Learning Visual Locomotion with Cross-Modal Supervision

Antonio Loquercio, Ashish Kumar, Jitendra Malik

Presenter: Ken Shao
Discussion: Jiahang Li

3/15/2023

“We see in order to move, and we move in order to see”

- **Vision gives us “look-ahead,”**
  - When we can see the ground ahead of us, our foot placement is smoother and the walking is more efficient.
  - Our gaze distributions vary with terrain.

Gaze distribution of a participant wearing eye tracker, Bonnen et al.

- **and “look-ahead” is tightly coupled to the walking strategy**

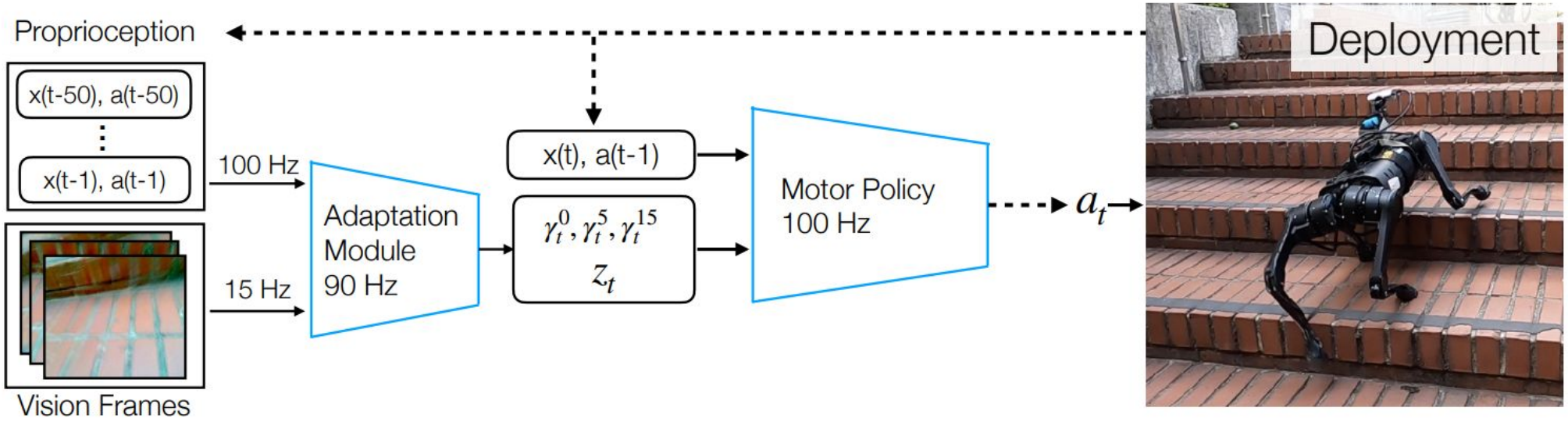
Major Contribution

1. **Look-ahead visual walking policy**
   - Only uses a single monocular RGB camera and proprioception.
   - The vision policy anticipate the geometry and adapts the robot's stride accordingly.
2. **Cross-Modal Supervision (CMS) from proprioception**
   - Proprioception supervises vision system with a time lag.
3. **Lifelong learning of on-the-fly terrain prediction**
   - The quality of the vision predictor improves with experience.

