Chapter Outline

Common Perception Sensors

Chapter 8 Multimodal Perception and Intelligent Decision-Making

8.1 Image Processing Fundamentals

8.2 Recognition, Tracking, Optical Flow

8.3 Localization, Mapping, and Planning

8.4 Common Perception Sensors

8.5 Coordinate Transformations

8.6 Fusion and Control Decision-Making

8.7 Tools and Framework Interfaces

8.8 Experiments and Real-World Transfer

8.9 Exercises and Classroom Discussion

5 Units in This Section

8.4.1 Vision Cameras
Textbook: Sec. 8.1.4.1

8.4.2 LiDAR
Light Detection and Ranging; Textbook: Sec. 8.1.4.2

8.4.3 Radar
Textbook: Secs. 8.1.4.3–8.1.4.5

8.4.4 Acoustic Sensors
Textbook: Secs. 8.1.4.6–8.1.4.8

8.4.5 Other Sensors
Textbook: Sec. 8.1.4.9

8.4.1 Vision Cameras: Sensor Overview

Sensor Outputs Determine What We Can Infer



Fig. 8.4 Common Multimodal Perception Sensors

- **Images:** Color, texture, and thermal features support target recognition and scene understanding.
- **Range and bearing:** Ranging and spatial scans support obstacle detection, localization, and mapping.
- **Self-motion and environment:** Motion, pressure, magnetic field, temperature, and humidity aid navigation and operating adjustments.

8.4.1 Vision Cameras: Imaging and Outputs

Color, Depth, and Pose Are Different Outputs



(a) Monocular RGB

SY011HD
Red, green, blue (RGB) pixels show appearance. Metric range needs extra cues.



(b) Depth Camera

D435
Color and stereo depth for 3D reconstruction and obstacle detection.



(c) Stereo Visual Odometry

T265
Own position and attitude. Localizes and tracks motion.

Fig. 8.4(a–c) Vision Cameras and Visual Odometry

8.4.1 Vision Cameras: RGB Cameras

Optical Imaging Preserves Appearance but Loses Range

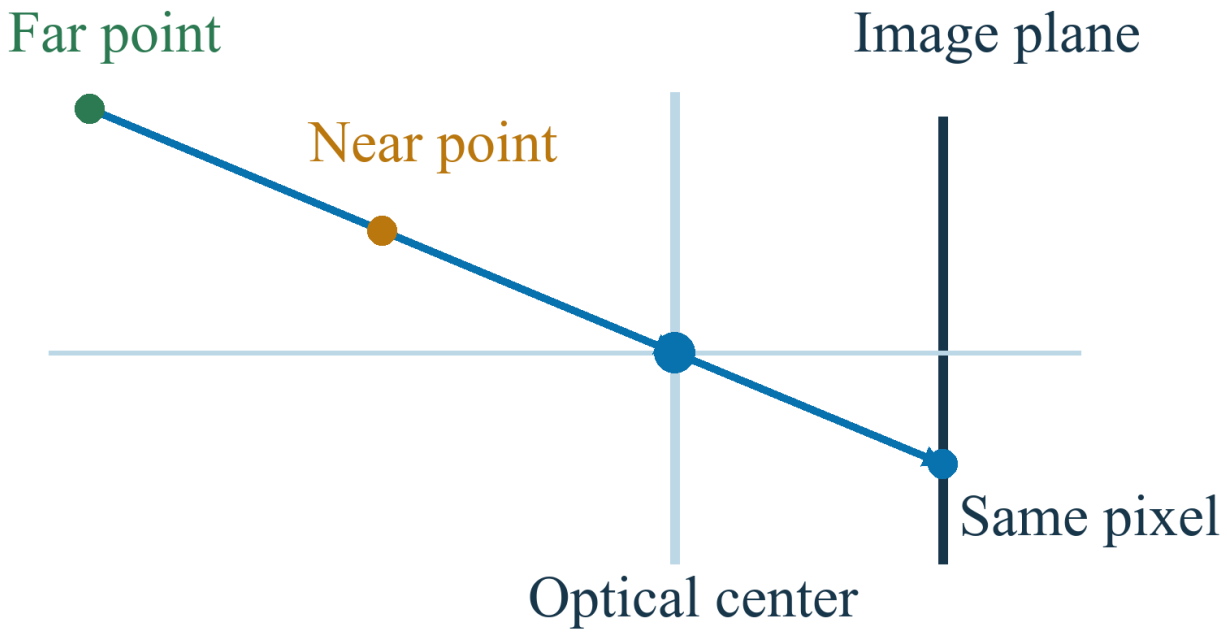


Diagram: Different Ranges Along One Imaging Ray

- **Imaging:** A lens focuses light. Complementary metal-oxide-semiconductor (CMOS) or charge-coupled device (CCD) chips digitize it.
- **Appearance:** Color, texture, and shape support classification, detection, tracking, and visual navigation.
- **Range:** One RGB image gives no direct metric depth. Known sizes or other sensors supply scale.

8.4.1 Vision Cameras: RGB-D Cameras

Depth Maps Depend on the Ranging Mechanism

Mechanism	Range Measurement	Outputs and Conditions
Structured Light	Use deformation of a known projected pattern and calibrated geometry.	Pixel depth; strong ambient light can wash out the pattern.
Time of Flight (ToF)	Emit light; use round-trip time or phase difference for range.	Range/depth; multipath and return strength affect results.
Active Stereo	Stereo matching gives disparity; infrared projection adds texture.	Examples: D435/D435i and Gemini 335/335L.

RGB-D combines color (RGB) and depth (D). After time and spatial alignment, the Point Cloud Library (PCL) supports 3D reconstruction, obstacle analysis, and interaction.

8.4.1 Vision Cameras: Stereo and Multicamera Systems

Disparity Connects Two Images to Depth

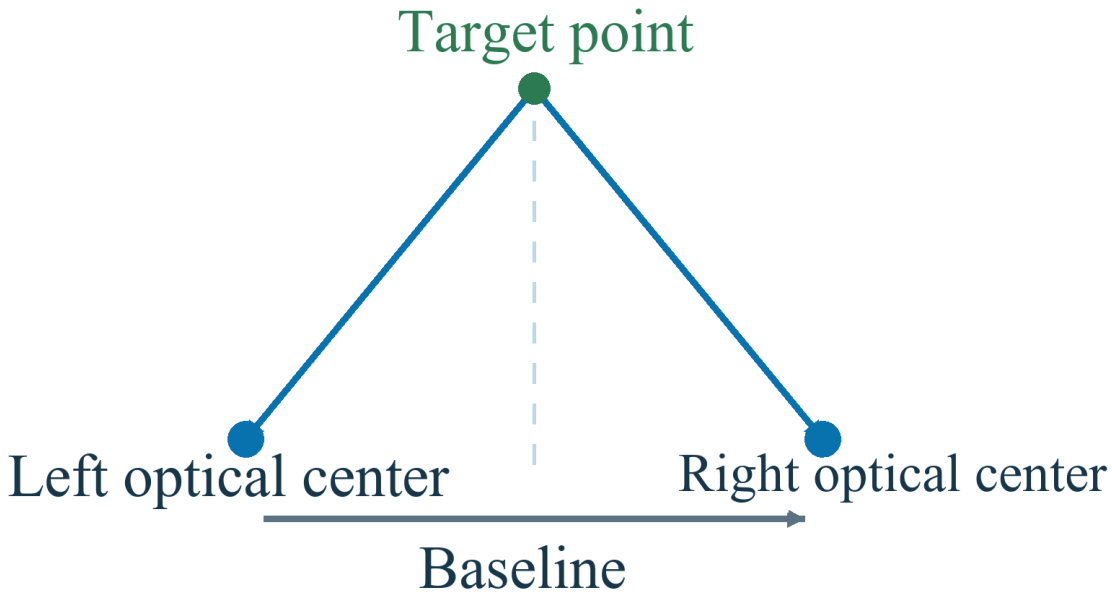


Diagram: Two Optical Centers and a Target

OAK-D: stereo. OAK-4P coverage depends on configuration. Passive stereo needs light/texture.

$$Z = \frac{fB}{d}$$

Z: optical-axis depth (m); f: focal length (px); B: baseline (m); d: disparity (px).

Disparity is the horizontal pixel offset after rectification. Longer baselines improve far-range resolution but add occlusion and mounting demands. Synchronization, calibration, and computing also matter.

8.4.1 Vision Cameras: Depth Example

Distant Targets Are More Sensitive to Disparity Error

Rectified stereo: $f = 800$ px, $B = 0.12$ m. Track the same feature over three frames.

