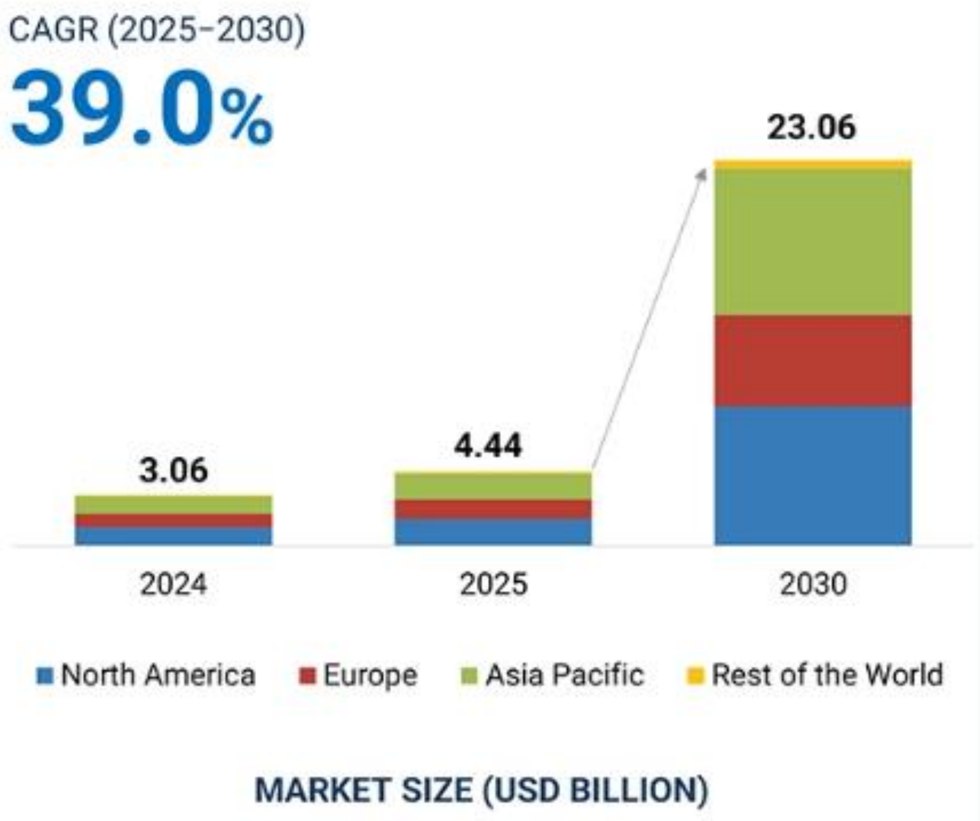
Embodied AI and TSN Potentials

Tongtong Wang, Lihao Chen
Huawei Technologies

IEEE 802.1 Interim Meeting, Sept 2025

HUAWEI

Embodied AI in rapid global growth

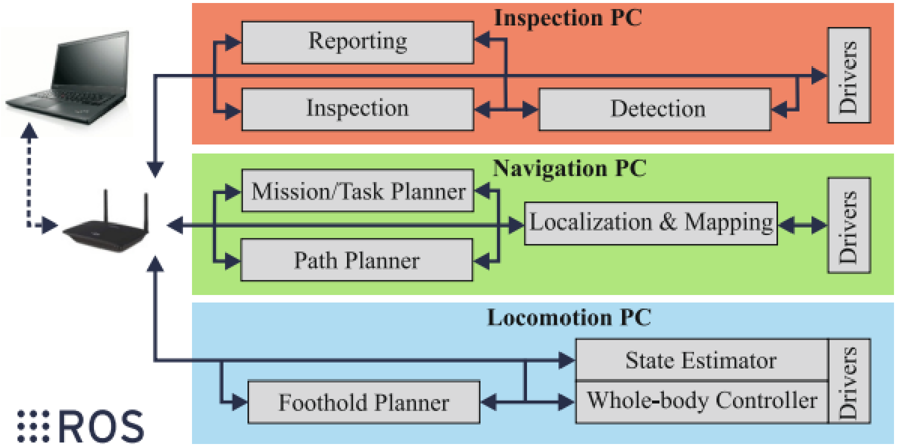


Key trends fueling this growth include:

