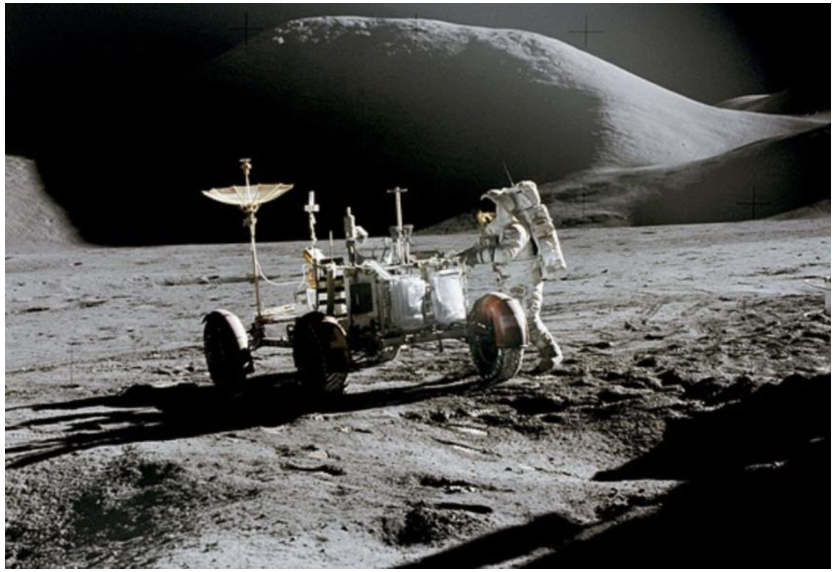
Visual Navigation System Overview

Graduate Course INTR-6000P

Week 1 - Lecture 2

Changhao Chen
Assistant Professor
HKUST (GZ)

The Motivation of Intelligent Vehicles



1) Enhanced Safety & Reduced Congestion: Mitigate human error—the cause of most accidents—and optimize traffic flow to reduce frustrating traffic jams.

2) Efficiency & Accessibility: Revolutionize logistics for faster delivery times and save labor costs.

3) Operation in Extreme Environments: Perform tasks in hazardous or inaccessible locations, such as deep-sea exploration or space missions.

The need of enhancing the safety, efficiency, accessibility of vehicles

The History of Self-Driving (1980s - 2000s)

The Rise of Autonomy and Machine Vision

Researchers began developing vehicles that could perceive and interpret their environment independently using sensors and computers.

1980s: The DARPA Autonomous Land Vehicle (ALV) project in the US used computer vision, lidar, and robotic control to navigate off-road terrain. It was slow but groundbreaking.

1987: Ernst Dickmanns, a German engineer, transformed the field. He equipped a van with cameras and computers, using "machine vision" to interpret the world in real-time.

1995: Dickmanns' improved car, VaMP, drove over 1,000 miles on a Paris multi-lane highway in traffic, including executing automated lane changes and passing other cars.

The History of Self-Driving (2000s)

The Grand Challenges

Acceleration Through Competition

2004: DARPA Grand Challenge (Mojave Desert): Result: Total Failure. No car finished the 150-mile course. The best car made it only 7.3 miles. It proved how incredibly difficult the problem was.

2005: DARPA Grand Challenge: Result: Stunning Success. Stanford's "Stanley" won, completing the course. It used sensors, computers, and sophisticated software, proving the feasibility of autonomous navigation in harsh environments.

2007: DARPA Urban Challenge: Cars now had to navigate a mock urban environment, obeying traffic laws and interacting with other robots. Carnegie Mellon's "Boss" won. This challenge proved autonomous systems could handle complex, real-world rules.

The History of Self-Driving (2010 - Present)

Industry Race & AI Dominance

2009: Google secretly starts its self-driving car project, later becoming Waymo. They began testing on public roads, accumulating millions of real-world miles.

2010s- Present: A massive industry race begins. Traditional automakers, tech giants, and dedicated startups all enter the fray.