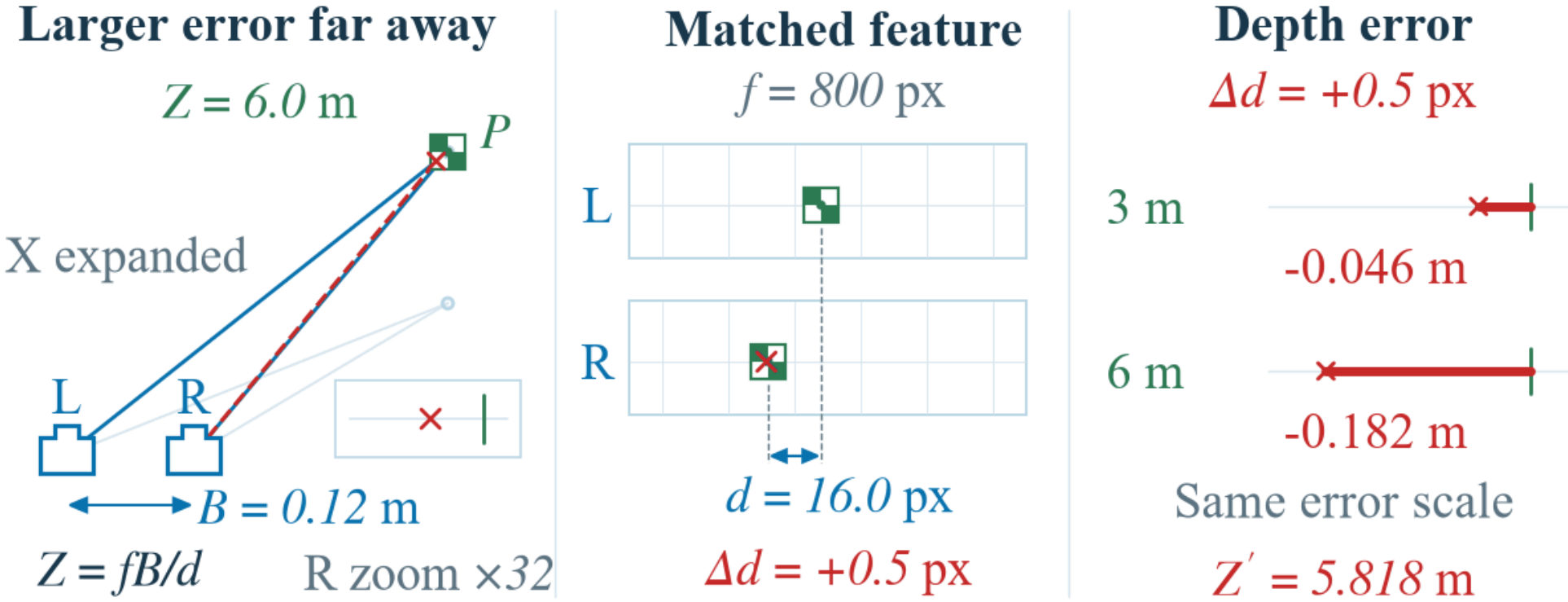
$$Z = \frac{800 \times 0.12}{d} = \frac{96}{d} \text{ m}$$

Frame	Disparity (px)	Depth (m)	Depth With +0.5 px	ΔDepth (m)
1	32	3.000	2.954 m	-0.046
2	24	4.000	3.918 m	-0.082
3	16	6.000	5.818 m	-0.182

Same +0.5 px bias: Larger depth error farther away. Increasing depth means relative recession.

8.4.1 Vision Cameras: Disparity and Depth Error

Why Does the Same Half-Pixel Error Matter More Farther Away?



Rectified-image disparity gives depth; the same +0.5 px error causes a larger depth error farther away.

8.4.1 Vision Cameras: Infrared and Thermal Imaging

Thermal Radiation Provides Nighttime Target Cues



Fig. 8.4(d) Thermal Camera

WIRIS Pro

Near-infrared reflection and thermal emission differ.

Thermal Radiation → **Infrared Optics** → **Detector Image**

- **Thermal imaging:** Radiation patterns support night navigation, search and rescue, fire prevention, and heat-source monitoring.
- **Temperature:** Emissivity, reflected background radiation, and calibration affect inference. Brightness alone is not true temperature.
- **Environment:** Some smoke and dust still allow target cues; rain and fog also weaken infrared propagation.

8.4.1 Vision Cameras: EO/IR Gimbal Pods

A Stabilized Gimbal Keeps Multiple Sensors on Target



Fig. 8.4(e) MOES-188

Electro-Optical (EO) → Color, Texture, and Contours

Thermal Infrared (IR) → Thermal Cues and Night Targets

Laser Range (Optional) → Target Range

Gimbal stabilization and tracking keep the line of sight on target. Gimbal rotation changes sensor orientation; localization also needs range, body attitude, and gimbal angles.

Exercises and Discussion

Choose an Imaging Method for the Task

Task	Sensor / Measurement	Conditions
Color/Class ID	RGB: Color and texture.	Light, exposure, and blur affect appearance.
Near-Range 3D Interaction	RGB-D: Color and depth.	Range limits and multipath affect valid pixels.
Outdoor Stereo	Stereo/multicamera: Disparity.	Baseline, texture, and occlusion affect scale.
Night Heat Search	Thermal camera: Thermal radiation.	Emissivity, rain, and fog affect thermal cues.
Distant Observation	EO/IR pod: Stabilized line of sight.	Gimbal attitude and ranging support localization.

Appearance, geometry, and thermal cues require complementary observations.

Chapter Outline

Common Perception Sensors

Chapter 8 Multimodal Perception and Intelligent Decision-Making

8.1 Image Processing Fundamentals

8.2 Recognition, Tracking, Optical Flow

8.3 Localization, Mapping, and Planning

8.4 Common Perception Sensors

8.5 Coordinate Transformations

8.6 Fusion and Control Decision-Making

8.7 Tools and Framework Interfaces

8.8 Experiments and Real-World Transfer

8.9 Exercises and Classroom Discussion

5 Units in This Section

8.4.1 Vision Cameras
Textbook: Sec. 8.1.4.1

8.4.2 LiDAR
Light Detection and Ranging; Textbook: Sec. 8.1.4.2

8.4.3 Radar
Textbook: Secs. 8.1.4.3–8.1.4.5

8.4.4 Acoustic Sensors
Textbook: Secs. 8.1.4.6–8.1.4.8

8.4.5 Other Sensors
Textbook: Sec. 8.1.4.9

8.4.2 LiDAR: 2D and 3D

Scan Planes and 3D Point Clouds Support Different Maps



Fig. 8.4(f) S2L



Fig. 8.4(g) Mid-360

2D scans: Range and bearing within one scan plane. Useful for planar mapping, ground robots, and planar obstacle avoidance at low altitude.

3D point clouds: Range, azimuth, and elevation form spatial points. Support terrain mapping, obstacle heights, relative localization, and local planning.

8.4.2 LiDAR: Time-of-Flight Ranging

Round-Trip Time Determines a Laser Beam's Range

Emit Near-Infrared Pulse → **Target Surface Reflects** → **Receive and Time Round Trip**

$$R = \frac{c\tau}{2}$$

$$R = \frac{3 \times 10^8 \times 100 \times 10^{-9}}{2} = 15\ \text{m}$$

R is target range (m); c is light speed (m/s); τ is round-trip time (s). This example uses $c = 3 \times 10^8$ m/s.