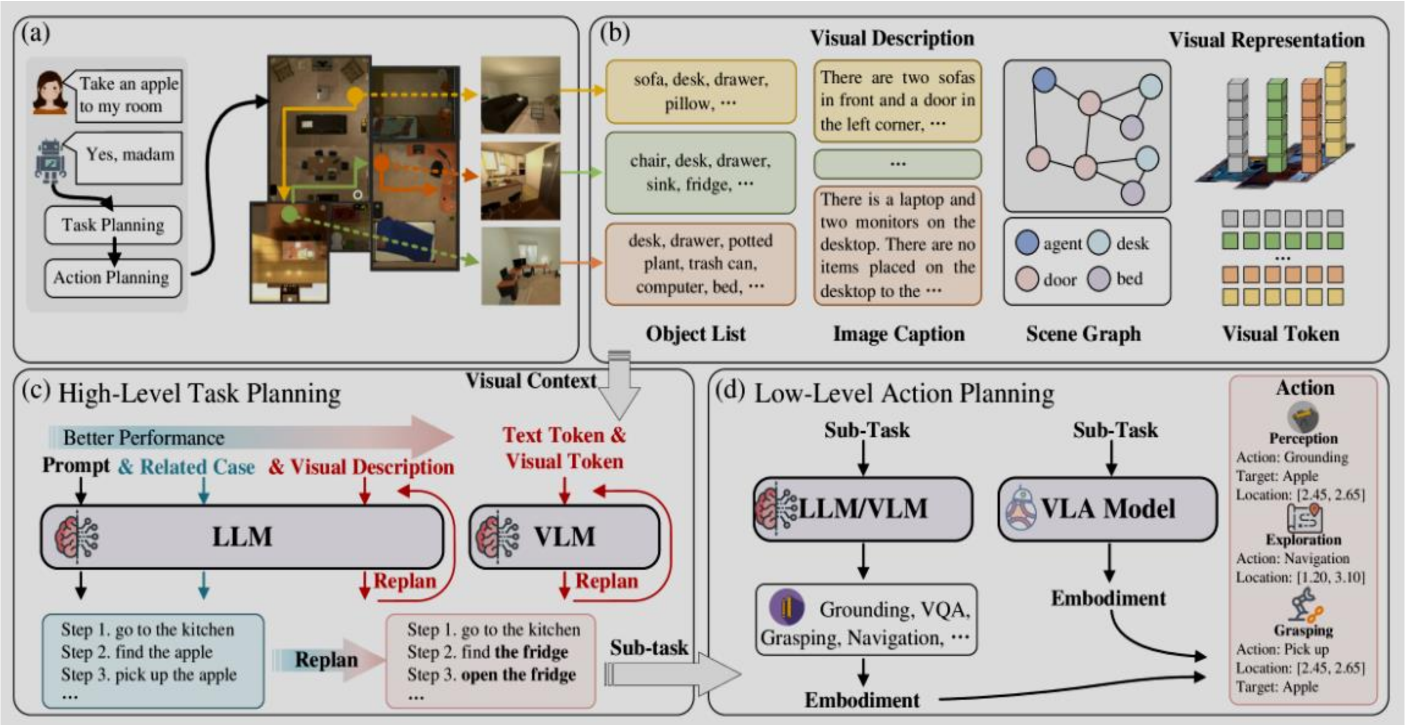
- Integration with large language and vision models, enabling robots to better perceive, understand, and interact with complex real-world environments.
- Expansion into diverse sectors such as manufacturing, warehousing, healthcare, home services, and education.
- Commercial deployment of humanoid and service robots by leading global tech and robotics companies （Telsa, FigureAI, 1X, Unitree, UBtech …） , signaling a shift from research to large-scale, real-world applications.

https://finance.yahoo.com/news/embodied-ai-research-global-forecast-104700343.html