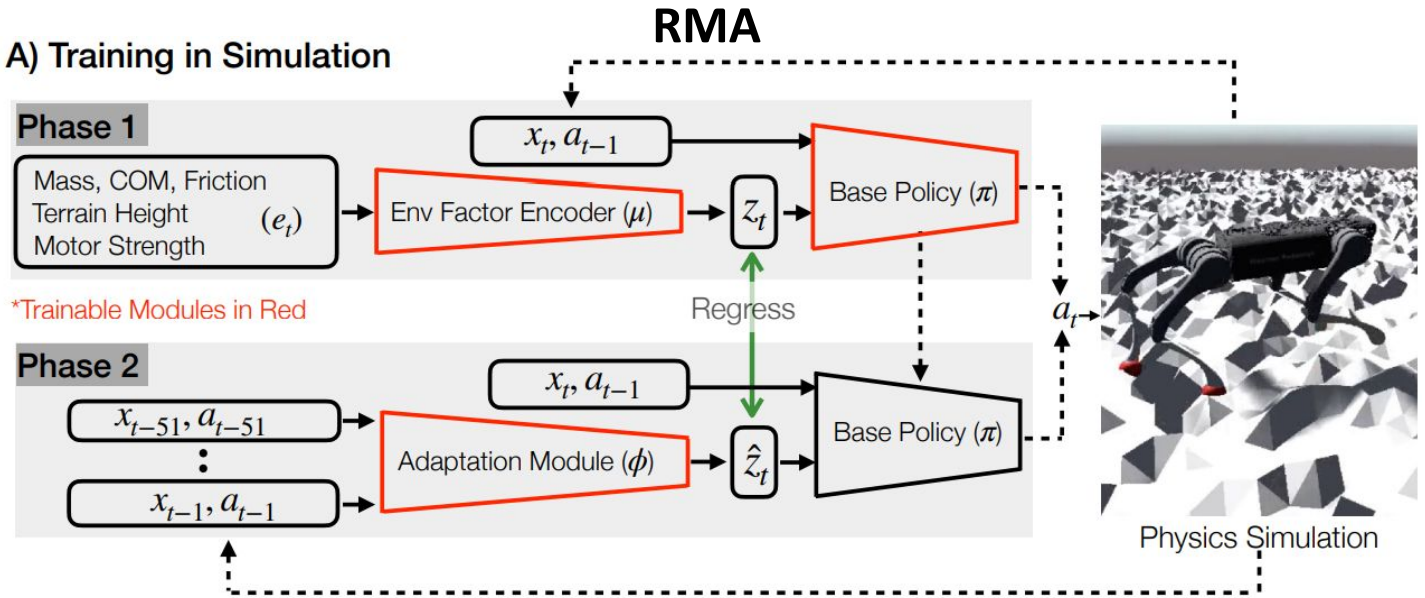
Vision System

- From previous papers, we know that proprioception can get an accurate estimate of the terrain -> ***a blind policy***
  - We can not “look-ahead” with proprioceptive model.

 Train a “look-ahead” vision module to predict future proprioceptive estimates, and then supervise the “look-ahead” module with actual proprioception (CMS).

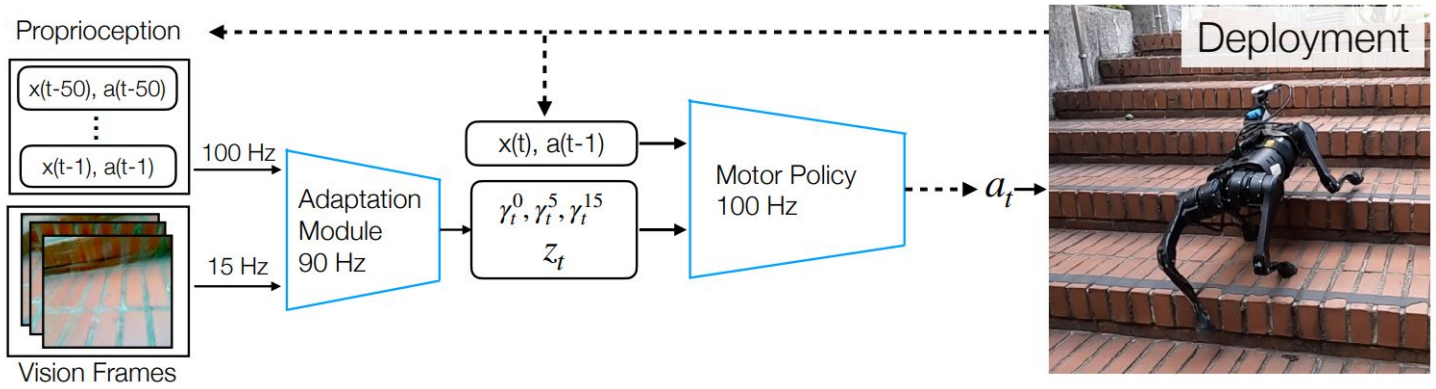


Vision RMA vs. RMA



RMA Base Policy

$$z_t = \mu(e_t)$$
$$a_t = \pi(x_t, a_{t-1}, z_t)$$



Vision RMA Blind Policy

$$\boldsymbol{z}_t = \mu(\boldsymbol{e}_t)$$
$$\boldsymbol{\gamma}_t = \delta(\boldsymbol{h}_t)$$ Encodes GT geometry
$$\boldsymbol{a}_t = \pi_{blind}(\boldsymbol{x}_t, \boldsymbol{z}_t, \boldsymbol{\gamma}_t),$$

Vision System – Blind Policy

- A handcrafted reward function is used to promote the initial blind motor policy
  - Note that this blind motor policy is only trained during stimulation, and it is frozen during deployment.