Range → **Add Scan Direction** → **3D Geometric Point**

8.4.2 LiDAR: Scanning and Maps

Unscanned Space Remains Unknown

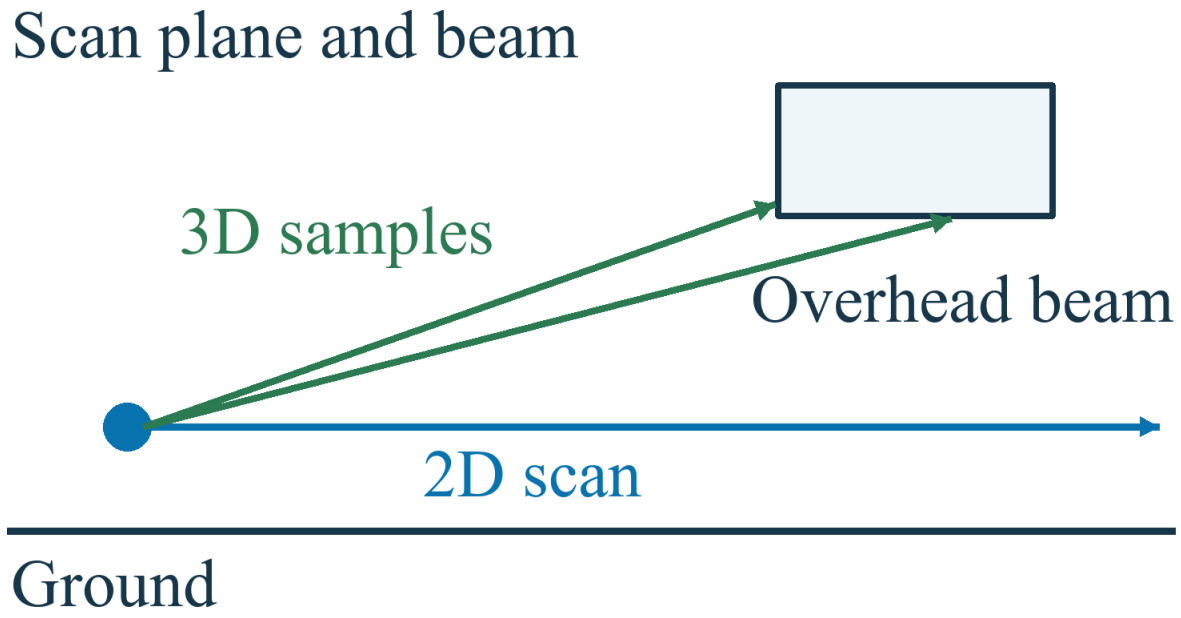


Diagram: Scan Plane and Overhead Beam

- **2D maps:** GMapping and HectorSLAM build planar maps; heights require vertical samples.
- **3D maps:** OctoMap represents occupancy; Voxblox uses a distance field to support planning.
- **Coverage/cost:** 3D also has occlusion and coverage limits. More points cost computation, storage, and data sharing.

8.4.2 LiDAR: Point-Cloud Example

Range and Direction Jointly Locate a Scan Point

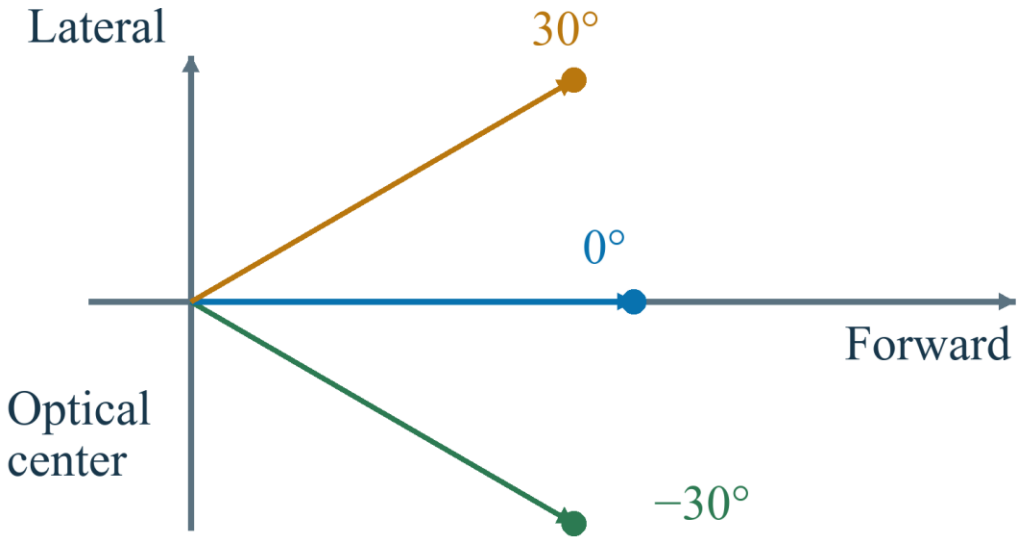


Diagram: Three Directions at 5 m

Sensor plane: $R = 5$ m; θ runs from forward toward positive lateral.

$$x = R\cos\theta$$

$$y = R\sin\theta$$

Scan Angle	Forward x (m)	Lateral y (m)
-30°	4.330	-2.500
0°	5.000	0.000
30°	4.330	2.500

8.4.2 LiDAR: Mutual Interference

Identifying Our Own Returns Reduces Ghost Points

Own Transmission → **Receive Mixed Returns** → **Match Own Signal**

Method	Identifying Own Returns	Design Conditions
Pulse Dithering	Randomize emission intervals to disrupt fixed time correlation.	Needs timing, detection, and multiframe criteria.
Distinct Signal Codes	Match the sensor's own code to reject external pulses.	Code design and the decoding receiver work together.
Frequency-Modulated Continuous Wave (FMCW)	Coherent detection compares frequency/phase with a local reference.	Strong selectivity; signal and hardware design govern effectiveness.

Exercises and Discussion

What Does a Missed Overhead Beam Mean?

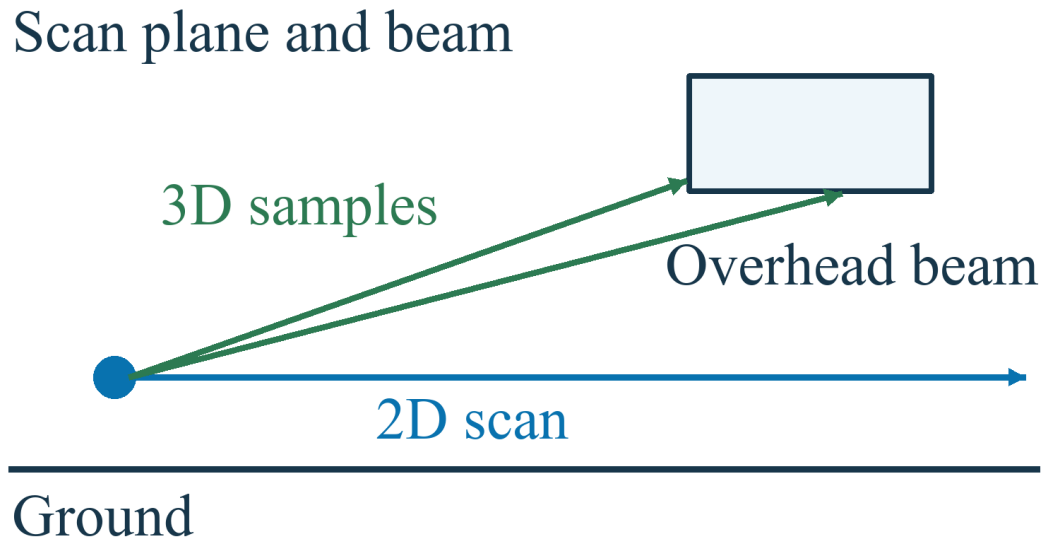


Diagram: A 2D Scan Passes Below the Beam

- **Question:** If a 2D scan has no forward return, can we conclude the flight path above has no overhead beam?
- **Answer:** No. The beam is above the scan plane and unobserved. Add vertical sampling or another observation.
- **Check:** Scan plane, field of view, occlusion, reflection, and interference explain missing or abnormal points.

A missing return does not directly imply traversable space.

Chapter Outline

Common Perception Sensors

Chapter 8 Multimodal Perception and Intelligent Decision-Making

8.1 Image Processing Fundamentals

8.2 Recognition, Tracking, Optical Flow

8.3 Localization, Mapping, and Planning

8.4 Common Perception Sensors

8.5 Coordinate Transformations

8.6 Fusion and Control Decision-Making

8.7 Tools and Framework Interfaces

8.8 Experiments and Real-World Transfer

8.9 Exercises and Classroom Discussion

5 Units in This Section

8.4.1 Vision Cameras
Textbook: Sec. 8.1.4.1

8.4.2 LiDAR
Light Detection and Ranging; Textbook: Sec. 8.1.4.2

8.4.3 Radar
Textbook: Secs. 8.1.4.3–8.1.4.5

8.4.4 Acoustic Sensors
Textbook: Secs. 8.1.4.6–8.1.4.8

8.4.5 Other Sensors
Textbook: Sec. 8.1.4.9

8.4.3 Radar: Types and Classification

Detection, Scanning, and Imaging Describe Different Capabilities



(h) Millimeter-Wave Radar
Phoenix 4D

Returns give range, radial velocity, and direction.



(i) Active Phased Array
DJI RD241608

Element phase/delay steers the beam electronically.



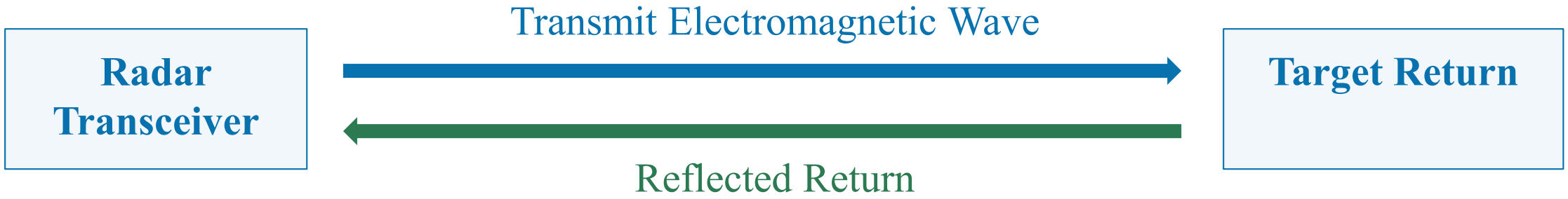
(j) Synthetic Aperture Radar
SAR RK-12

Focus coherent returns collected along a moving trajectory.