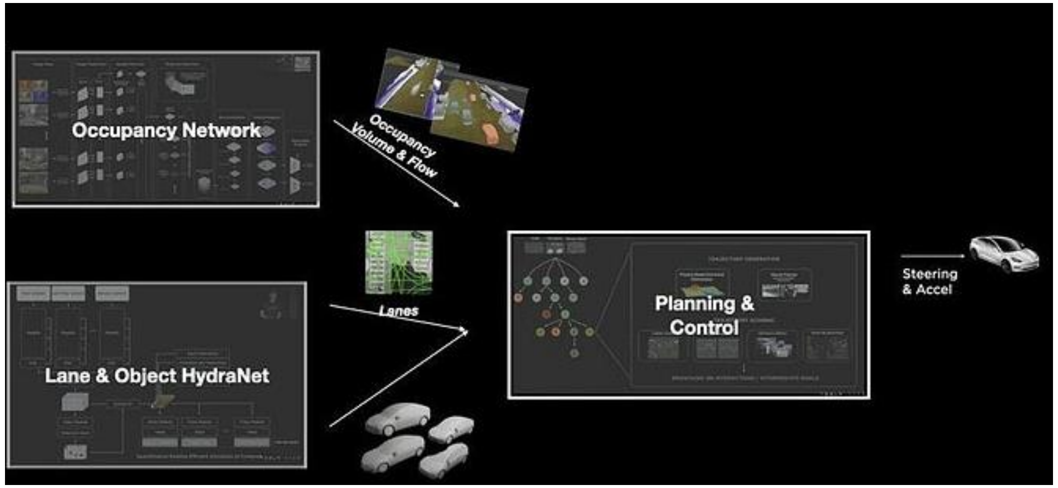
In China,

Baidu, PonyAI, WeRide

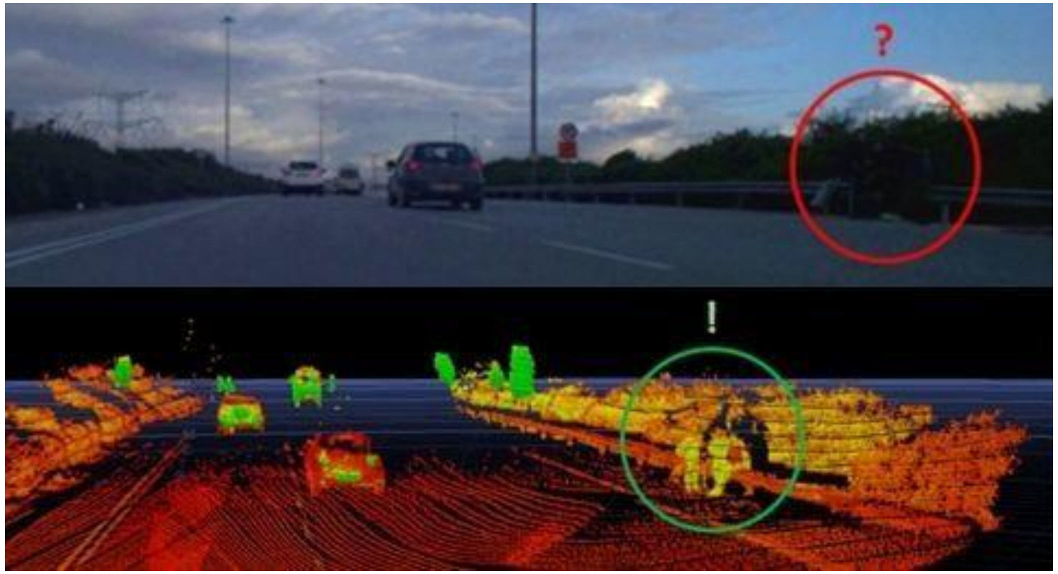
NIO, Xpeng, Li Auto, Didi, BYD

Discussion

Rules vs. Deep Learning: The approach shifts from hand-coded rules to end-to-end deep learning systems, where neural networks learn to drive from vast amounts of data.



The Lidar vs. Vision Debate: Two main sensor philosophies emerge: relying heavily on expensive lidar (Waymo, most others) vs. using primarily cameras and AI, championed by Tesla with its "Full Self-Driving" system.



Key Techniques - Sensors

Sensors are the hardware that collects raw data about the environment. Each type has unique strengths and weaknesses, making them complementary.

Cameras:

Provide rich semantic information (color, texture, text). They are essential for reading traffic signs, recognizing traffic light colors, detecting lane markings, and classifying objects (e.g., pedestrian vs. cyclist).

Challenge:

They produce 2D images, making it difficult to accurately judge distance and speed, and their performance degrades in poor lighting (darkness, glare, fog).

Key Techniques - Sensors

Monocular Camera:

Single lens, low-cost

Rich information

Limitation: lacks depth perception

Stereo Camera:

Mimics human eyes → depth from disparity

Useful for near-field obstacle detection

More expensive & calibration-sensitive

Sensor Placement:

Front-facing cameras - lane keeping, forward vision

Surround-view (4–8 cameras) - 360° coverage

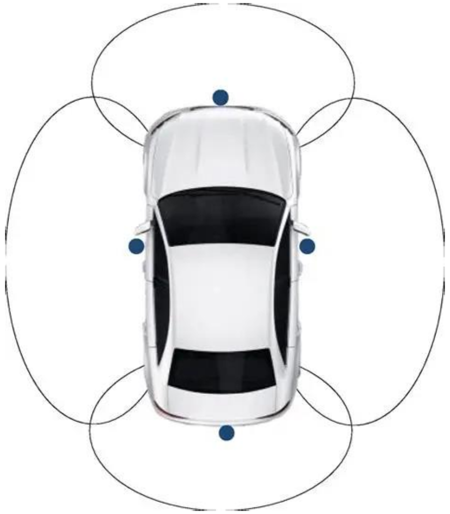
Key Techniques - Sensors

Fisheye & Surround-View Cameras:

