ICRA 2021 PAPERS OVERVIEW

-

Presented by Yannis He

TABLE OF CONTENT

TABLE OF CONTENT (CONT')

I. DESIGN OF THE HIGH-PAYLOAD FLAPPING WING ROBOT E-FLAP

- Pros of flapping-wing robot:
  - Safe and affordable for rapidly deploying robots around humans and in complex environments
  - Absence of propellers: 1. resistant to physical contact, 2. Able to fly in cluttered environment, 3. Safer to work with
- Cons:
  - Low payload available: 1. high torque required, 2. Airflow over wings is highly turbulent, i.e. hard to model
- Paper's Contribution:
  - Create a 510g flapping-wing robot with a 100% of payload with a high degree of autonomy
    - Flight angle up to 50° and can speed between 2 ~ 6m/s
- Approach:
  - Lightweight integrated electronic framework permitting autonomous flights
  - New wing skeleton with optional camber
  - A modular design for the structure, wings, and the tails

Figure 1. Front view of the flying E-Flap robot during a downstroke.

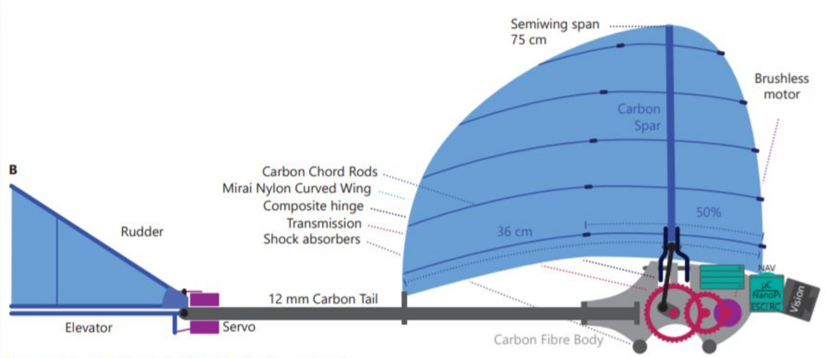
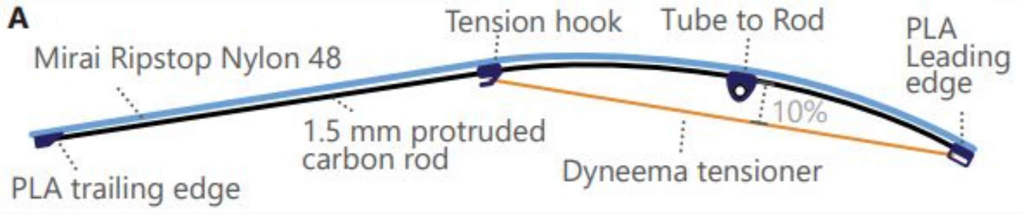


Figure 2. Schematic of the robotic bird with its wing in upwards position.



https://ieeexplore.ieee.org/document/9354891

Presented by Yannis He

2. GROUND-AWARE MONOCULAR 3D OBJECT DETECTION FOR AUTONOMOUS DRIVING

- Context: Estimating 3D position orientation of objects with ONE RGB camera (low-cost) is challenging
- Paper's Approach:
  - Inspiration: monocular 6D pose estimation
  - Utilize ground plane as additional clues in depth reasoning in 3D detection in driving scenes
    - 3D anchors with a deep neural network module through application-specific priors
  - 1 assumption & 2 procedures:
    - Assumption of "important objects are on a ground plane".
    - 1. Anchor filtering:
      - Given a prior distance between an anchor and its distance to the camera, we back-project the anchor to 3D.
    - 2. Ground-aware convolution module:
      - a) identifying contact points to the ground b) compute 3D position c) gather info w.r.t. field focusing downwards
- Contribution:
  - Proposed two networks achieving state-of-the-art performance on the KITTI 3D object detection and depth prediction benchmarks, respectively

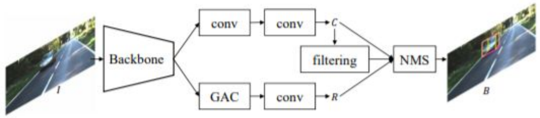


Fig. 2: Network structure for 3D object detection. We extract features from image I and predict classfication tensor C and regression tensor R. We filter anchors far from the ground before post-processing and produce the final bounding boxes B.