Fig. 8.4(h–j) Radar Hardware and Its Different Capabilities

8.4.3 Radar: Propagation Delay and Doppler

Delay Measures Range; Frequency Shift Measures Radial Velocity



$$R = \frac{c\tau}{2}$$

R is range; τ is round-trip time; c is light speed. Delay reflects propagation distance.

$$f_D = \frac{2v_r}{\lambda}$$

f_D is Doppler shift (Hz); v_r is radial velocity toward the radar (m/s); λ is wavelength (m).

8.4.3 Radar: Frequency-Modulated Continuous Wave

The Transmit–Return Frequency Difference Encodes Range

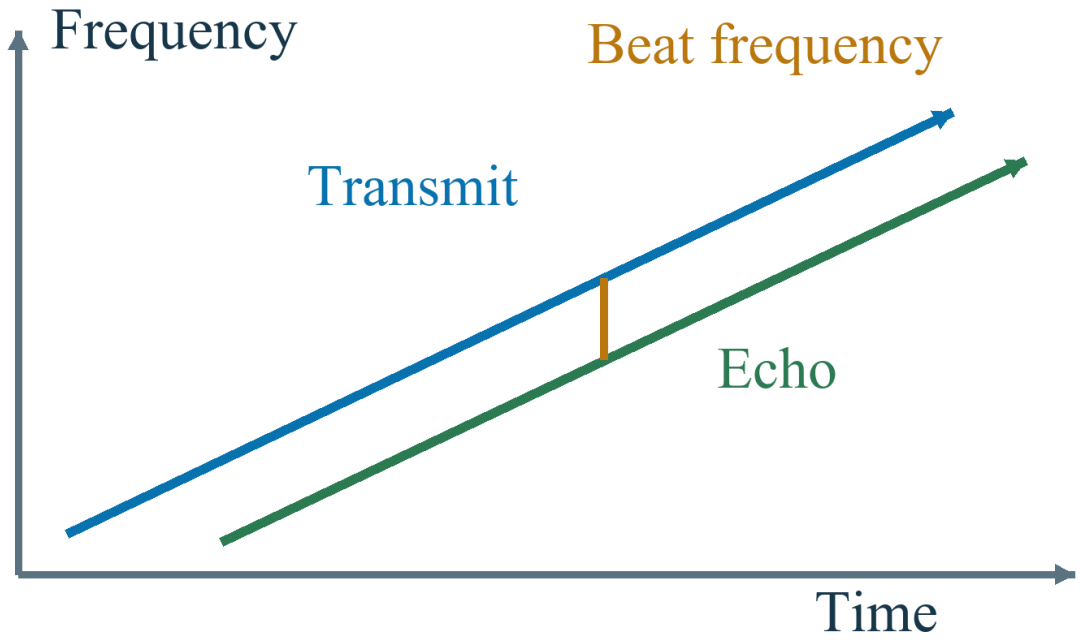


Diagram: Linear Chirp and Delayed Return

$$f_b = S\tau$$

$$R = \frac{cf_b}{2S}$$

S is the frequency ramp slope (Hz/s); f_b is beat frequency (Hz); τ is round-trip delay (s). This single-target ranging model neglects velocity coupling.

Repeated-chirp phase gives radial velocity; receive-array spatial phase gives direction.

8.4.3 Radar: Airborne Millimeter-Wave Sensing

Range and Motion Cues Remain Useful in Reduced Visibility

24 GHz is centimeter-wave; engineering usage groups 24 and 77 GHz as millimeter-wave radar.

$$\lambda_{24} = \frac{3 \times 10^8}{24 \times 10^9} = 12.5\ \text{mm} \qquad \lambda_{77} = \frac{3 \times 10^8}{77 \times 10^9} \approx 3.90\ \text{mm}$$

- **Altitude and obstacles:** Ground range supports height observations; obstacle range and radial velocity support dynamic-scene sensing.
- **Light and weather:** No visible light is needed. Useful returns can remain in some rain or fog, but atmospheric loss, wet surfaces, and clutter still matter.
- **Targets and motion:** Scattering strength, aspect, antenna coverage, and velocity range determine detectability. Radar complements cameras and LiDAR.

8.4.3 Radar: Ground Surveillance

Motion Cues Help Separate Targets From Ground Clutter

On an unmanned ground vehicle (UGV) or fixed site, X/Ku and Ka/W bands trade propagation loss and antenna size against mission range.

Static Ground Clutter	Whole-Target Motion	Component Motion
Trees and rocks form clutter; time/frequency differences support moving-target processing.	Range and radial velocity provide cues; radar cross-section (RCS) describes return strength.	Micro-Doppler describes moving parts; models use it to support classification.

Sector/full-azimuth scans find targets and cue electro-optical viewing; configuration sets coverage.

8.4.3 Radar: Active Phased Arrays

Element Phase Differences Steer the Beam

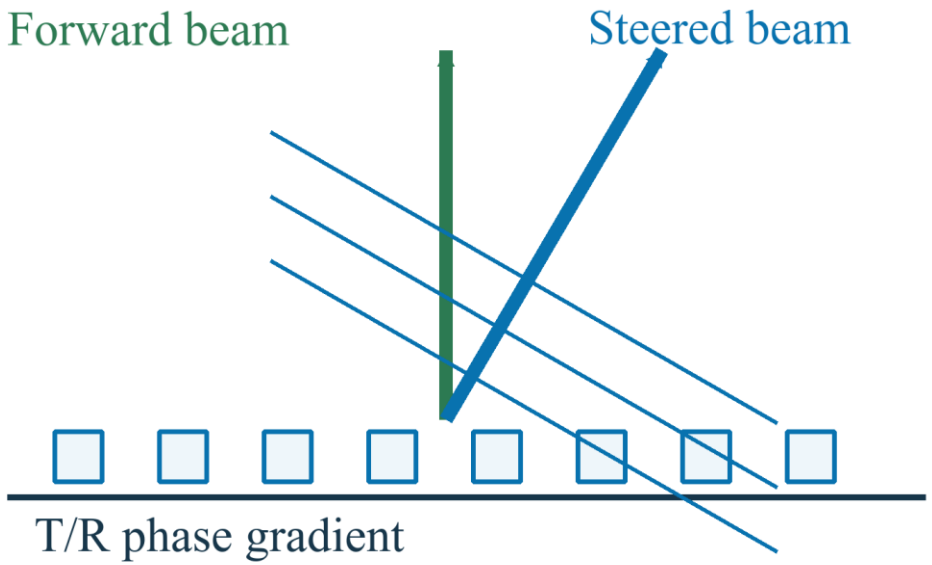


Diagram: Electronic Beam Steering

$$|\Delta\varphi| = \frac{2\pi d|\sin\theta|}{\lambda}$$

d is element spacing; λ is wavelength; θ is angle from the array normal; $\Delta\varphi$ is the phase step between adjacent elements (rad).

$$d = \frac{\lambda}{2},\ \theta = 30^\circ \Rightarrow |\Delta\varphi| = 90^\circ$$

Active phased array radar (APAR) uses independent transmit/receive modules for electronic scanning and multitarget scheduling.
Coverage, resolution, and redundancy depend on configuration.

8.4.3 Radar: Low-Altitude Terrain Following

Terrain Following Needs Range and Beam Direction

T40/T20P crop-spraying tasks: Forward obstacle observations support avoidance; ground range supports altitude hold and following uneven terrain.

Radar Measures Slant Range → **Combine With Beam Direction** → **Vertical Ground Clearance**

$$h = R\cos\beta \qquad R = 10\ \text{m},\ \beta = 30^\circ \Rightarrow h \approx 8.66\ \text{m}$$

R is slant range; β is measured from vertically downward. Ground is approximately level here. T40 horizontal obstacle sensing covers 360° azimuth and ±45° elevation; other channels and mounting also affect coverage.