- Forward: $min({v_x}^d, v_x)$
- Lateral: $\|v_y\|$
- Angular: $-\|{w_z}^d - w_z\| + {w_x}^d$
- Work: $-\|\boldsymbol{\tau}^\top \cdot (\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1})\|$
- Foot Slip: $-\|\text{diag}(g)_t \cdot \boldsymbol{v}_t^f\|$

Promotes the agent to move with a user-defined forward and angular speed.

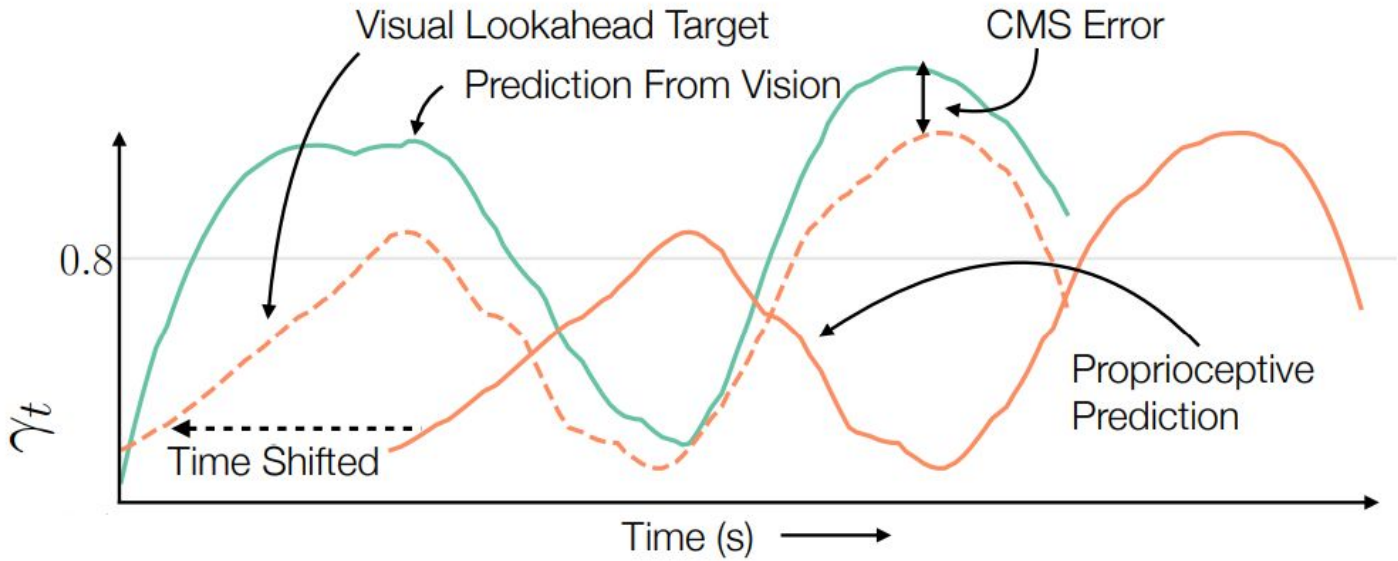
Penalizes lateral speed and jerky motions.

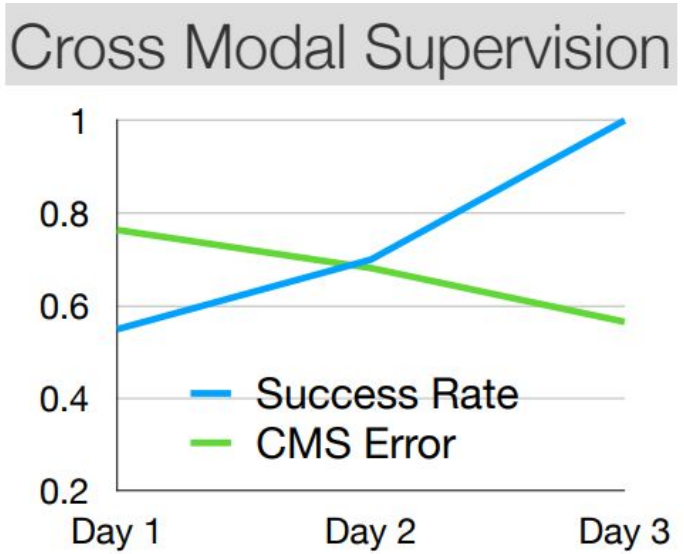
Cross-Model Supervision

- To initialize the CMS improvement procedure, a blind policy $\pi_{\mathbf{blind}}$ is trained in stimulation.
  - Instead of using visual "look-ahead," the blind policy uses proprioception to estimate local geometry.