Embodied AI Control Architecture



Classical robot control



Embodied Robot control

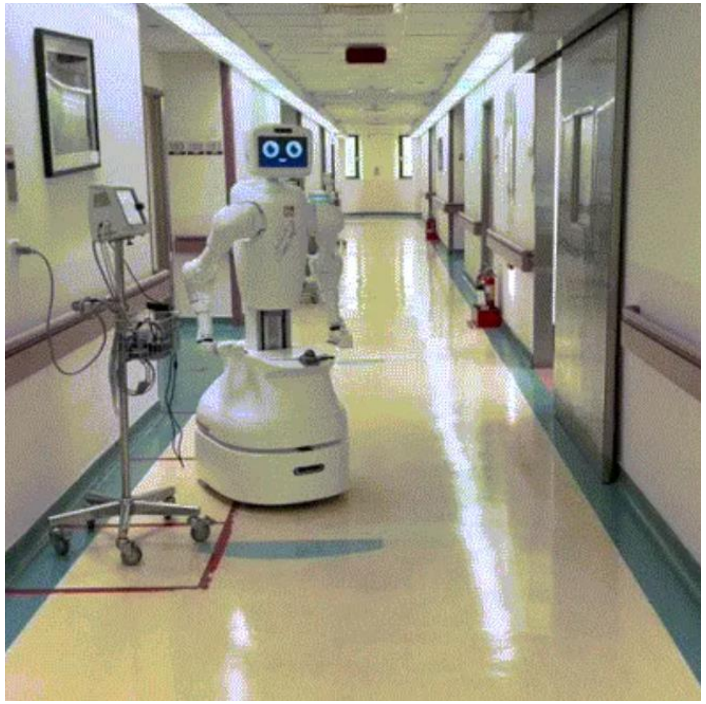
- Classical robot control methods based on kinematics and dynamics utilize control theory (such as PID, Model Predictive Control (MPC), and Whole-Body Control (WBC)) to achieve multi-objective control of position, velocity, attitude, or torque.
- Embodied AI robot is a general-purpose intelligent agent that can understand, learn, plan, make decisions and act across diverse complex scenarios like a human. It integrates perception, reasoning, planning, and action in an end-to-end manner. Most current prototypes are based on large Transformer models with multi-modal perception and autonomous decision modules, though other architectures like world models are also being explored.

HUAWEI

https://arxiv.org/html/2407.06886v1

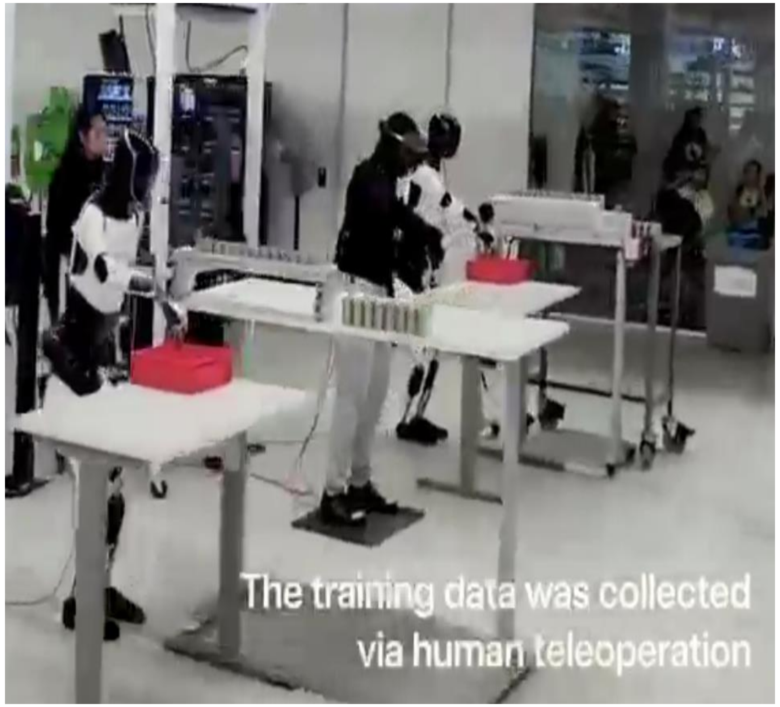
Embodied AI Use Cases

Generally anticipated that humanoid robots will begin to be deployed in education and retail sectors from 2025, achieve adoption in industrial manufacturing within 3 to 5 years, and gradually enter household general-purpose services in 10 years.



Companion robots in healthcare are capable of performing simple tasks such as conversation and navigation, while also incorporating tactile and emotional perception capabilities.

Material handling, logistics, and sorting scenarios in factories.



Teleoperation is a key method for data collection in embodied intelligence, providing ground-truth demonstrations across real-world and simulated environments