Wide field of view (180°)

Typically used for parking, blind-spot monitoring

Enable bird’s-eye surround-view systems

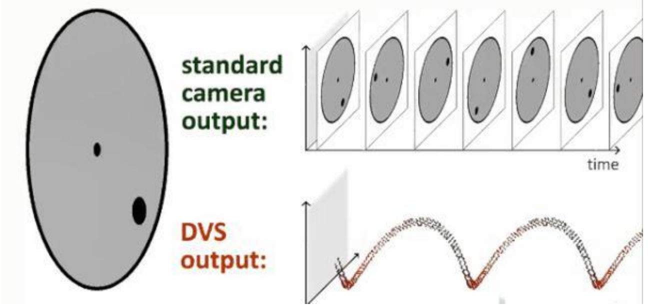


Thermal & Event Cameras:

Thermal cameras: robust to low light, fog, night driving

Event cameras: capture fast motion with low latency

Emerging but not yet mainstream



Key Techniques - Sensors

GPS (Global Positioning System):

Primary Role: Provides absolute global positioning (latitude, longitude, and altitude).

Key Strength: It gives a vehicle its precise location on a global map, which is essential for determining the overall route from origin to destination.

Critical Weakness: Its signal is unreliable in urban canyons, tunnels, and dense cities. Tall buildings can block, reflect, or multipath signals, leading to inaccurate or lost positioning. It also updates at a slow rate (~1-10 Hz), making it insufficient for real-time vehicle control.

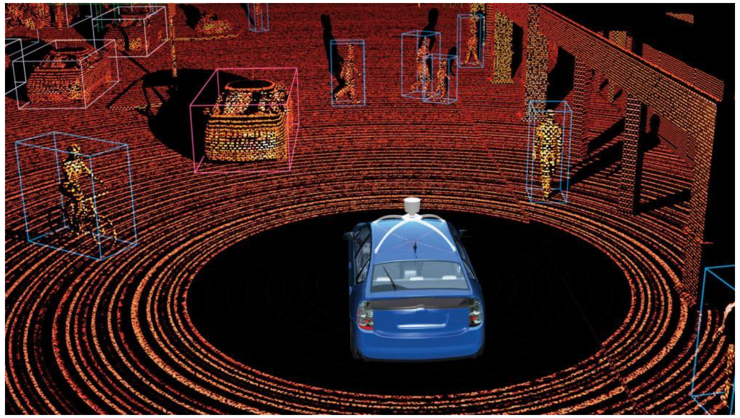
Key Techniques - Sensors

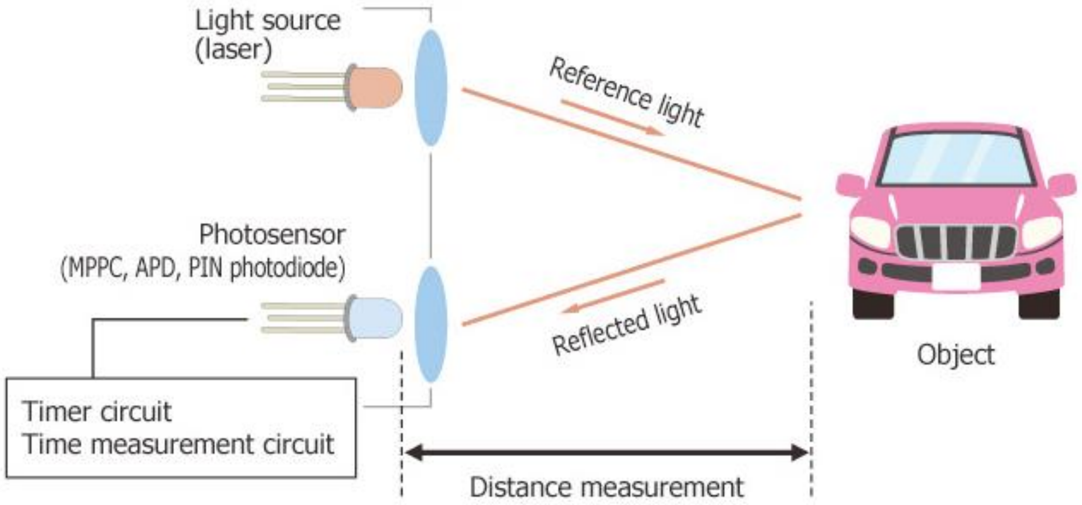
LiDAR (Light Detection and Ranging):

Primary Role: Creates a high-resolution 3D point cloud map of the immediate environment by measuring the distance to objects with laser pulses

Key Strength: Provides extremely accurate geomet information. It is excellent for understanding the precise shape, size, and distance of obstacles, othe vehicles, and road contours. It works well in the dark.