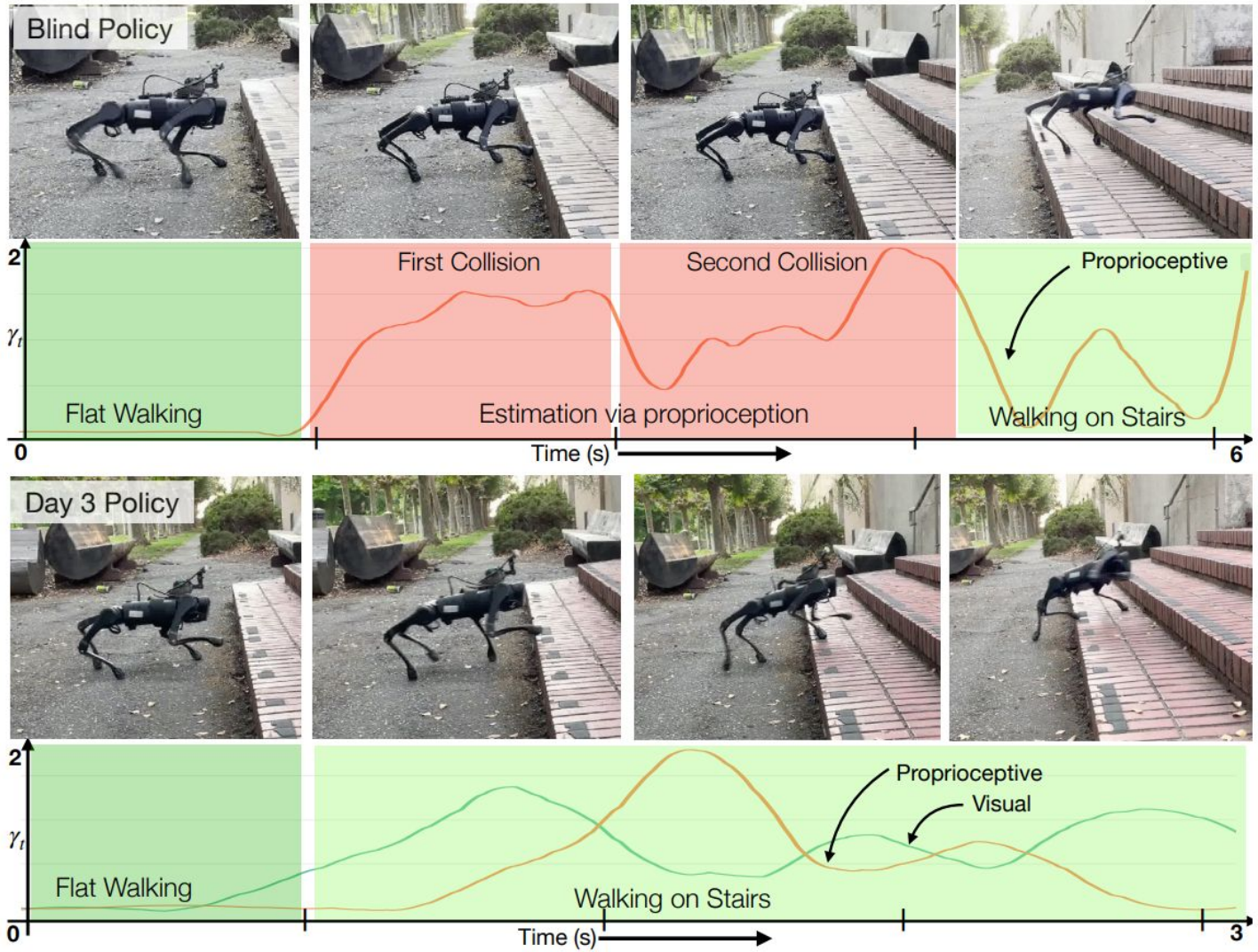
$$\boldsymbol{a}_t = \pi(\boldsymbol{x}_t, \boldsymbol{z}_t, \boldsymbol{\gamma}_t, \boldsymbol{\gamma}_{t+\Delta t})$$

$$\boldsymbol{a}_t = \pi_{blind}(\boldsymbol{x}_t, \boldsymbol{z}_t, \boldsymbol{\gamma}_t),$$