Fig. 1: Contact points with the ground plane are important in inferreing 3D information of an object. Predicting depths of background pixels (e.g., the brown point) also rely on the geometry of the ground plane. Best viewed in color.

https://arxiv.org/abs/2102.00690

Presented by Yannis He

3. HOW TO TRAIN YOUR HERON

- Context:
  - Navigation in natural environment requires a robot's behavior model or local planner needs to know how changes in environment impact the system.
  - Most of the current approach requires to measure the state system precisely to depict a robot's dynamics
    - I.e. high cost and hard to embedded to small size system due to the size of sensors
- Contribution:
  - Use Deep Reinforcement Learning and Domain Randomization to solve a navigation task in a natural environment
    - Rely solely on 2D laser scanner & never trained in real world
  - Show that RL agent is more robust, faster, and more accurate than state-aware Model-Predictive-Controller
  - First time DREAMER is applied onto robots
- Approaches:
  - DREAMER, a model-based RL technique
  - Use RL to reduce the amount of interaction with environment during the model training
  - DREAMER uses latent imagination to natively learn the dynamics of the system to build its world model

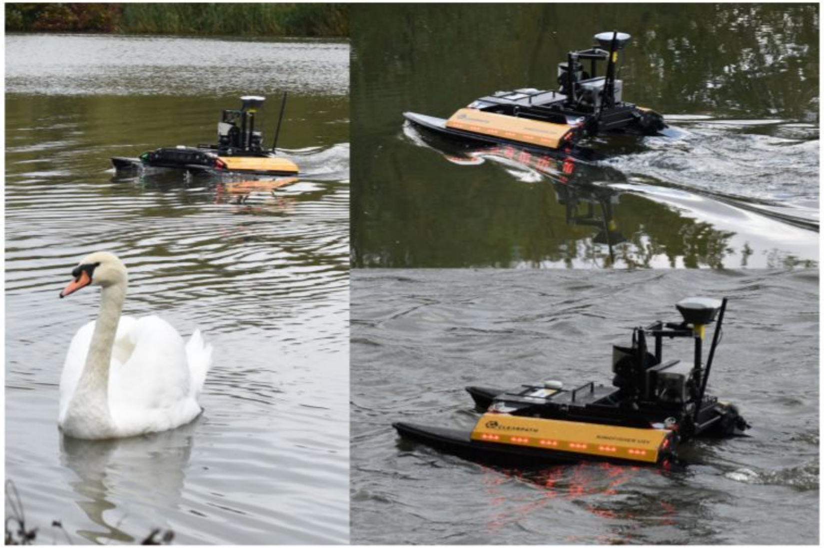
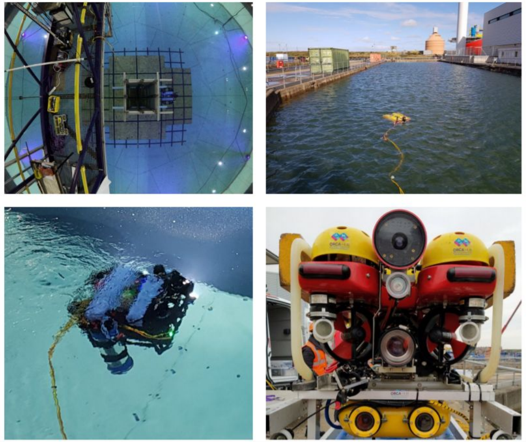
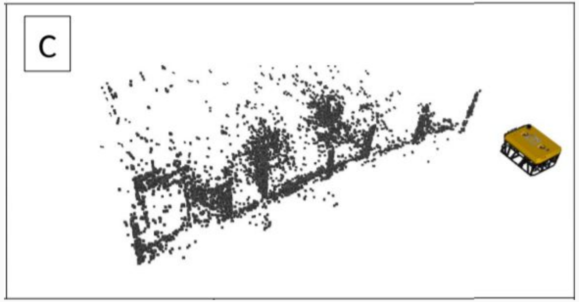