Critical Weakness: Historically expensive and can be bulky. Performance can be degraded by heavy rain, fog, or snow, which scatter the laser beams. Lower-resolution LiDARs may struggle with fine details..





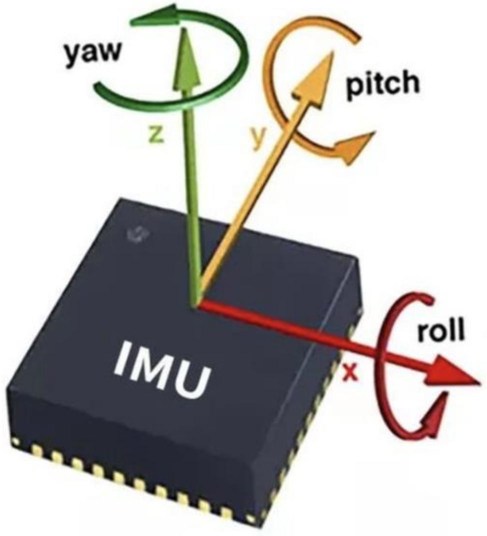
Key Techniques - Sensors

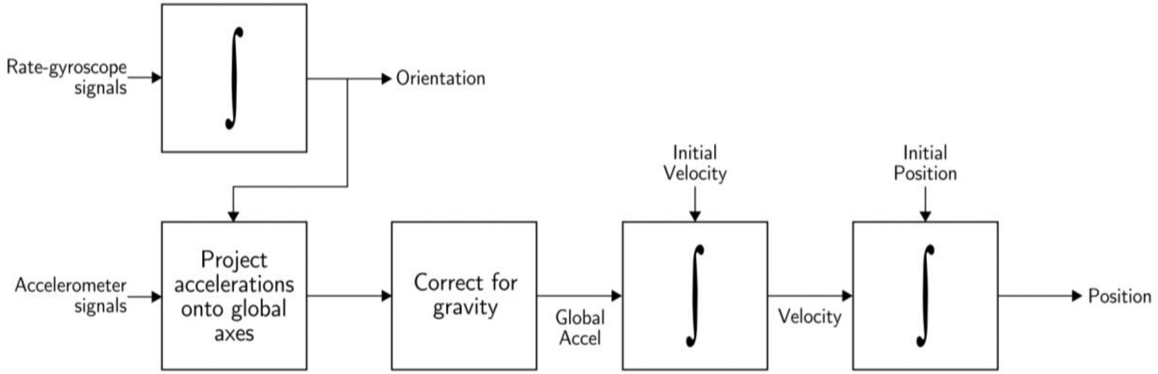
IMU (Inertial Measurement Unit):

Measures the vehicle's short-term motion and dynamics.

Key Strength: It provides high-frequency data (~100-1000 Hz) on acceleration and rotational rate. This is crucial for tracking the vehicle's movement between updates from slower sensors like GPS or cameras (e.g., during quick turns or sudden braking). It is not affected by external environmental conditions like weather or lighting.

Critical Weakness: Its estimates drift over time. Any small error in measuring acceleration or rotation is integrated into the position estimate, leading to exponentially growing inaccuracies if not corrected by other sensors.





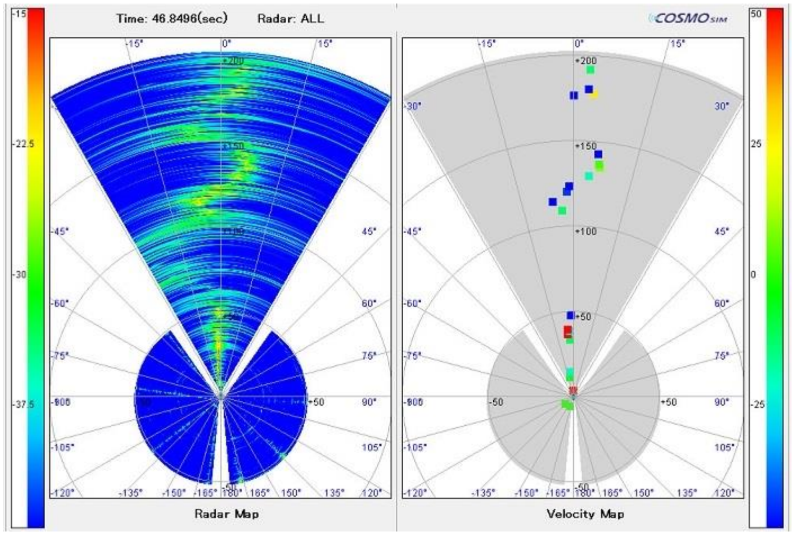
Key Techniques - Sensors

Radar (Radio Detection and Ranging):

Detects objects and precisely measures their relative distance and velocity.