Cross-Model Supervision

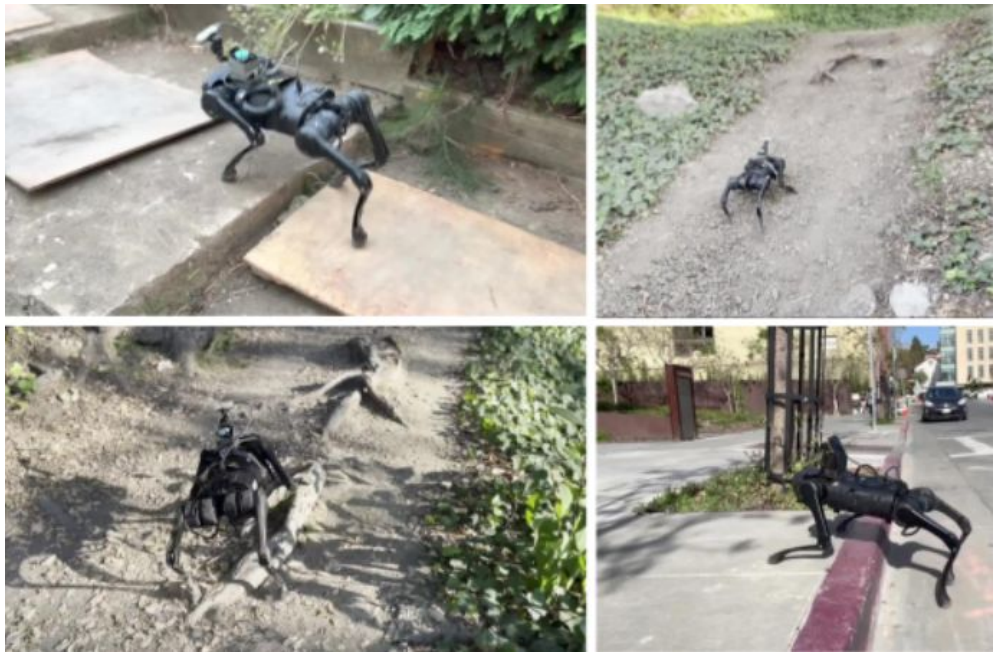


The blind policy is clumsy and exploratory. The robot needs to “tap” the stairs before walking on it

The “look-ahead” policy anticipates the geometry of the terrain and adapts its stride accordingly.

Lifelong Vision Learning

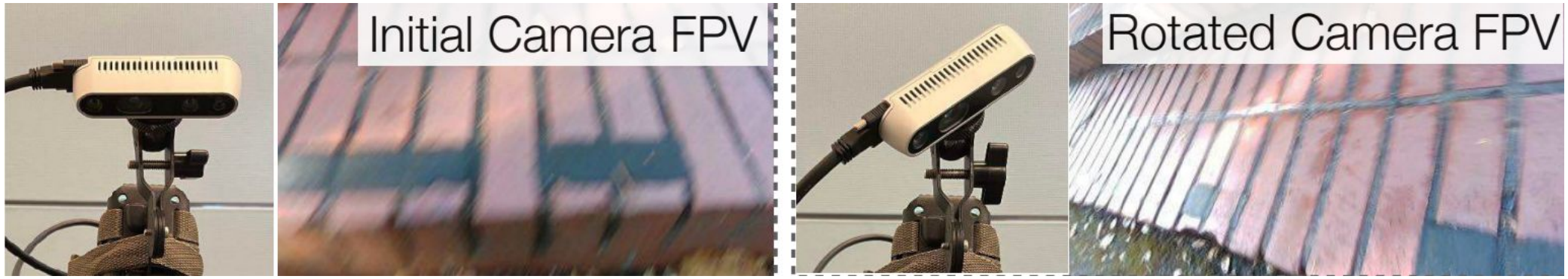
- CMS enables lifelong learning by:
  - *1. continuously collecting data with real-world experience, and improving the quality of the look-ahead predictor*
  - *2. collecting smooth and uncorrupted data (locomotion is smooth due to look-ahead).*



Generalization experiments on challenging terrains (steep inclines, discrete terrains, curbs). The proposed approach achieves high success rate on unseen terrain.