HUAWEI

Statistics on Sensors in Embodied Intelligent Robots

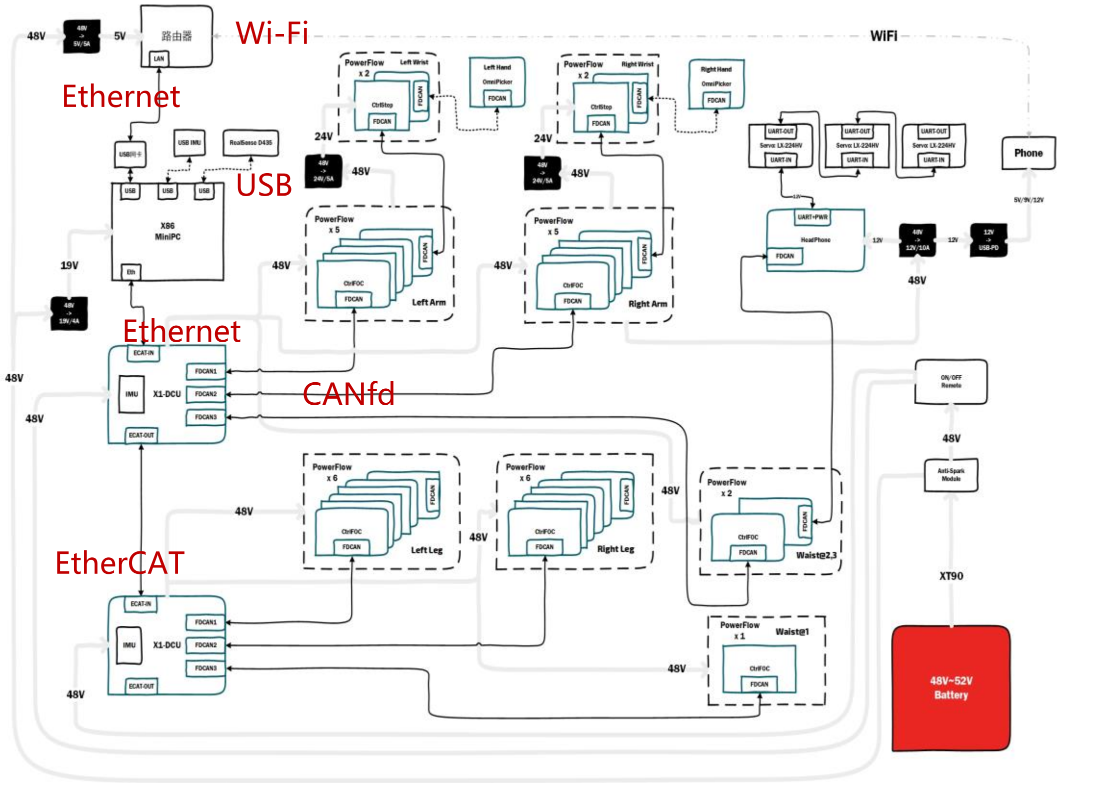
Sensor	Usage	Count	Bandwidth	Note
Depth Camera	Scene Perception, Obstacle Detection, and Foothold Estimation	2–3	30fps, 10–200Mbps/device（RealSense 848×480 @ 30fps , one depth frame 16bit+2 RGB frame 24bit ~**488 Mbps , compressed to 50-200Mbps**）	
Wide-Angle RGB Camera	Semantic Understanding / Eye–Hand Coordination	1–2	1080p (1920×1080) @30 fps， each pixel 24 bit（3B RGB）≈ 1920×1080×3×30 ≈ 178 MB/s ≈ 1.4 Gbps Compressed to（H.264/H.265， 10:1～20:1） → 50–200 Mbps	Integrated with semantic perception
3D LiDAR	Mapping / Long-Range Obstacle Detection	0–1	10–30Mbps	Essential for outdoor or low-light environments
IMU	Posture, Balance, and Visual–Inertial Odometry (VIO)	1–2	200–1000Hz, kB/s	Tightly coupled with VIO (Visual–Inertial Odometry)
Joint Encoder	Whole-Body Control	24–30	500–1kHz, ~12kB/s/Joint	EtherCAT/CAN-FD
Foot Force Sensor	Contact Detection and Zero Moment Point	8–12点	200–500Hz， 16 **bit×500 Hz ≈ 1 kB/s** **Total 10–20 kB/s**	Event-triggered reporting supported
Downward-Facing Depth Camera	Grasping and Foothold Planning	1-2	30fps, 10–200Mbps/device	For grasping actions or foothold planning Agility Robotics Digit： NASA Valkyrie：
Tactile Skin Sensor	Collision Avoidance / Human–Robot Interaction	100–300 Taxels	100–200Hz， 16 bit×200 Hz ≈ 0.4 kB/s Total 40KB/s ~ 4.8 Mb/s （on each arm）	Covers load-bearing / contact surface for manipulation
Microphone Array	Speech Interaction / Alarm Notification	4–8	Audio 0.5–2Mbps	