Key Strength: Highly robust in adverse weather conditions like rain, fog, and snow. It directly measures the speed of approaching objects using the Doppler effect, making it exceptional for adaptive cruise control and collision warning.

Critical Weakness: Offers very low resolution compared to LiDAR and cameras. It typically struggles to discern the exact shape or type of an object (e.g., distinguishing a car from a guardrail) and can have difficulty with stationary objects.



Key Techniques – Computer Vision

Turning raw pixels into understanding:

- Humans drive mainly through vision
- Cameras replicate human perception for vehicles
- Enables understanding of lanes, obstacles, traffic signs

Object Detection and Recognition

Semantic Segmentation

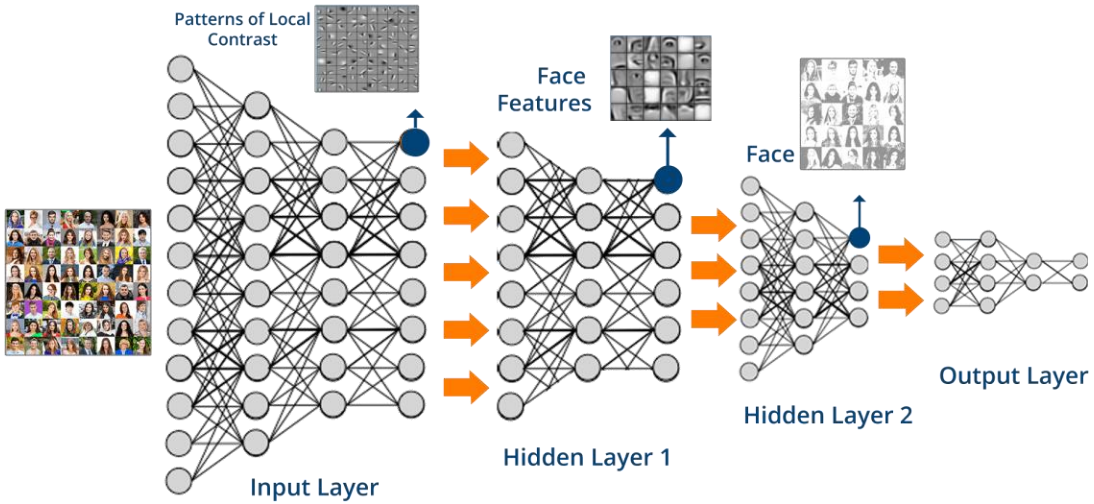
Visual Localization and Mapping

Depth Estimation

Etc.

Key Techniques – AI and Deep Learning

AI and especially Deep Learning provide the "brain" that learns to use those tools effectively. Instead of relying on hand-coded rules, these systems learn directly from vast amounts of data.



