MARCUS PROJECT

Progress & Achievement Log

YS Lootah Technology | Robotics & AI Division | April 2026

STATUS: ROBOT WALKING + AI VISION CONTROL — FULLY WORKING (April 4, 2026)

1. Executive Summary

Project Marcus achieved two major milestones across April 3-4, 2026. On April 3, the Unitree G1 EDU humanoid robot achieved full AI locomotion control via Holosoma at 50Hz. On April 4, the complete autonomous AI stack was deployed: LLaVA 7B vision model running offline on the Jetson Orin NX, understanding natural English commands, seeing through the RealSense camera, and controlling Holosoma via ZMQ. The robot now responds to multi-step natural language commands and executes them in sequence.

April 4 breakthrough: 'move forward and then turn right' — robot executed both steps in sequence. LLaVA sees camera, understands English, moves robot. Fully offline, no cloud, no remote.

2. April 4, 2026 — Full Day Timeline

Time	Activity	Status
09:00	**ZMQ input system built for Holosoma** Created custom ZMQInput class, registered in Holosoma input factory. Robot received start/walk/stop commands via ZMQ.	**DONE**
10:30	**Robot turns confirmed via ZMQ** vyaw -0.3 sent via PUB socket, Holosoma receives via SUB, robot physically turns right. First AI-controlled movement.	**DONE**
11:00	**Multi-step ZMQ commands working** start + walk + velocity sequence executes cleanly. gradual_stop() prevents jerking. Robot walks and turns on command.	**DONE**
13:00	**Ollama + LLaVA 7B deployed on Jetson** ollama serve running, llava:7b pulled (4.75GB). GPU: Orin CUDA, 15GB VRAM. LLaVA loads in 3.8s.	**DONE**
14:00	**LLaVA + camera integration** RealSense D435I frames (640x480) captured in background thread, encoded as base64 JPEG, sent to LLaVA with command. Response 4-7s.	**DONE**
15:00	**Full Marcus brain working**	**DONE**

	LLaVA sees camera + understands English → JSON action list → ZMQ → Holosoma → G1 moves. 'turn left', 'walk forward', 'what do you see?' all working.	
16:00	**Multi-step action chains** Prompt updated to return actions list. 'move forward and then turn right' → Step 1/2 forward, Step 2/2 right. Both execute in sequence.	**DONE**
17:00	**Inspect office command tested** Robot scans office, describes people, furniture, layout. LLaVA identifies scene content accurately.	**DONE**
20:00	**Goal-based navigation designed** navigate_to_goal() function designed: continuous camera loop, LLaVA checks if goal reached each step. Not yet deployed.	**DESIGN**

3. Confirmed Working Capabilities

Command	Robot Action	Confirmed
hi	No movement — greets	April 4
what do you see?	Describes scene from camera	April 4
turn left / turn right	Rotates 2 seconds	April 4
turn left three steps	Rotates 3 seconds	April 4
walk forward	Forward 2 seconds	April 4
step one step	Forward 1 second	April 4
move back	Backward 2 seconds	April 4
move forward and then turn right	Forward then right — multi-step	April 4
move backward and step three steps to the right	Backward then right	April 4
inspect the office	Scans and describes room	April 4

4. Current System Architecture (April 4, 2026)

```
You type: "move forward and then turn right"
        |
        v
RealSense D435I (background thread, 640x480 continuous)
        |
        v
LLaVA 7B via Ollama (offline, Jetson GPU, 4-7s)
sees camera + understands command
        |
        v
JSON action list:
```

```
{"actions":[{"move":"forward","duration":2.0},{"move":"right","duration":2.0}],"spe
ak":"..."}
        |
        v
marcus_llava.py executor
runs steps in sequence, gradual_stop() between
        |
        v
ZMQ PUB tcp://127.0.0.1:5556
        |
        v
Holosoma RL policy (50Hz, fastsac_g1_29dof.onnx)
        |
        v
G1 EDU 29 joints — robot physically moves
```

5. Performance Metrics

Metric	Value	Notes
Holosoma RL rate	50.00 Hz	Stable, zero drops
ONNX inference	0.5-2ms	Per locomotion step
LLaVA response time	4-7 seconds	First call 9s warmup
Camera resolution	640x480 15fps	Default config
ZMQ latency	<1ms	Local loopback
LLaVA model size	4.75 GB	llava:7b quantized
GPU VRAM used	~5GB of 15GB	Orin NX unified memory
Camera reconnect time	~2 seconds	Auto on USB drop

6. Next Steps

6.1 Immediate (Tomorrow)

- Wire G1 Arm SDK — raise, wave, point arm commands
- Add Whisper STT for voice input (pip install openai-whisper)
- Implement goal-based navigation loop (navigate_to_goal)
- Wire /dog_odom for precise distance walking
- Add LiDAR obstacle avoidance (front < 0.5m stop)

6.2 This Week

- Add Arabic language support via Whisper + LLaVA
- Implement patrol route (waypoint list loop)
- Wire YOLO PPE detections into LLaVA context

- Add Piper TTS for robot voice responses
- Test NaVILA after download completes (16GB checkpoint)

6.3 Month 2

- Install GR00T N1.6 for arm manipulation
- AMASS dataset retargeting for custom WBT motions
- Fine-tune LLaVA/Qwen on ADNOC facility data
- Full autonomous HSE inspection loop
- ADNOC client demonstration

7. Key Commands Reference

7.1 Start Marcus AI Brain

```
# Tab 1 — Holosoma ZMQ mode
source ~/.holosoma_deps/miniconda3/bin/activate hsinference
cd ~/holosoma && sudo jetson_clocks
~/.holosoma_deps/miniconda3/envs/hsinference/bin/python3 \
  src/holosoma_inference/holosoma_inference/run_policy.py \
  inference:g1-29dof-loco \
  --task.model-path
src/holosoma_inference/holosoma_inference/models/loco/g1_29dof/fastsac_g1_29dof.onn
x \
  --task.velocity-input zmq --task.state-input zmq --task.interface eth0

# Tab 2 — Marcus brain
ollama serve &
conda activate marcus
/home/unitree/miniconda3/envs/marcus/bin/python3 ~/Models_marcus/marcus_llava.py
```

7.2 Network Reference

Device	IP Address	Notes
Jetson Orin NX	192.168.123.164	unitree / 123
Workstation RTX 4060	192.168.123.222	zedx
G1 Locomotion Computer	192.168.123.161	proprietary
Livox Mid-360 LiDAR	192.168.123.120	ping alive

Marcus is now an AI-controlled robot. From joystick to full natural language vision control in 2 days. YS Lootah Technology | Kassam | April 3-4, 2026