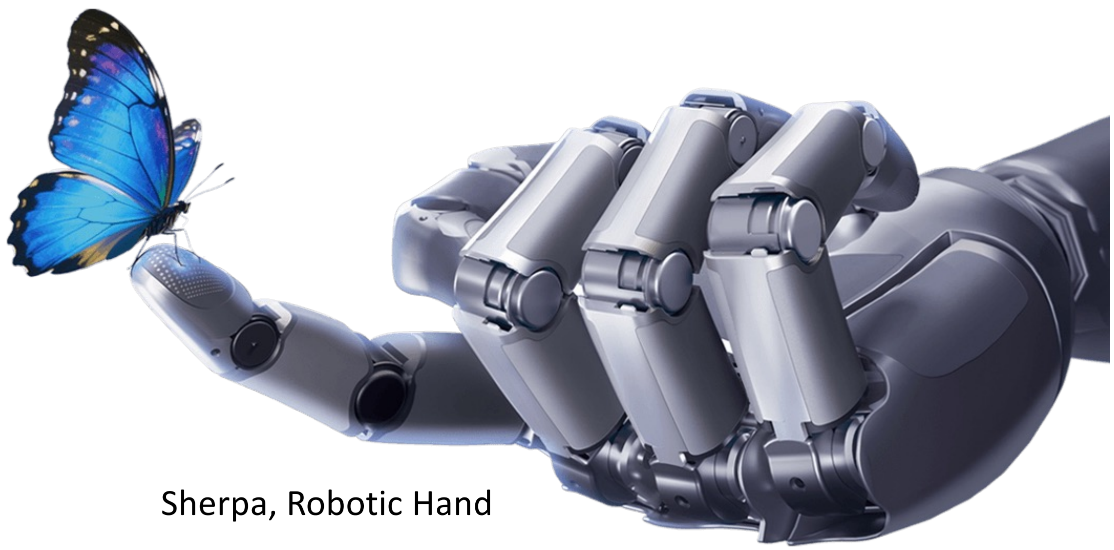
Design Methodology for Magnetic Field-Based Soft Tri-Axis Tactile Sensors, Wang et al, *Sensors*, 2016.

AnySkin: **Plug-and-play Skin Sensing for Robotic Touch**



Piezoelectric Sensors

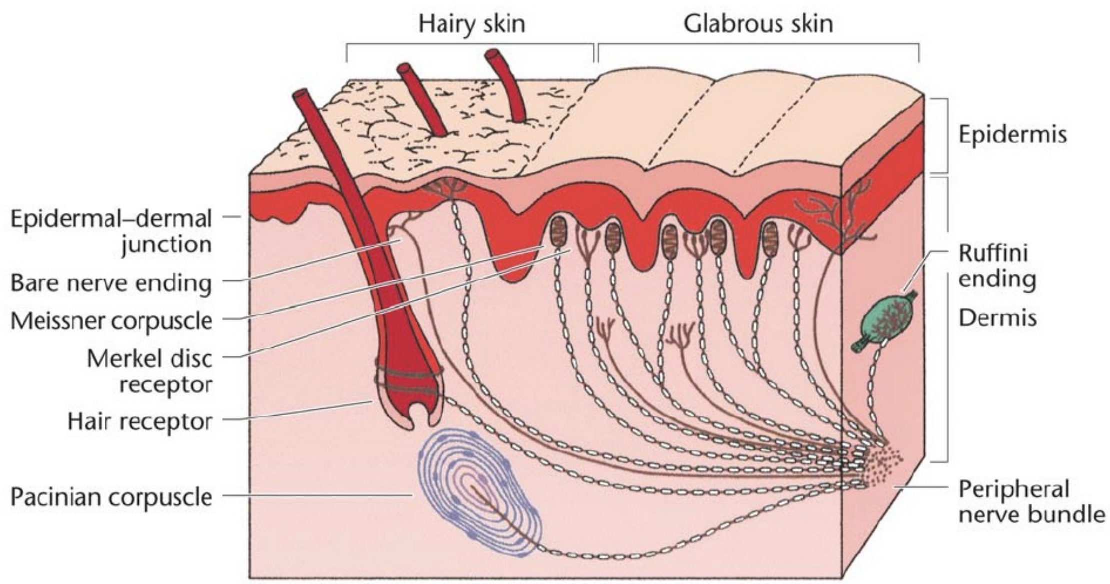
- They convert various physical forces, including pressure, vibration, and temperature, into electrical charges that can be measured.
- High sensitivity and durability, small size, and a wide frequency range.
- High cost (needs specialized electronics), susceptible to temperature.