T40 Radar Fields of View: https://www.dji.com/t40/specs

8.4.3 Radar: Synthetic Apertures

Platform Motion Supplies Multiple Coherent Viewpoints

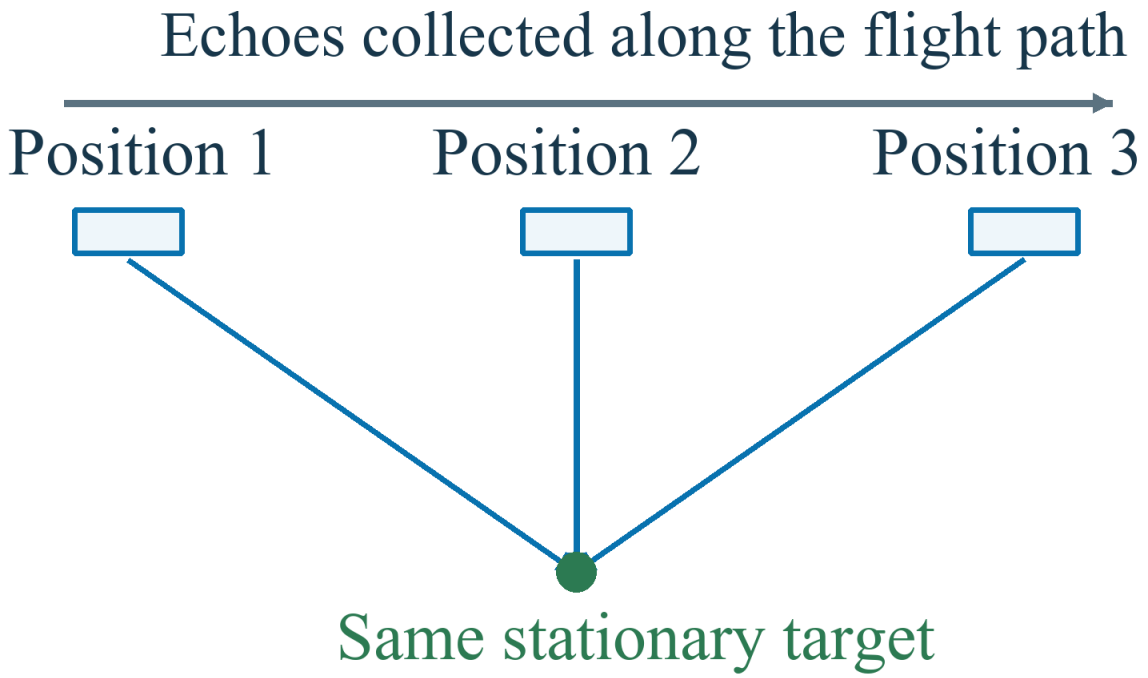


Diagram: Repeated Views of One Ground Feature

- **Synthetic aperture radar (SAR):** Along a moving path, record one ground feature's amplitude and phase; compensate phase history to focus it.
- **Conditions:** Known motion, coherent returns, and motion compensation. Start with stationary terrain.

$$L = vT_{\text{int}} = 30 \times 2 = 60\ \text{m}$$

L: trajectory length (m); v: speed (m/s); T_{int}: coherent integration time (s).

8.4.3 Radar: Remote-Sensing Applications

Scattering Images and 3D Terrain Need Different Processing

Focus SAR Returns	→ Amplitude and Phase Image	→ Interpret Terrain/Scattering
Two Images of the Same Area	→ Register and Compare	→ Area Changes and Disaster Assessment
Two Views/Repeated Phase	→ Interferometric Geometry	→ Elevation and 3D Terrain

Scattering/change cues monitor buildings, infrastructure, and vegetation. Long-endurance fixed-wing aircraft or high-altitude pseudo-satellites (HAPS) supply stable tracks for wide-area imaging, assessment, and aided localization.

8.4.3 Radar: Range and Velocity Example

Beat Frequency and Doppler Give Complementary Results

Single target: $S = 20\ \text{MHz}/\mu\text{s} = 2 \times 10^{13}\ \text{Hz/s}$; carrier 77 GHz.
Beat frequency $f_b = 4$ MHz; Doppler $f_D = 1$ kHz. Range/velocity are decoupled; approaching is positive.

$$R = \frac{cf_b}{2S} = 30\ \text{m} \qquad v_r = \frac{f_D\lambda}{2} \approx 1.95\ \text{m/s}$$

Time (s)	Range (m), Model	Beat (MHz)	Interpretation
0.0	30.000	4.000	Initial single target.
0.1	29.805	3.974	Radial approach: Range falls.
0.2	29.610	3.948	Lower beat agrees with approach.

8.4.3 Radar: Comparing Methods

Combine Radar Capabilities Around the Measurement Task

Capability	Measurement and Processing	Tasks and Main Conditions
Return Measurements	Delay/beat → range. Doppler → radial velocity.	Avoidance, altitude, distant detection. Bandwidth, update rate, scattering matter.
Active Phased Array	Module phase/delay control rapidly selects beam direction.	Multiple targets, terrain following. Array, coverage, mounting matter.
Synthetic Aperture	Record amplitude/phase along a path; compensate motion and focus coherently.	Terrain/area/disaster monitoring. Trajectory and coherence matter.

FMCW waveforms, phased array scans, and SAR imaging can combine; they are not exclusive levels.

Rapid response and wide-area imaging require different observation processes.

Exercises and Discussion

Choose Radar Observations for the Task

Task	Suggested Observations	Explanation and Checks
Dynamic Avoidance	Rapid range/direction updates with radial velocity.	Check forward coverage, scattering, velocity range, and update delay.
Uneven Fields	Ground range and directed sensing; a suitable phased array is an option.	Use slant range and beam direction; check height channels and obstacles.
Disaster Survey	SAR scattering and time comparisons; 3D elevation needs extra geometry.	Check trajectory/coherence/registration; separate image and elevation changes.

Chapter Outline

Common Perception Sensors

Chapter 8 Multimodal Perception and Intelligent Decision-Making

8.1 Image Processing Fundamentals

8.2 Recognition, Tracking, Optical Flow

8.3 Localization, Mapping, and Planning

8.4 Common Perception Sensors

8.5 Coordinate Transformations

8.6 Fusion and Control Decision-Making

8.7 Tools and Framework Interfaces

8.8 Experiments and Real-World Transfer

8.9 Exercises and Classroom Discussion

5 Units in This Section

8.4.1 Vision Cameras
Textbook: Sec. 8.1.4.1

8.4.2 LiDAR
Light Detection and Ranging; Textbook: Sec. 8.1.4.2

8.4.3 Radar
Textbook: Secs. 8.1.4.3–8.1.4.5

8.4.4 Acoustic Sensors
Textbook: Secs. 8.1.4.6–8.1.4.8

8.4.5 Other Sensors
Textbook: Sec. 8.1.4.9

8.4.4 Acoustic Sensors: Media and Signals

Ultrasound, Sonar, and Audio Use Different Observations

Sensor	Propagation and Measurement	Outputs and Scenarios
Ultrasonic Ranging in Air	Emit pulses above 20 kHz; range from round-trip delay.	Nearby obstacles/ground range. Air at 20°C: ≈343 m/s.
Underwater Sonar	Actively emit and measure echoes, or passively receive underwater sources.	Water: Range, bearing, imagery. Sound speed ≈1500 m/s.
Environmental Audio	Microphones sample sound pressure; arrays compare arrival times.	Source direction, events, semantics. Passive: No sound emitted.

Ultrasonic = frequency; sonar = use. High-frequency underwater sonar can use ultrasound.

8.4.4 Acoustic Sensors: Ultrasonic Ranging in Air

Pulse Echoes Suit Short-Range, Low-Speed Sensing

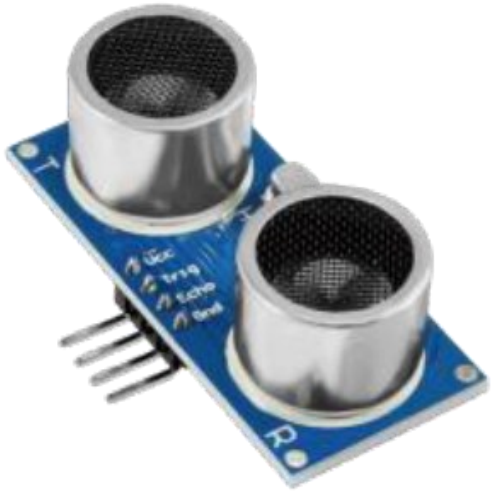


Fig. 8.4(k) HC-SR04

Emit 40 kHz Pulse → **Receive Target Echo**

$$R = \frac{c_a \tau}{2}$$

c_a is sound speed in air; τ is round-trip time. Residual vibration after transmission creates a near blind zone. Tilt, absorption, multipath, and noise alter echoes.

Uses: Parking blind spots, ground proximity in landing, indoor avoidance, hovering, and fine low-altitude motion; suitable for low speeds and short response distances.