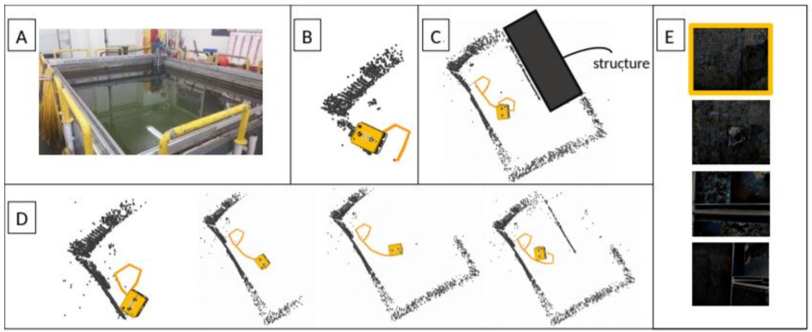
Fig. 1. The Heron dashing around the Symphonie lake.

4. ROBUST UNDERWATER VISUAL SLAM FUSING ACOUSTIC SENSING

- Context: Underwater SLAM is challenging
  - Poor visibility by suspended particles, lack of light, insufficient texture in scenes.
  - Many state-of-the-art approaches rely on acoustic sensing instead of vision
- Contribution:
  - Robust visual SLAM in underwater environments leveraging acoustic, inertial and altimeter/depth sensors
  - Improve the robustness of camera pose estimation in underwater environments
  - Enable the system to create a new map whenever it encounters a new scene where visual odometry can work.
- Approach:
  - Estimate the 6 Degree-of-freedom robot pose from fusion of a Doppler Velocity Log (DVL), gyroscope, and altimeter. This is the first time of such fusion.
    - Unlike a sonar, it produces explicit linear velocity estimate







https://evargasv.github.io/publications/conferences/vargas2021_icra.pdf

Presented by Yannis He

5. LOCOMOTION AND CONTROL OF A FRICTION-DRIVEN TRIPEDAL ROBOT

- Contribution:
  - Presents a omnidirectional gait design and feedback control of a radially symmetric tripedal friction-driven robot.
    - 3 servo motors: for every step, 2 motors rotate the legs and push a robot towards the 3rd leg's direction
    - This design allows the robot to translate in any direction
    - Introduce a Proportional-Integral (PI) feedback control framework that enables the robot to closely follow a desired path.
    - Reduced the tracking error by ~46% with live feedback from an overhead tracking camera
    - Reduced the tracking error by ~65% in a 5.5m/s turbulent wind field
- Benefit of this design:
  - The motion strategy can be easily modified by the detachable limbs
  - No rotation of the body is necessary
  - Radially-distributed limbs also creates more contact points for adhesion, which is important in the case of resisting exogenous forces like fluid forcing

Fig. 1. Minimally-actuated tripedal robot demonstrating curve following capabilities.
Demo: *https://youtu.be/F9UxznYtJGM*

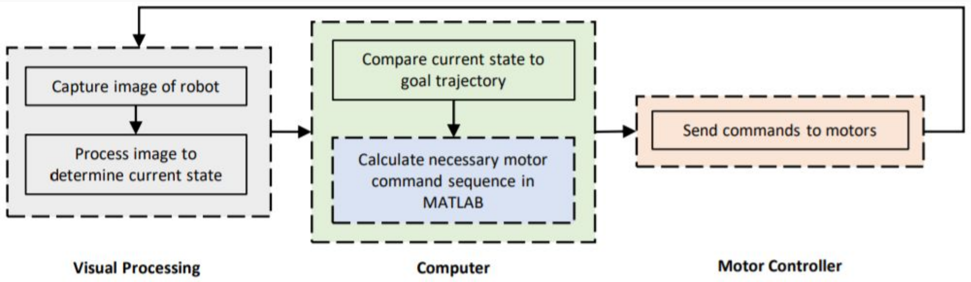
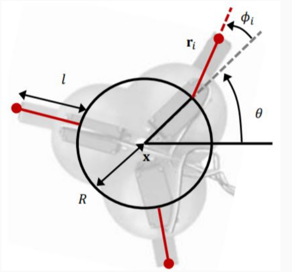


Fig. 2. x is the position coordinate of the center of mass with respect to a fixed reference frame, ξ is the rotation of the body based on the hinge point of a specified limb, and φ_i is the local limb rotation.