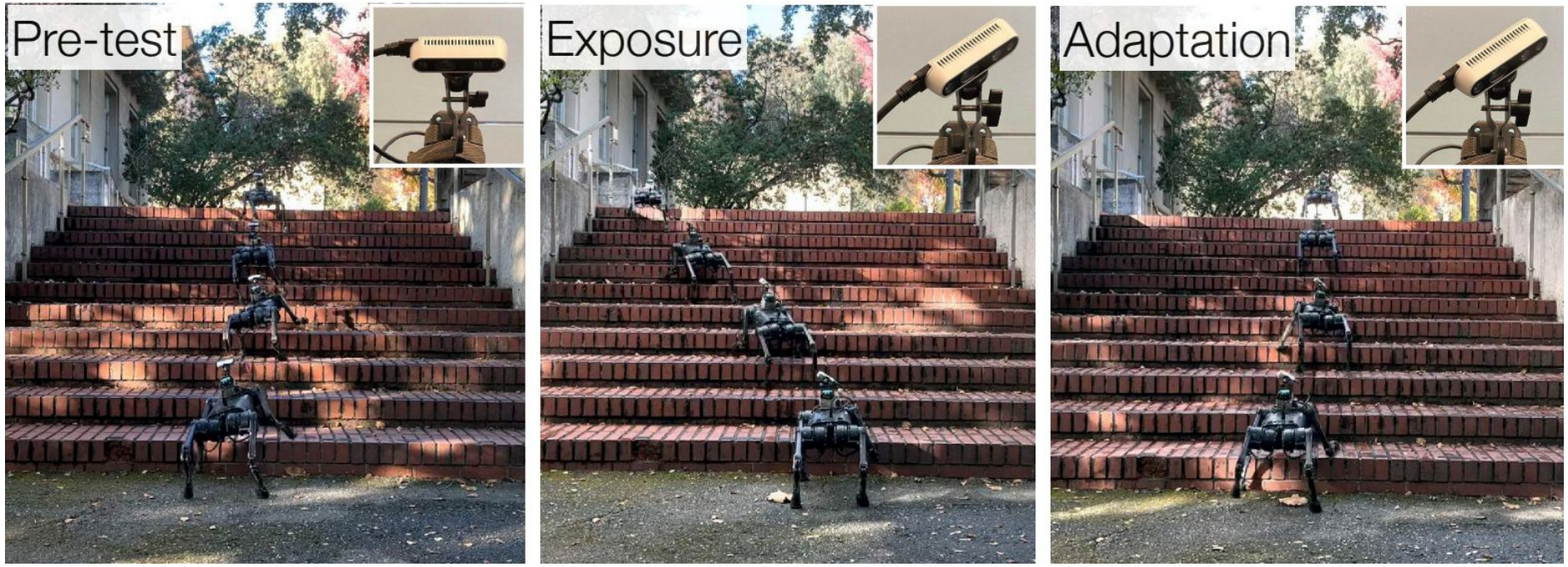
Visual Plasticity

- The prism test:
  - *Pre-test*: the subject performs a task without any disturbance
  - *Exposure:* the subject preforms the same task under a horizontal shift of the visual field.
  - *Adaptation:* the subject adapts to the new visual field and can perform the task at *pre-test* level.
  - Shifts the visual field by rotating the camera on its yaw axis by 30 degree.



Visual Plasticity

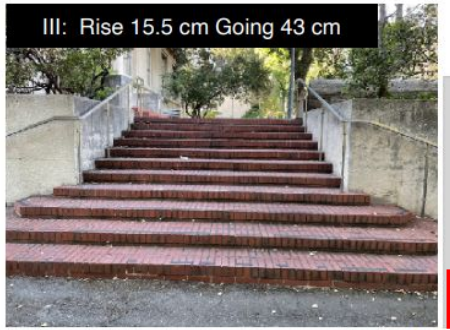
- *Pre-test* vs. *Adaptation*: large variation in the field of view; the robot cannot see the terrain directly in front of it after rotation.
- After training with 80 seconds of experience (*adaptation*), the robot is able to recover its *pre-test* behaviors.