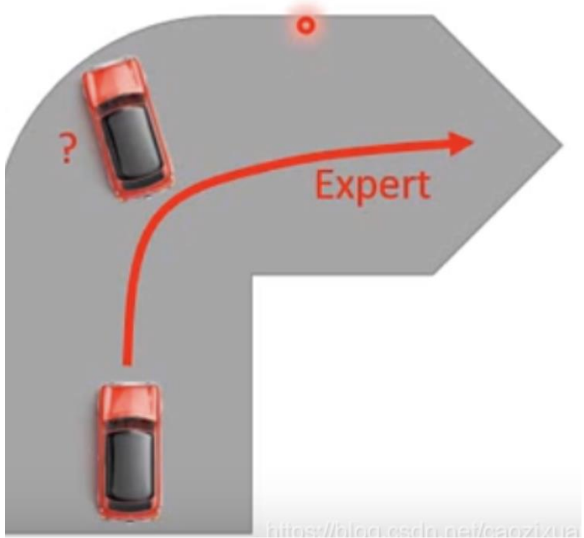
Convolutional Neural Networks

Reinforcement Learning

Imitation Learning

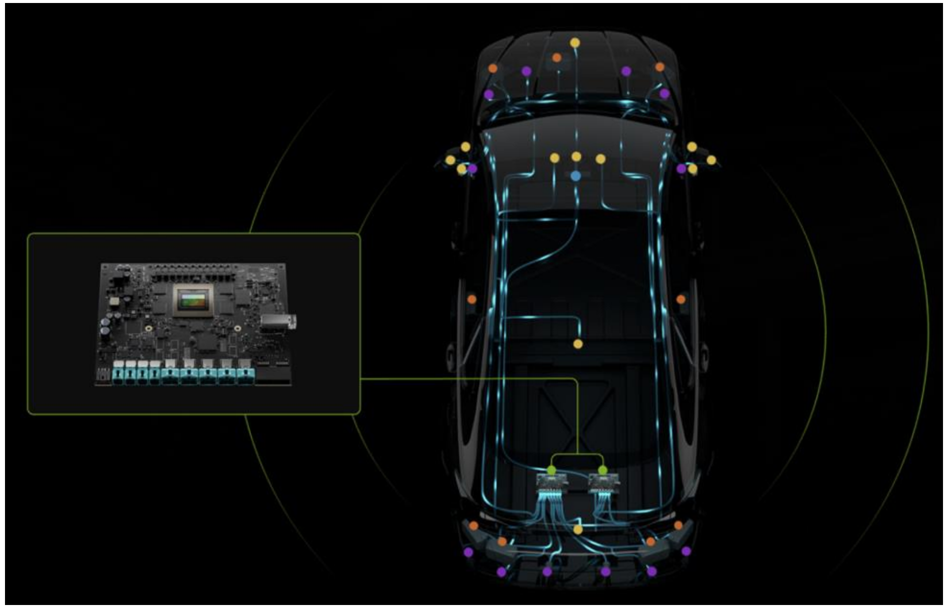
Transformers

End-to-end Driving



Key Techniques – Computing Hardware

The sophisticated algorithms for perception and decision-making require immense computational power. The computing hardware is the central nervous system of the intelligent vehicle, responsible for processing massive data streams in real-time under rigorous constraints.

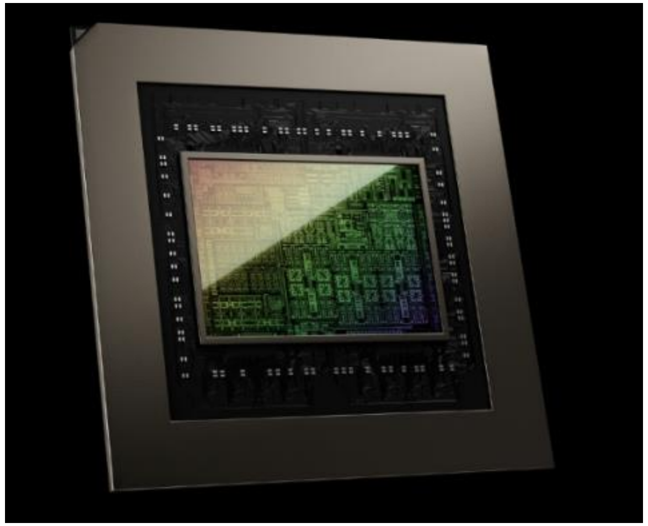


High-Performance Computing

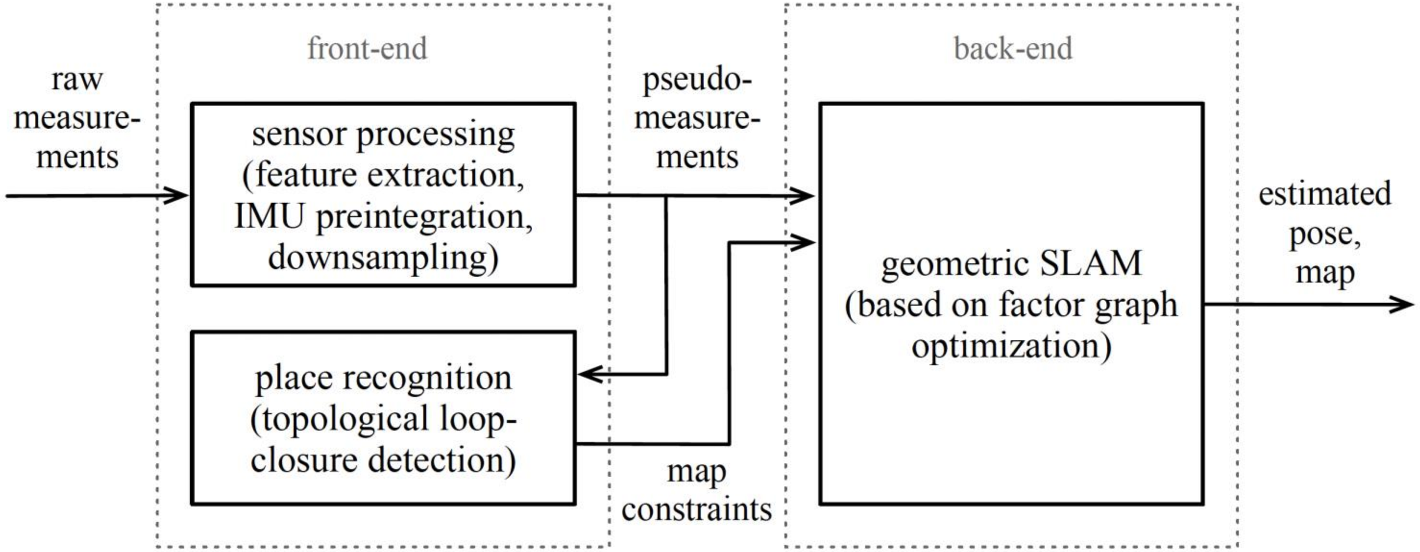
AI Accelerators (GPUs & NPUs)