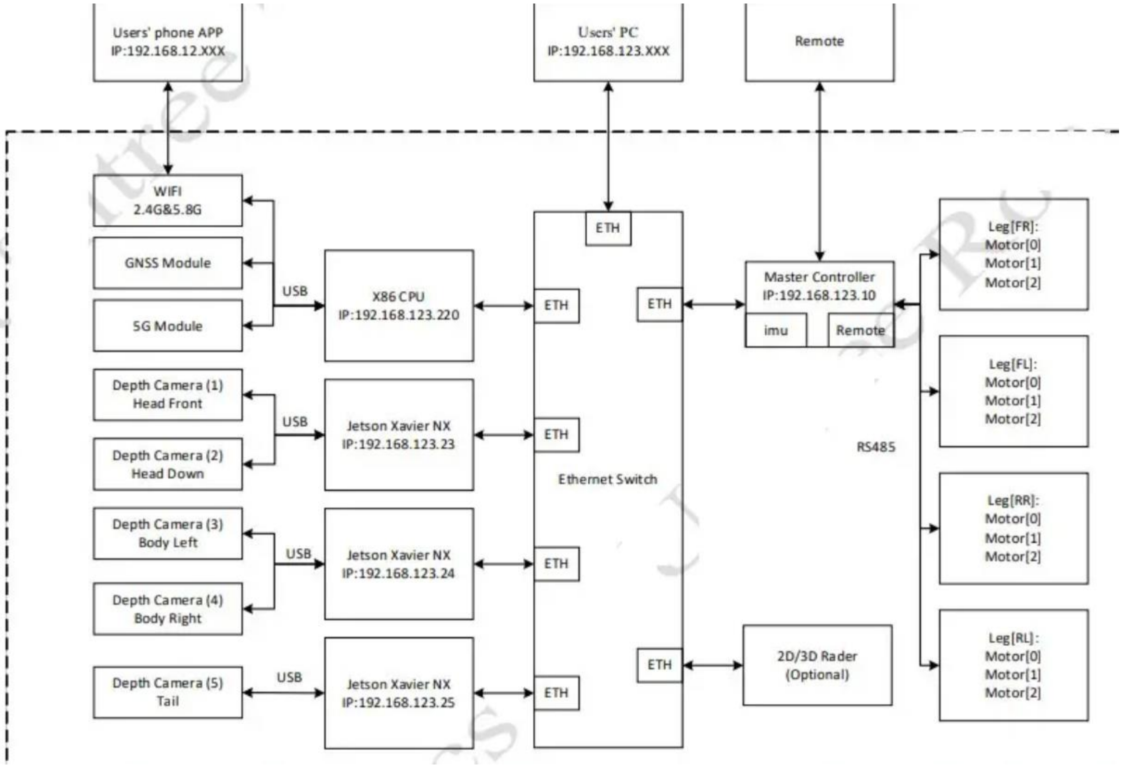
Xinbu HB03

(using the bipedal as an example)

https://www.intel.cn/content/dam/www/central-libraries/cn/zh/documents/2025-04/25-eai-accelerating-industry-innovation-and-solution-delivery-with-body-intelligent-brain-convergence-solutions-whitepaper.pdf

Internal Communication in Robots





At present, the most common internal communication setup for robots

- EtherCAT for drive control, Gigabit Ethernet as the perception backbone,
- GMSL/MIPI for camera links,
- CAN/CAN-FD for low-speed peripherals.

https://github.com/AgibotTech/agibot_x1_hardware/
https://device.report/manual/8384149

TSN Potentials in Embodied AI

Advantages
Unified Ethernet backbone (control + sensing + video + management) : reduces cabling, hardware costs, and infrastructure complexity, also allows priority management so time-critical traffic (e.g., motion control signals) is never interrupted by Non-RT data.
Deterministic & real-time (sub-ms latency, <50 µs **sync**): guarantees that control and sensor data are delivered within strict timing limits, crucial for real-time coordination between robot components like controllers, actuators, and sensors; 802.1AS ensures high-precision synchronization.
High bandwidth scalability (1–10 Gbps, suitable for vision/AI), much larger than EtherCAT and CAN.
Standards-based, interoperable with OT platform and compatible with DDS, which is popular for drone and robot systems.
Supports security, SDN, and centralized management

Improvement needed:

- Higher Implementation and Configuration Complexity → find ways to simplify configurations
- More Reliability Mechanisms → on PHY layer and MAC layer
- Immature Ecosystem in Robotics → collaboration with robot component suppliers and system integrators

Contact author for more discussion and joint innovation. *Tongtong.wang@Huawei.com*

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home and organization for a fully connected, intelligent world.

Wireless Connectivity Requirements for Robots in Factory

- **Latency**
  - 10 ms for real-time control
  - 50 ms for perception and sensor feedback
- **Bandwidth**
  - 50 Mbps–1 Gbps per robot
  - 10 Gbps aggregate throughput for multi-robot collaboration
- **Reliability**
  - ≥99.999% system availability
  - Packet loss rate $<10^{-6}$
- **Scalability & QoS**
  - Support for high-density device access
  - Guaranteed Quality of Service for critical flows
- **Security**
  - End-to-end encryption
  - Strong authentication
  - Anti-interference and resilience against cyber-physical threats