8.4.4 Acoustic Sensors: Temperature and Range Example

Sound Speed Changes the Round-Trip Ranging Scale

A stationary planar target is 2 m away. Dry-air approximation; enter temperature in °C.

$$c_a \approx (331 + 0.6T_C)\ \text{m/s}$$

$$\tau_{20} = \frac{2 \times 2}{343} \approx 11.66\ \text{ms}$$

Temp.	Actual Sound Speed	Round Trip for 2 m	Estimate Using 343 m/s
0°C	331 m/s	12.08 ms	2.073 m: Overestimate
20°C	343 m/s	11.66 ms	2.000 m: Correct
40°C	355 m/s	11.27 ms	1.932 m: Underestimate

Temperature corrects scale; wind, angle, absorption, and multipath cause separate echo errors.

8.4.4 Acoustic Sensors: Active and Passive Sonar

Active Sonar Measures Echoes; Passive Sonar Listens

Transmit Sound → **Underwater Target Reflects** → **Receive and Analyze Echo**

Mode	Measurement and Outputs	Operating Conditions
Active Sonar	Known transmit time gives slant range; arrays give bearing. Intensity/frequency shift add cues.	Sounding, obstacles, imaging. Medium affects propagation and reflection.
Passive Sonar	Listen to target/environmental sound. Array gives bearing and acoustic features.	Quiet listening; range needs multiple sites, motion, or near-field geometry.

Water sound speed ≈1500 m/s; temperature, salinity, and pressure affect scale and propagation paths.

8.4.4 Acoustic Sensors: Underwater Imaging and Navigation

Depth-Sounding Beams and Side-Scan Images Differ

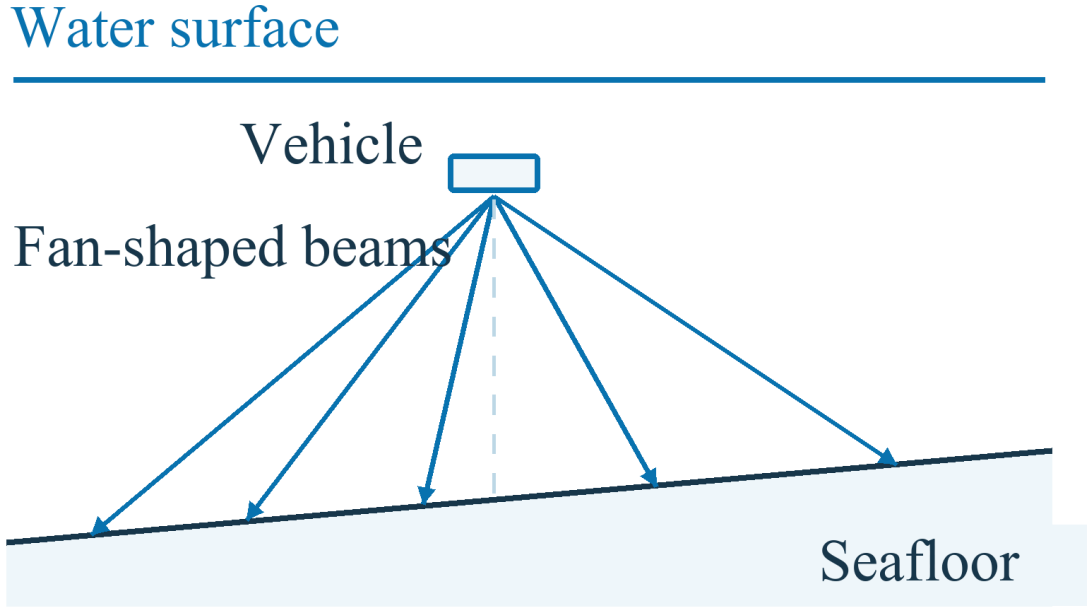


Diagram: Fan Beams Measure the Seabed Profile

- **Multibeam:** Direction and slant range recover seabed geometry. Pose and sound-speed correction affect 3D maps.
- **Side scan:** Intensity shows target/seabed texture. Low frequency favors range; high frequency gives finer detail but more loss.

- Unmanned underwater vehicles (UUVs) and remotely operated vehicles (ROVs) use sonar for navigation, avoidance, rescue, and facility monitoring.

8.4.4 Acoustic Sensors: Underwater Depth Example

Project Slant Range Onto the Vertical Direction

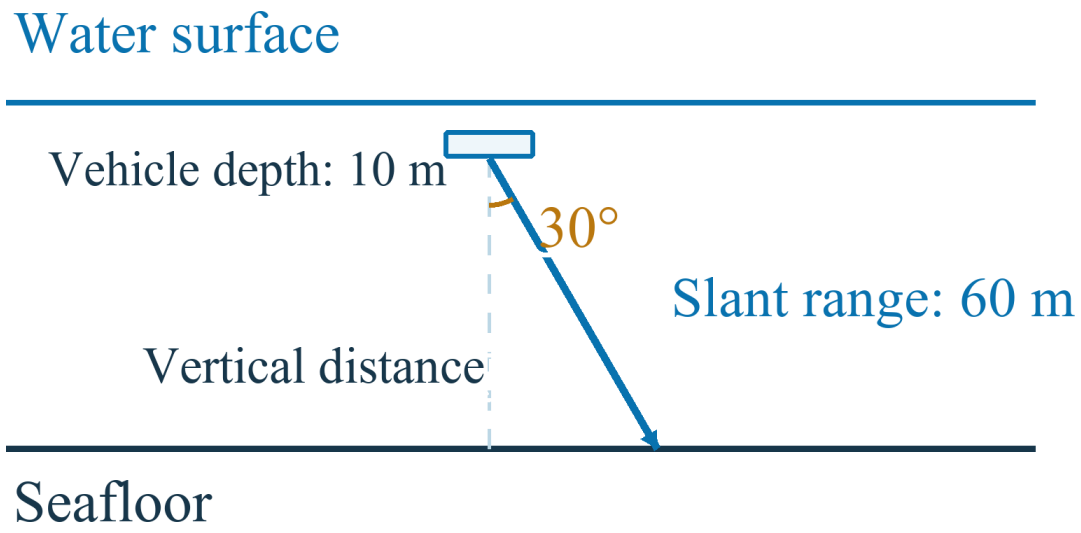


Diagram: Slant Range and Depth References

$$R = \frac{1500 \times 0.080}{2} = 60\ \text{m}$$

$$h = R\cos 30^\circ \approx 51.96\ \text{m}$$

$$H = 10 + h \approx 61.96\ \text{m}$$

R is slant range; h is vertical distance below the vehicle; H is surface-referenced water depth. Vehicle depth is 10 m. Assume constant sound speed and straight rays.

Range, direction, and own depth locate the seabed; sound-speed and pose errors affect results.

8.4.4 Acoustic Sensors: Microphone Arrays

Arrival-Time Differences Indicate Sound Direction

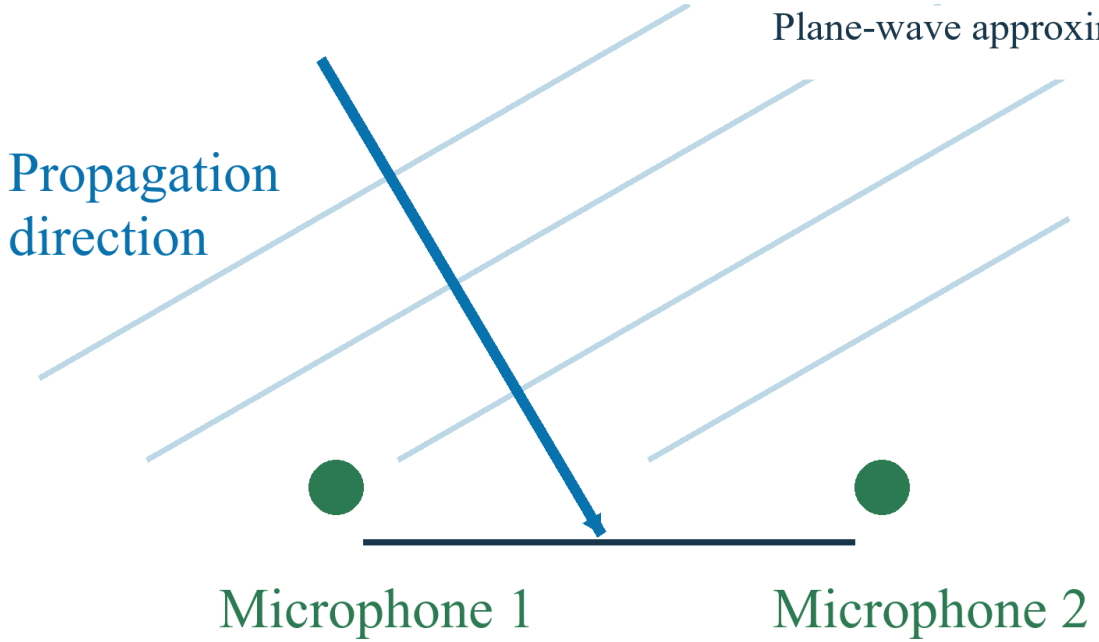


Diagram: Wave Arrival at Two Microphones

Arrays may be linear, circular, or spherical.

$$\tau = t_2 - t_1 = \frac{d \sin\theta}{c_a}$$

Time difference of arrival (TDOA): τ, channel arrival delay; d, microphone spacing; θ, angle from array normal; c_a, air sound speed.