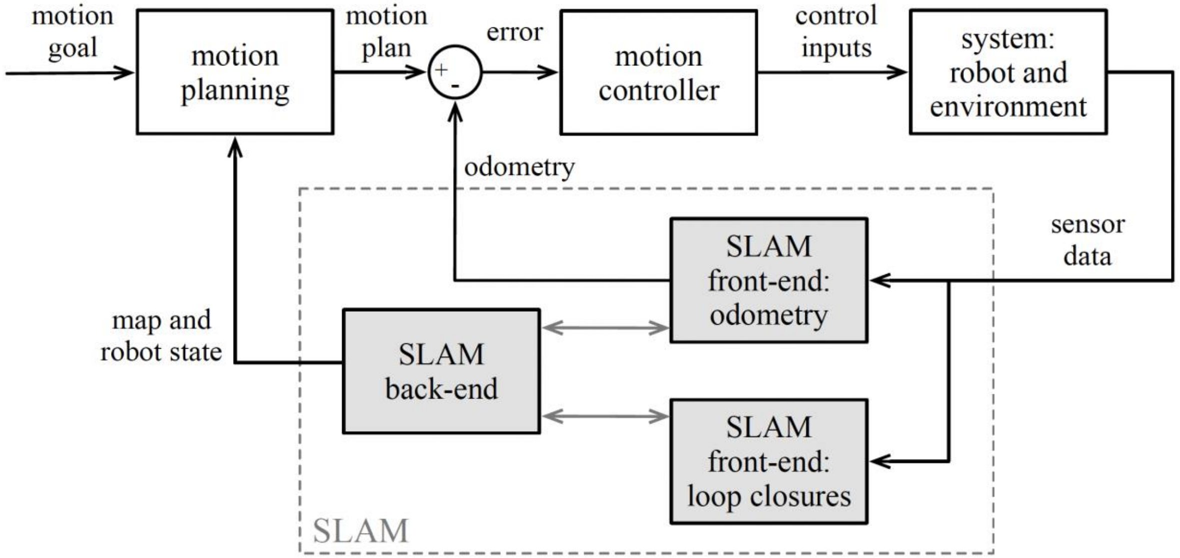
System-on-a-Chip (SoC) Integration

Edge Computing



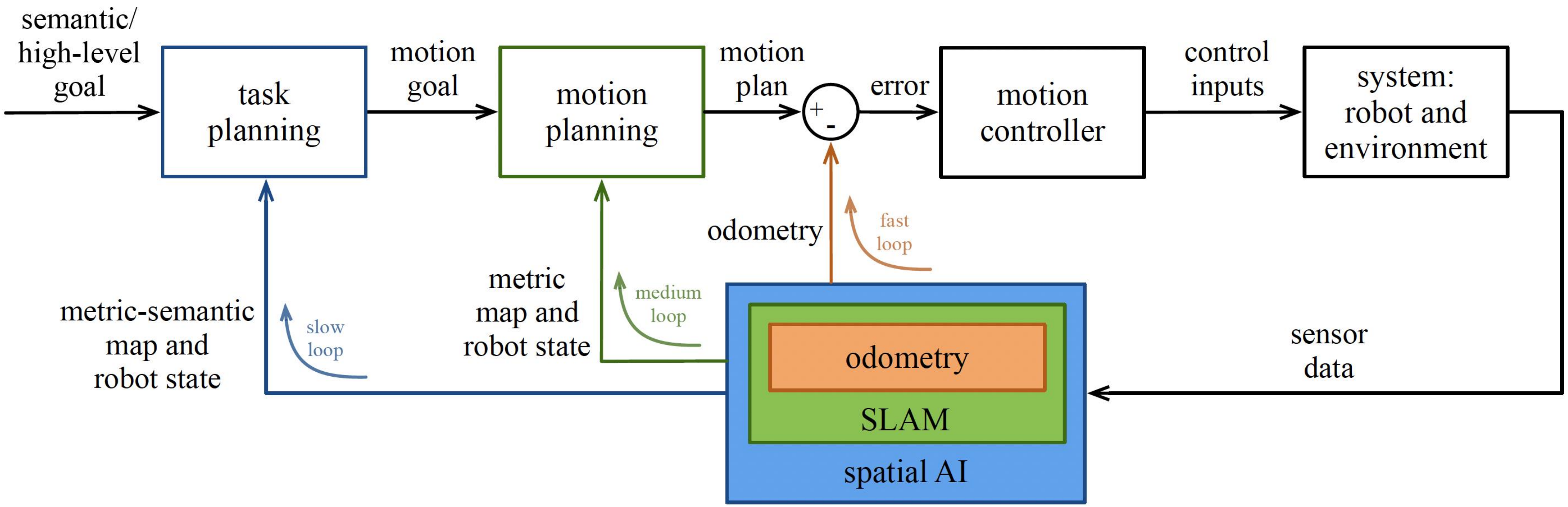
Visual Navigation





Source: SLAM Handbook

Visual Navigation



Source: SLAM Handbook

Key Challenges and Research Frontiers

1) Robustness to Conditions: Dealing with bad weather (rain, snow, fog), low light, and sun glare.