https://arxiv.org/abs/2102.10357

Presented by Yannis He

6. SMILE LIKE YOU MEAN IT: DRIVING ANIMATRONIC ROBOTIC FACE WITH LEARNED MODELS

- Context:
  - Able to generate intelligent and generalizable facial expressions is essential for building human-like social robot
  - Current robotic facial expressions are programmed by humans
- Contribution:
  - Designed a physical animatronic robotic face with soft skin and by developing a vision-based self-supervised learning framework for facial mimicry
  - Can generalize well to different human subjects and diverse expressions
  - Algorithm does not require require any knowledge of the robot's kinematic model, camera calibration or predefined expression set.
  - More flexible and stable control
- Approaches:
  - Decomposing the learning process into a generative model
  - Two stages:
    - 1. Use a generative model to synthesize a corresponding robot self-image with the same facial expression, given normalized human facial landmarks,
    - 2. leverage an inverse network to output the set of motor commands from the synthesized image.

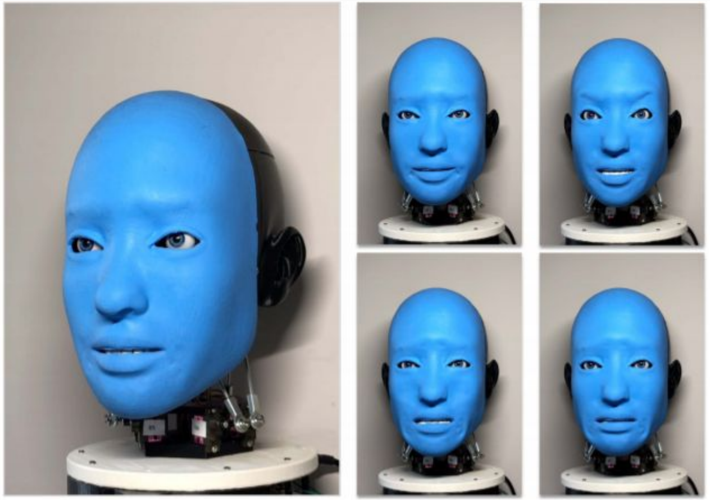


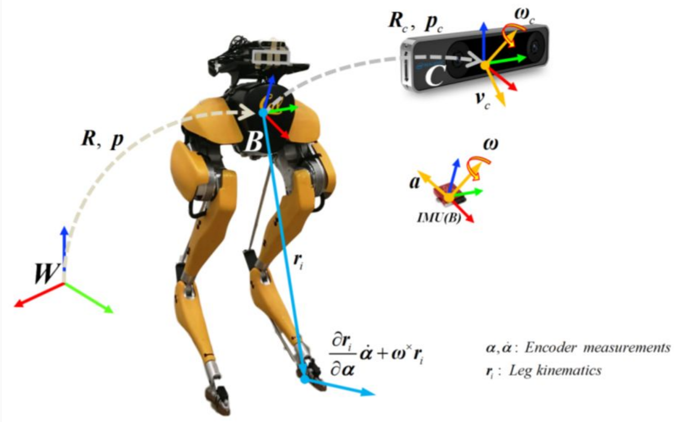
Fig. 1: **Eva 2.0** is a general animatronic robotic face for facial mimicry. The robot does so by learning the correspondence between facial landmarks and self-images as well as a learned inverse kinematic model. The entire learning process relies on the robot's motor babbling in a self-supervised manner. Our robot can mimic varieties of human expressions across many human subjects.