Far field gives bearing, with two-microphone ambiguity. Range needs near-field geometry, multiple sites, or motion.

8.4.4 Acoustic Sensors: Bearing Example

Synchronization Error Directly Changes Estimated Bearing

Given $d = 0.20$ m, $c_a = 343$ m/s; one far-field source.
Microphone 2 is later by $\tau = 0.2915$ ms. Added channel delay changes the bearing.

$$\theta = \arcsin\left(\frac{c_a \tau}{d}\right) \approx 30^\circ \qquad |\tau| \le \frac{d}{c_a} \approx 0.583 \text{ ms}$$

Sampling Case	Delay (ms)	Bearing/Check	Interpretation
Synchronized	0.2915 ms	About 30.0°	Valid geometric bearing.
Extra 50 μs Delay	0.3415 ms	About 35.9°	About 5.9° bias.
Abnormal Peak	0.7000 ms	Physically Invalid	Check timing, reverberation, and peak selection.

8.4.4 Acoustic Sensors: Localization and Semantics

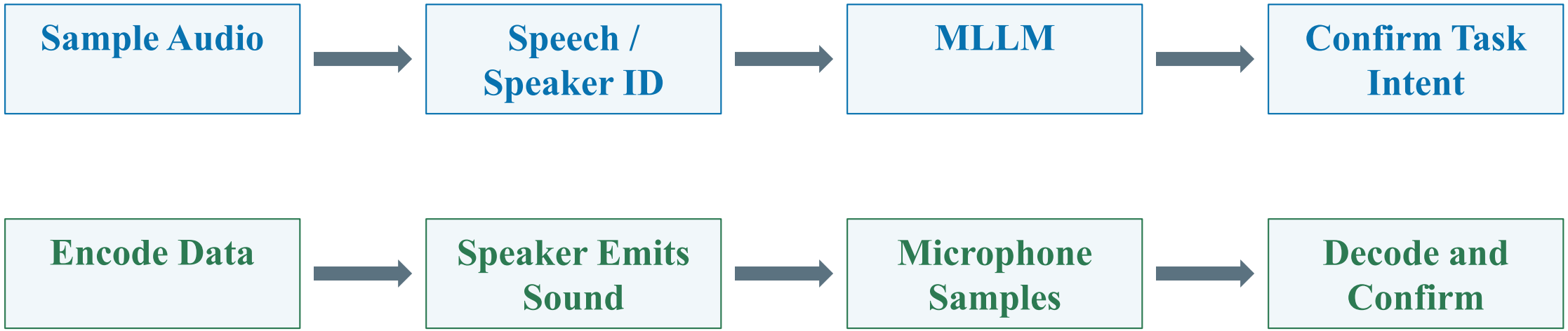
Localization, Separation, and Recognition Are Different Tasks

Sound Source Localization (SSL)	Beamforming and Source Separation	Event and Semantic Recognition
Timing/spatial cues give bearing and a continuous source track.	Delay-and-sum emphasizes direction; blind separation splits sources.	Models identify sirens, abnormal noises, speech, and meaning.

- **Occluded sirens:** Sound may still propagate around buildings and provide early cues. Reflections affect bearing estimates.
- **Search and tracking:** Calls or abnormal sounds guide searches or cameras. Rotor self-noise and reverberation can mask the source.
- **Environmental semantics:** Machine noises and tire friction can aid fault or road-state analysis; conclusions depend on models and operating conditions.

8.4.4 Acoustic Sensors: Interaction and Audio Links

Semantic Interaction and Acoustic Links Need Added Modules



A multimodal large language model (MLLM) combines audio, images, and task context. Encoded audio or biomimetic sounds can support experimental auxiliary links under interference, limited by propagation, noise, bandwidth, and confirmation.

Open-Source Encoded-Audio Communication: https://github.com/ggerganov/ggwave

Exercises and Discussion

Combine Acoustic Sensors by Medium and Task

Scenario	Suggested Observations	Interpretation and Limits
Windy Landing	Ultrasound: Near-ground range; add vision, LiDAR, or suitable radar.	Wind, noise, angle, and absorption can invalidate ultrasound.
UUV in Turbid Water	Active sonar: Obstacle/seabed slant range. Pressure: Vehicle depth from surface.	References differ. Sound speed/pose affect geometry; turbid water limits cameras.
Occluded Siren	Passive array: Recognize event, estimate arrival bearing, cue observation.	Reflections/self-noise affect bearing. One far-field delay gives no direct range.

Chapter Outline

Common Perception Sensors

Chapter 8 Multimodal Perception and Intelligent Decision-Making

8.1 Image Processing Fundamentals

8.2 Recognition, Tracking, Optical Flow

8.3 Localization, Mapping, and Planning

8.4 Common Perception Sensors

8.5 Coordinate Transformations

8.6 Fusion and Control Decision-Making

8.7 Tools and Framework Interfaces

8.8 Experiments and Real-World Transfer

8.9 Exercises and Classroom Discussion

5 Units in This Section

8.4.1 Vision Cameras
Textbook: Sec. 8.1.4.1

8.4.2 LiDAR
Light Detection and Ranging; Textbook: Sec. 8.1.4.2

8.4.3 Radar
Textbook: Secs. 8.1.4.3–8.1.4.5

8.4.4 Acoustic Sensors
Textbook: Secs. 8.1.4.6–8.1.4.8

8.4.5 Other Sensors
Textbook: Sec. 8.1.4.9

8.4.5 Other Sensors: Inertial Measurement Units

Inertial Measurements Provide High-Rate Self-Motion Cues

An inertial measurement unit (IMU) uses three-axis accelerometers/gyroscopes to measure specific force/angular velocity. An inertial navigation system (INS) estimates motion from them.

Accelerometer: Specific Force → **Handle Attitude and Gravity** → **Motion Acceleration, Velocity, and Position**

Gyroscope: Angular Velocity → **Integrate and Correct Errors** → **Estimate Attitude**

$$\mathbf{f} = \mathbf{a} - \mathbf{g}$$

Idealized, same frame: $\mathbf{f}$ is specific force, $\mathbf{a}$ is motion acceleration, $\mathbf{g}$ is gravity. At rest, specific-force magnitude $\approx 9.81\ \mathrm{m/s^2}$.

8.4.5 Other Sensors: Inertial Drift Example

Constant Bias Makes Position Error Grow Quadratically

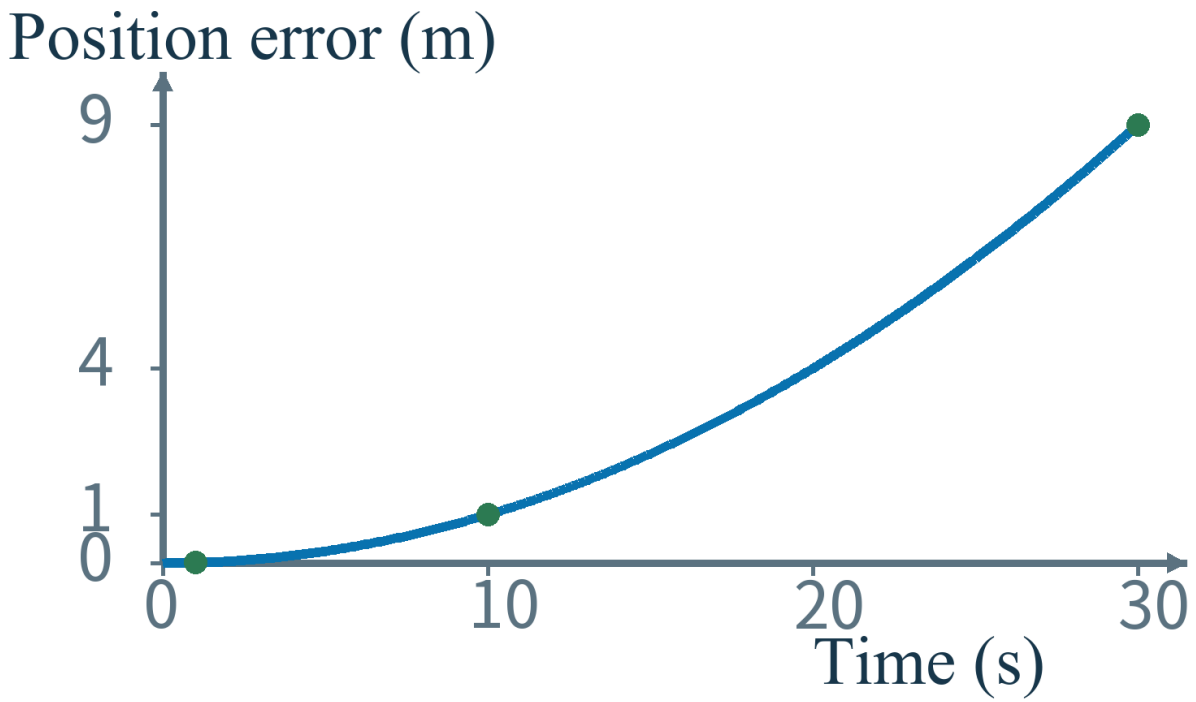


Diagram: Theoretical Position Error From Constant Bias

Stationary axis: $b_a = 0.02\ \mathrm{m/s^2}$.
$T = 10\ \mathrm{s}$; initial errors zero.
No attitude error or random noise.

$$\Delta v = b_a T = 0.20\ \mathrm{m/s}$$

$$\Delta p = \frac{1}{2} b_a T^2 = 1.00\ \mathrm{m}$$

At 1/10/30 s: 0.01/1/9 m.
External data limit long-term drift.