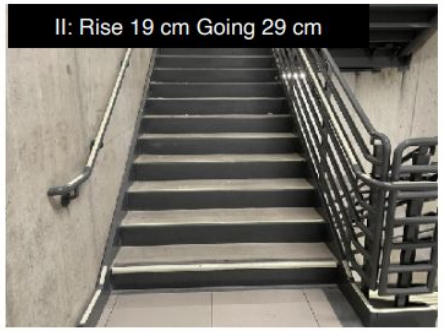
Summary of results

This is ~75% of the robot standing height!

	Success	Distance	TTC
Blind	40%	0.51	10.5 s
Day I	40%	0.71	9.55 s
Day II	60%	0.75	10.4 s
Day III	100%	1	9 s

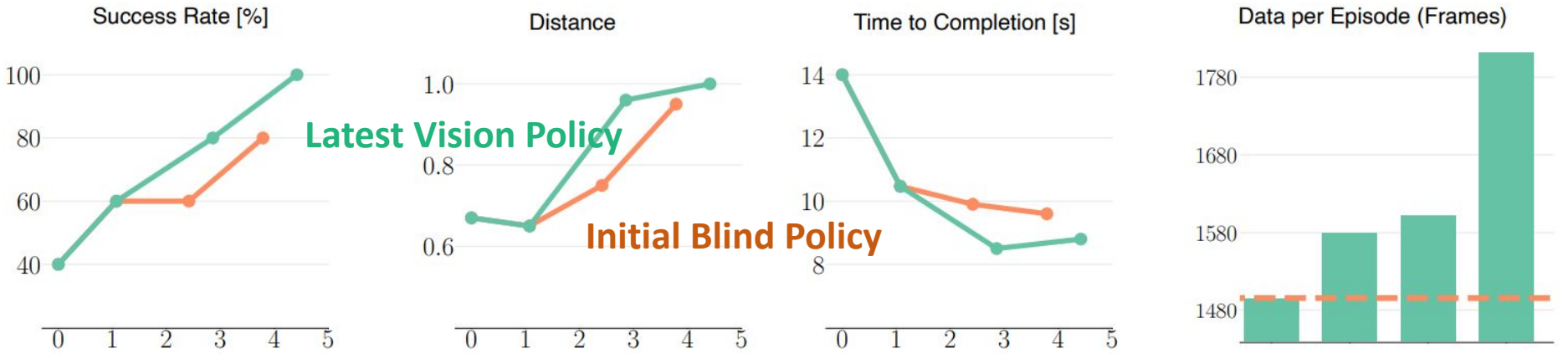


	Success	Distance	TTC
Blind	40%	0.67	14 s
Day I	60%	0.89	13.69 s
Day II	80%	0.81	10.24 s
Day III	100%	1	9 s



	Success	Distance	TTC
Blind	40%	0.58	7.25 s
Day I	60%	0.76	7.44 s
Day II	60%	0.85	6.74 s
Day III	100%	1	6.58 s

	Success	Distance	TTC
Blind	60%	0.82	28 s
Day I	60%	0.85	26.58 s
Day II	80%	0.86	25.6 s
Day III	100%	1	23 s



Discussion #1 - @84_f2

The visual information increase the success rate and speed of walking over obstacles such as stairs. The additional sensor (RGB) camera provides additional information to the motor models. The robots do not need to bump its leg to stairs several times to gain information of its height.

- In the end, the author suggests that the algorithm does not improve the motor system in the real world. What does he mean?

- Can you think of any techniques that may improve the motor policy along with CMS ?

Discussion #2 - @84_f5

Time shifting the proprioceptive latent vector to supervise the vision based system was a really interesting idea.

They mentioned that the final policy uses less than 30mins of real world data collected over 4 days. It would be interesting to see if letting the robot walk on more areas would lead to more generalizability of terrains.

- Is there a reason they only did a few minutes of training each day or what could they do with more training time ?

Discussion #3 - @84_f3, @84_f6

The idea of using a monocular RGB camera as a sensor is more closely related to how human works.

Supplement the policy with vision might contradict the advantages of RMA. The blind model could adjust themselves to sudden change or unexpectable situation, such as throwing a 5 pounds object on it and walking on the terrains with oil, both of which could not be captured perfectly by vision. With the vision input, the adaptation might generate wrong information about the further movements.

- What is your thought about using RGB inputs ?