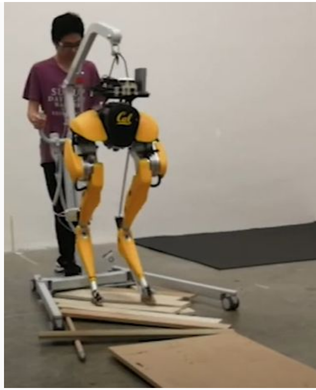
https://arxiv.org/abs/2105.12724

Presented by Yannis He

7. LEGGED ROBOT STATE ESTIMATION IN SLIPPERY ENVIRONMENTS USING INVARIANT EXTENDED KALMAN FILTER WITH VELOCITY UPDATE

- Contribution:
  - Propose a state estimator for legged robot operating in slippery environment
  - Implemented an Invariant Extended Kalman Filter (InEKF) to fuse inertial and velocity measurements from a tracking camera and leg kinematic constraints
  - Develop an online noise parameter turning method to adapt to the highly time-varying camera measurement noise
  - More robust to consistent slippages than prior methods
- Approaches:
  - Model the misalignment between the camera and the robot-frame to auto-calibrate camera pose
  - Formulate a right-invariant observation for the velocity-based leg kinematics measurement
  - Use a probabilistic method for slippage and contact detection since the force sensors and accelerometers are vulnerable to ground impact
- Discoveries:
  - Nonlinear observability analysis shows that other than the rotation around the gravity vector and the absolute position, all states are observable except for some singular cases.
  - Discrete observability analysis demonstrates that our filter is consistent with the underlying nonlinear system.
- Demo: https://youtu.be/VIqJL0cUr7s





https://arxiv.org/pdf/2104.04238.pdf

Presented by Yannis He

8. Human-like Artificial Skin Sensor for Physical Human-Robot Interaction

- Context:
  - Most of the robotic devices embed sensors rarely considers human factors
- Contribution:
  - Propose an approach to design and fabricate compliant Human-like artificial skin sensors for robots, with similar mechanical properties as human skin and capable of precisely detecting touch
  - Present the sensor & describe its fabrication process
    - Scalable, low-cost
    - Ensures flexibility, compliance and robustness
  - Introduce an open-source sensing development board, *Muca*
- Approaches:
  - The artificial skin relies on the use of different silicone elastomers to replicate the human skin layers
  - Comprises embedded electrodes matrix to perform mutual capacitance sensing

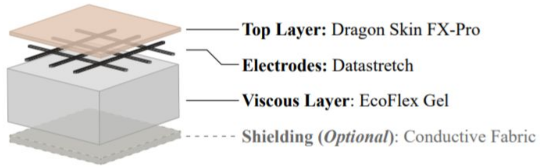


Fig. 2: Layered structure of the sensor. The sensing electrodes are situated in-between two layers of silicone with different mechanical properties.

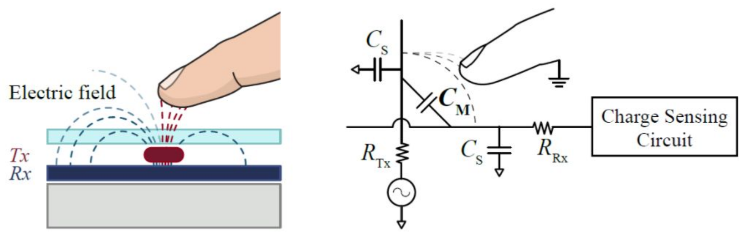
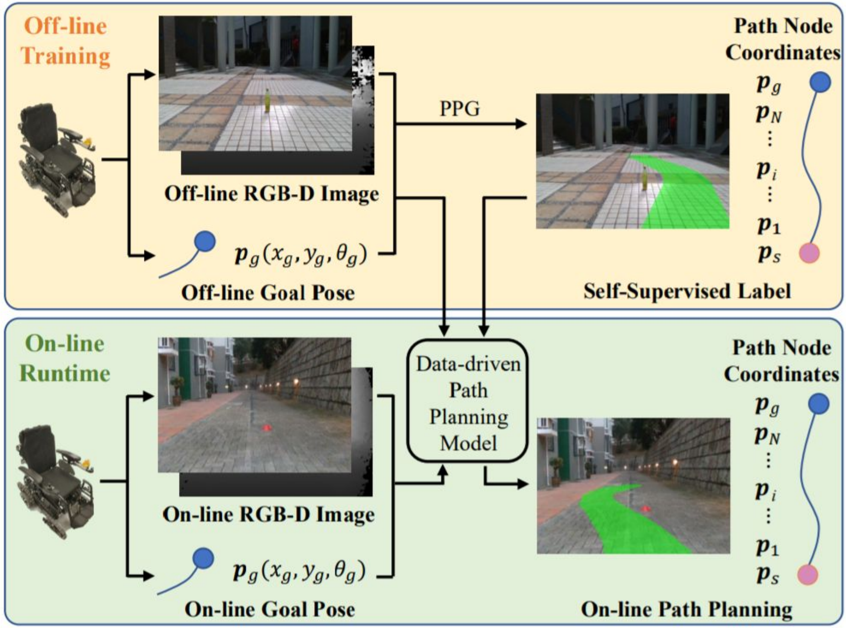


Fig. 11: The sensor can be shaped to conform to any robot. Here, a skin sensor is attached to a *Nao* social robot.