Global navigation satellite system (GNSS), vision, or LiDAR can constrain inertial drift.

8.4.5 Other Sensors: Barometric Altitude

Reference Pressure Converts Pressure Change Into Height

$$\Delta h = \frac{R_a T}{g} \ln\left(\frac{p_0}{p}\right)$$

Isothermal dry air: $R_a = 287\ \mathrm{J/(kg \cdot K)}$ is the specific gas constant; $T = 293\ \mathrm{K}$; $g = 9.81\ \mathrm{m/s^2}$. p_0 is reference pressure and p is current pressure.

Input or Result	Value and Meaning	Practical Conditions
Pressure Pair	101325 / 101205 Pa	Pressure falls; relative climb.
Relative Height	10.16 m	Temperature, airflow, and reference pressure matter.
Fusion Use	GNSS/vision add vertical position.	Terrain following needs ground range or a map.

8.4.5 Other Sensors: Water-Pressure Depth

Surface Pressure and Water Density Determine Depth

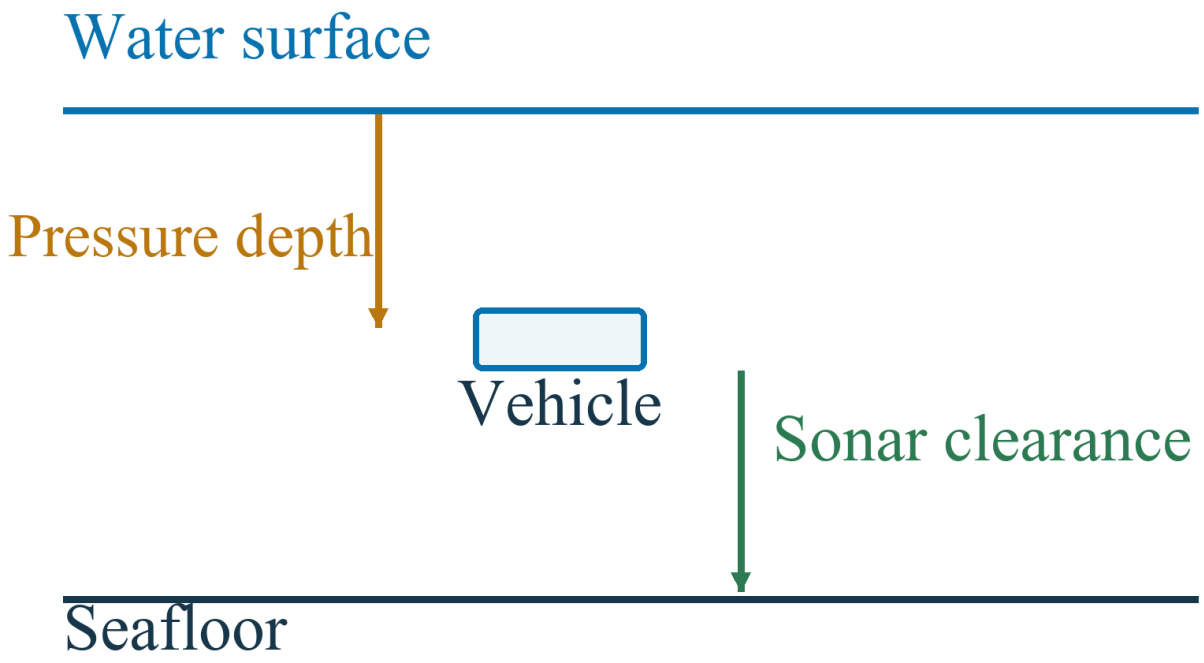


Diagram: Pressure Depth and Sonar Clearance

$$\Delta p = p - p_0 = \rho_w g h$$

$$h = \frac{100552.5}{1025 \times 9.81} = 10.0 \text{ m}$$

$\rho_w = 1025\ \text{kg/m}^3$, $g = 9.81\ \text{m/s}^2$. p_0 is surface reference pressure. Pressure increase is about 100.55 kPa. Assume hydrostatic water with constant density.

Pressure depth starts at the surface; sonar range starts at the vehicle. They have different references.

8.4.5 Other Sensors: Magnetometers

The Geomagnetic Field Provides a Heading Reference

Three-Axis Magnetic Field → **Tilt Compensation and Calibration** → **Magnetic-North Heading**

Influence	Effect on the Observation	Interpretation and Correction
Body Tilt	The measured field vector changes with body orientation.	Project using attitude. Use declination to convert magnetic to true north.
Hard/Soft Iron	Fixed sources add bias; materials distort direction or scale.	Calibrate bias, scale, direction; avoid magnetic contamination.
Current and Payload	Motors, power circuits, and metal structures add interference.	Correct using IMU, GNSS, and other heading cues.

8.4.5 Other Sensors: Temperature and Humidity

Environmental Conditions Affect Scale and Operation

Observation	Effect on the System	Adjustment or Interpretation
Ambient Temperature	Changes sound speed in air and device thermal drift.	Correct range scale; compensate drift using calibration models.
Ambient Humidity	Changes propagation, condensation, and operating conditions.	Use temperature too; protect equipment and adjust tasks.
Component Temperature	Indicates battery, motor, or drivetrain thermal state.	Component sensors support thermal management and output adjustment.
Temperature and Pressure	Together change air density and propulsion conditions.	Aid propulsion, energy management, and control adjustment.

Ambient readings do not replace component temperature; location and self-heating affect readings.

Exercises and Discussion

Build Observation Sets for Three Unmanned Systems

Scenario	Questions to Answer	Answer Requirements
Warehouse Platform	Target identity and nearby geometry? Motion estimates during satellite outage?	Use external and self-motion sensors; explain weak texture, occlusion, or drift.
Night Low-Altitude Flight	Thermal targets and forward obstacles? How to obtain attitude/reference altitude?	Check light, rain/fog, coverage, and scattering; separate altitude and ground clearance.
UUV in Turbid Water	Surface depth, seabed/obstacle range? How to maintain motion estimates?	Mark pressure/sonar origins; explain sound speed, pose, and camera limits.

Each design: Direct/estimated data, one failure, and coordinate, time, and reference conditions.

Suggested Answers

Combine Sensors to Fill Observation Gaps

Scenario	Sensors and Roles	Interpretation and Conditions
Warehouse	RGB: Targets. Stereo/RGB-D or LiDAR: Geometry. IMU: Specific force and angular velocity.	Low texture/occlusion limit geometry; inertial integration drifts. External positions constrain long-term errors.
Night Low Altitude	Thermal: Heat sources. Radar: Range/motion. IMU/barometer: Self-state cues.	Barometric altitude is not ground clearance. Coverage, scattering, weather limit obstacle detection.
Turbid-Water UUV	Active sonar: Obstacle/seabed slant range. Pressure: Own depth. IMU: Short-term motion.	Pressure/sonar origins differ. Temperature, salinity, pressure affect sound speed; turbidity limits cameras.

Section Summary

Mechanism, Reference Frame, and Environment Shape Usable Data

- **Start with the mechanism:** Light, electromagnetic waves, sound, inertia, and environmental sensing provide appearance, range, motion, or operating-condition data.
- **Distinguish outputs:** Color pixels, depth, point clouds, specific force, and estimated pose are different data. Semantics requires later processing.
- **Interpret errors and unknowns:** Coverage, occlusion, propagation, reflection, and mutual interference limit observations. No return does not directly mean free space.
- **Align references:** Slant range, water depth, ground clearance, and barometric altitude use different references. Fusion also needs consistent coordinates and time.

Complementary sensors need measurements that can be matched.