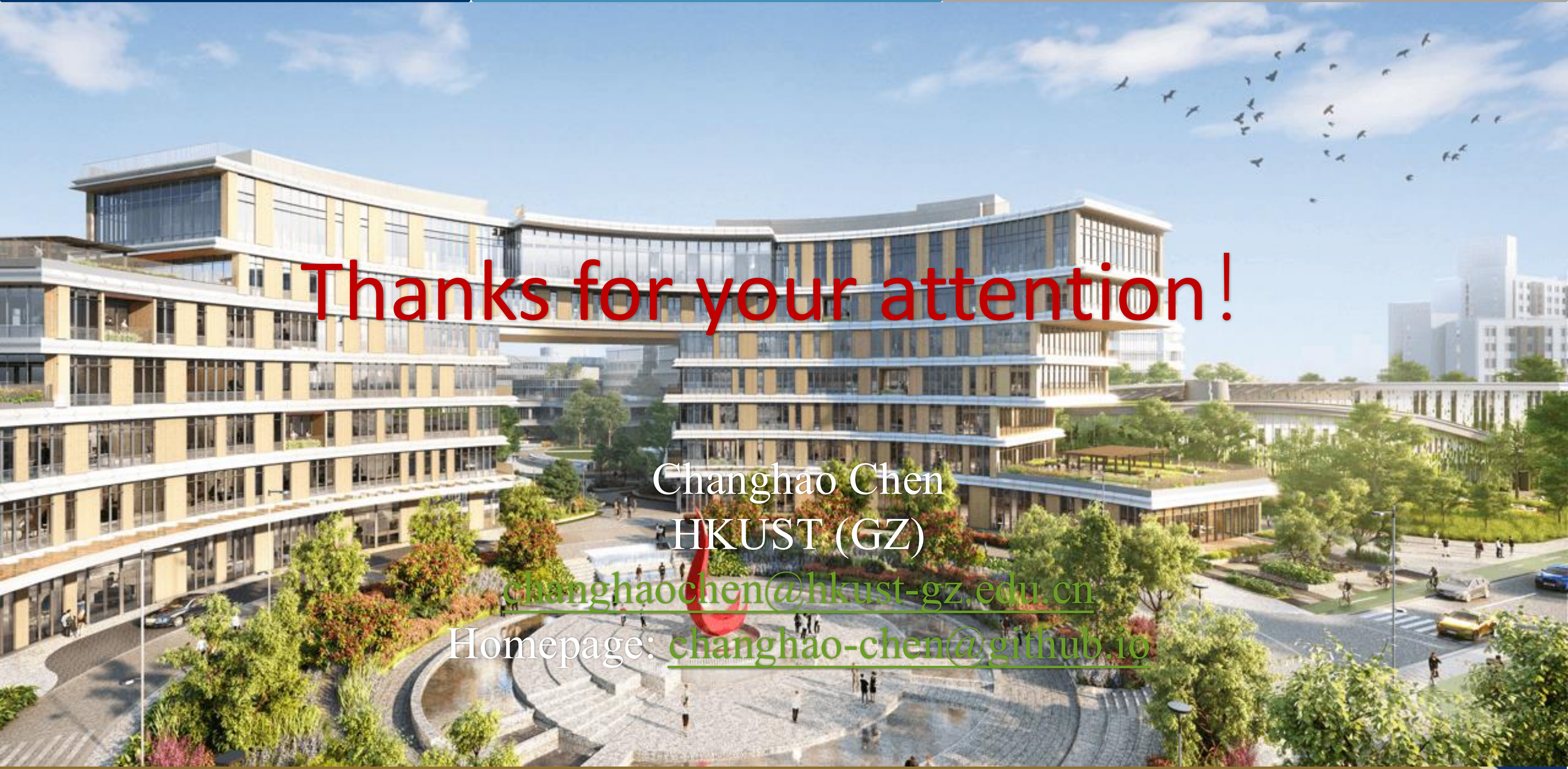
2) Safety and Reliability: The "edge cases" or corner scenarios that are difficult to handle. The need for rigorous testing and validation.

3) Computational Efficiency: Processing massive amounts of visual data in real-time on embedded hardware.

4) The "Sim-to-Real" Gap: Using simulation for safe and scalable testing, and transferring that learning to the real world.

5) Multi-agents

Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. IEEE Transactions on Robotics



Thanks for your attention！

Changhao Chen
HKUST (GZ)
changhaochen@hkust-gz.edu.cn
Homepage: changhao-chen@github.io