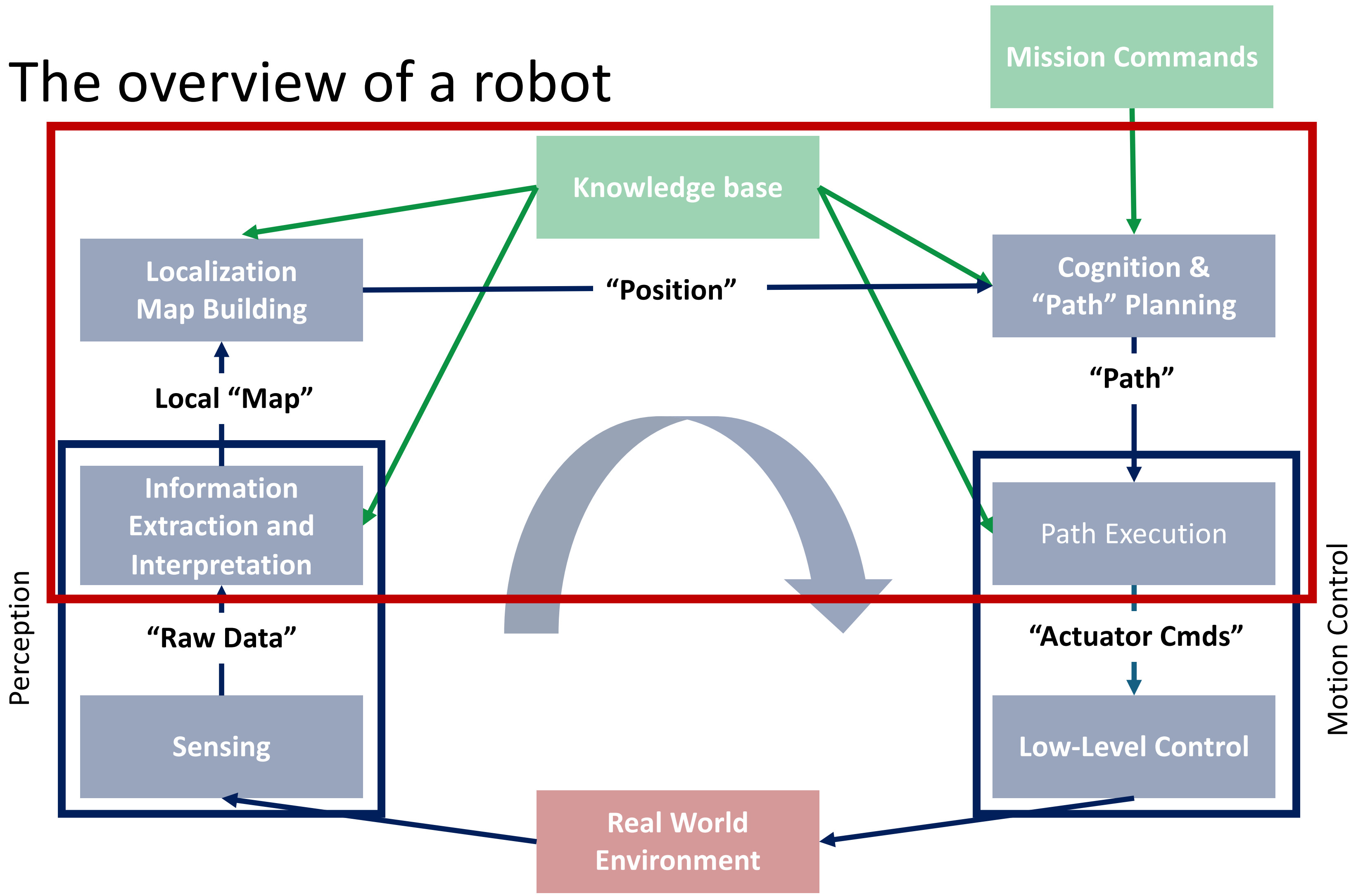
Sherpa, Robotic Hand

The Sense of Touch in Humans

- When the skin is contacted by an external stimulus, its surface is indented or stretched as the skin is flexible.
- The mechanical deformation is detected by mechanoreceptors that signal location of contact, force exerted, motion speed, and pressure.
- The closer to the surface, the more sensitive.
- Other receptors underneath the skin sense warmth, cold, pain, itching.



The overview of a robot



Questions?