9. S2P2: Self-Supervised Goal-Directed Path Planning Using RGB-D Data for Robotic Wheelchairs

- Context:
  - Path planning is a fundamental capability for autonomous navigation of robotics wheelchairs
  - Disadvantage of deep-learning + imitation learning-based path planning approach:
    - Need extensive time and labor to record export demonstration and data
    - Existing approaches could only receive high-level commands, such as turning left/right, which is less sufficient for the navigation of mobile robots, where usually require exact poses of goals
- Contribution:
  - Proposing S2P2: a self-supervised goal-directed path planning approach
  - Pipeline to automatically generate planned path labels given RGB-D images and poses of goals
  - Best-fit regression plane loss to train the data-driven path planning model

Fig. 1: The robotic wheelchair used in this work. It is equipped with an Intel RealSense RGB-D camera to collect data and an NVIDIA Jetson TX2 to run our S2P2 model.

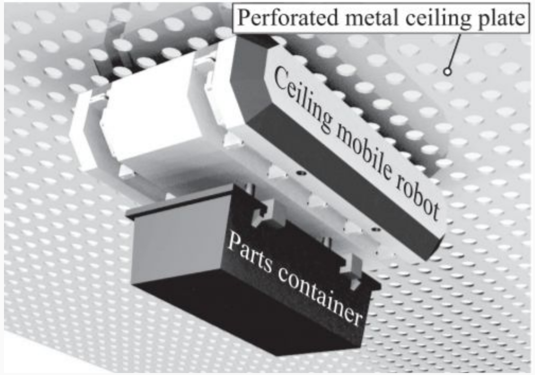


https://arxiv.org/abs/2103.10210

Presented by Yannis He

10. HanGrawler: Large-Payload and High-Speed Ceiling Mobile Robot Using Crawler

- Contribution:
  - Present a ceiling-mounted mobile robot, HanGrawler, with high-speed mobility
  - Able to freely select and adjust its route under a ceiling plate
  - Travel linearly at 0.1 m/s, turn at 8.5°/s, carry a maximum load of 60kg
- Approaches:
  - HanGrawler hangs from the holes of a perforated metal ceiling plate using newly developed mechanically constrained hanging mechanisms mounted on crawler-type traveling equipments
  - The mechanisms continuously grasp the ceiling holes while moving in a straight line
    - Continuously move by locking onto and releasing from the ceiling plate using the hanging mechanism
  - Use a mechanism that achieves both contact and fixing between the hanging mechanisms and ceiling plate



Demo: https://youtu.be/bqcb-qD5IcI

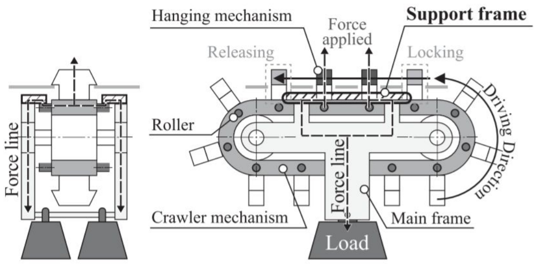


Fig. 4. Conceptual sketch of support frame.

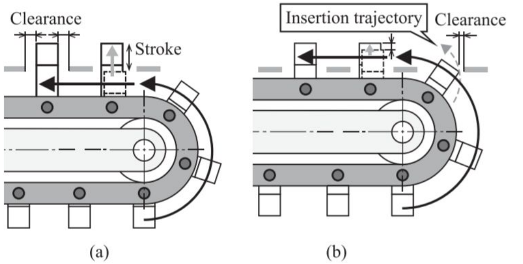
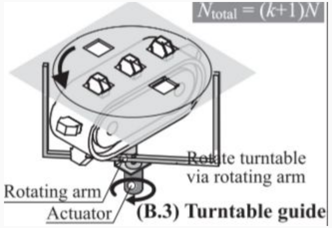


Fig. 5. Hanging-mechanism trajectory.



https://ieeexplore.ieee.org/ielx7/8860/9159950/09018154.pdf