2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Inchworm-inspired Adaptive Multimodal Neural Control for an Autonomous Inspection Robot

Wasuthorn Ausrivong[1], Arthicha Srisuchinnawong[1], Poramate Manoonpong[1,*]

***Abstract*—While inspection robots currently achieve autonomous locomotion and transitions using computationally expensive multi-dimensional data, such as point-cloud maps, invertebrates like inchworms can navigate complex tree branches effortlessly with a simple brain and little information. Inspired by this, an Inchworm-inspired Adaptive Multimodal Neural Control (IAMNC) is presented here. It relies on only nine exteroceptive sensors to achieve autonomous locomotion and gait transitions of a bipedal out-pipe inspection robot. The IAMNC consists of three interpretable modules: Mode Selection (MoSe) for autonomous mode selection, Locomotion Control (LoCo) for adaptive crawling, and Transition Control (TraCo) for adaptive transitions. Additionally, this work also proposes interpretable Adaptation Units (AUs, part of TraCo), which can be configured to obtain different functions (e.g., reflex, memory, and logic gate operations). Thus, connecting them as an Adaptation Unit Network (AUN) results in complex yet understandable adaptation signals for autonomous and stable gait transitions. With this control approach, the robot demonstrates autonomous locomotion, gait transitions, and adaptability across various pipe connection types. It can also autonomously step over an obstacle on the pipe, relying on seven infrared obstacle detection sensors (IR sensors) and two inductive sensors, rather than a computationally expensive camera sensor.**

***Index Terms*—Neural control, Climbing robots, Bio-inspired robotics, Pipe inspection, Inspection robots, Legged robots, Gait transition, Multi-modal locomotion.**

I. Introduction

INSPECTION ROBOTS are generally categorized as wheel-based or leg-based. While wheel-based robots achieve high speeds [1], they struggle with tall obstacles, non-planar pipe connections [1], and complex configurations (Fig. 1(a)). In contrast, leg-based robots trade speed for enhanced adaptability in navigating such challenges [2–6]. Prior studies have shown their ability to traverse 90° pipe connections [2–6] and step over obstacles or flanges [2, 6] using autonomous or semi-autonomous control.

Semi-autonomous robots [2, 6, 7] rely on a mix of manual commands and programmed trajectory generation. The 3DCLIMBER [7] executed predefined trajectories but lacked adaptability to environmental uncertainties (see Table I in Section V). AVIS [2] had a certain degree of adaptability but still required human intervention for obstacle crossing and manual gait transition control, which can lead to failure due to human errors. PiRob [6] used visual-based locomotion but needed repeated surface scans (at least 10 times), making the operation inefficient.

[1]All authors are with the school of Information Science and Technology, Vidyasirimedhi Institute of Science and Technology, Rayong, Thailand.
*PM is the corresponding author (e-mail: poramate.m@vistec.ac.th).

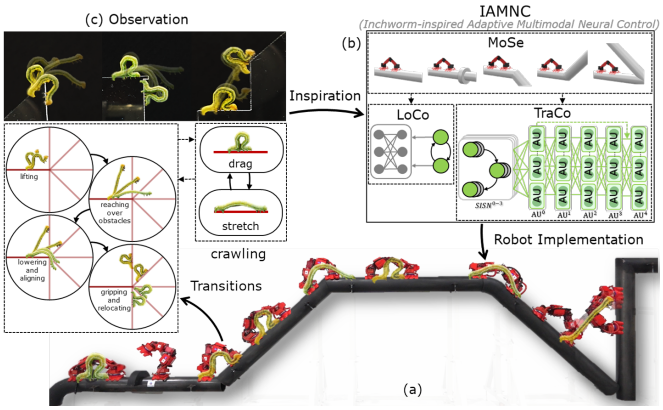

Fig. 1. (a) Physical AVIS-II robot, demonstrating autonomous multimodal locomotion (i.e., crawling on horizontal, vertical, and sloped pipes and obstacle crossing) and gait transitions on a complex industrial pipeline. (b) The Inchworm-inspired Adaptive Multimodal Neural Control (IAMNC) for generating autonomous locomotion and gait transitions in AVIS-II robot. (c) Inchworm behaviors observed on acrylic platforms with different slopes as an inspiration for the design of the IAMNC. A video demonstrating autonomous locomotion of AVIS-II is available at https://youtu.be/TiLU2V0leHs.

Fully autonomous robots [3–5] utilize depth image-based mapping for path planning (see Table I in Section V). However, structural complexity and control errors can distort maps, requiring frequent re-planning [3–5]. Despite verification methods such as point-cloud visualization [3–5], interpreting large datasets remains impractical [8], limiting reliability in critical applications such as oil and gas pipelines.

While robots rely on computationally expensive, high-dimensional feedback to navigate complex and discontinuous structural surfaces, inchworms, simple animals [5, 9], effortlessly traverse intricate tree branches. As can be observed in Fig. 1(c), they use proprioceptive feedback for adaptive crawling and obstacle crossing, while exteroceptive feedback (e.g., environmental contact area-based tactile sensing) aids frontal and surface detection. During crawling, two key steps are observed: (C1) bending (drag) pose and (C2) extending (stretch) pose. During transitions, four key steps are observed: (T1) lifting and reaching over obstacle, (T2) lowering and adaptively aligning, (T3) gripping the new surface, and (T4) relocating and resuming normal crawling, as shown in Fig. 1(c).

Inspired by this behavioral observation, we use the inchworm's simple yet effective and versatile locomotion control strategy as a principle for developing an interpretable neural locomotion controller using minimal sensory feedback for autonomous multimodal robot locomotion. To achieve this,

an Inchworm-inspired Adaptive Multimodal Neural Control (IAMNC) is developed with nine input signals: three IR sensors for frontal detection and six (four IR, two inductive) for surface detection and foot alignment. The key contributions of this work include:

1) A novel Inchworm-inspired Adaptive Multimodal Neural Control (IAMNC) for autonomous multimodal locomotion with crawling and obstacle crossing, as well as gait transitions on complex pipeline configurations,

2) A general Adaptation Unit (AU) for providing basic primitive neural functions, which can be combined to form an Adaptation Unit Network (AUN) to generate more complex robot locomotion behaviors, and

3) The real-world demonstration and performance evaluation of a bipedal out-pipe inspection robot for autonomous multimodal locomotion and gait transitions, using a minimal set of sensory feedback, its achieved capabilities surpass those of related robots, as detailed in Section V.

II. AVIS-II Bipedal Out-pipe inspection robot

In this study, we use the AdVance pipe Inspection robot System-II (AVIS-II), presented in Fig. 2, as a robot platform for revealing control development and experiments. It is an upgrade of the previous version [2]. The robot structure is made of aluminum, while the cover consists of PLA plastic. It has six revolute joints, driven by XM540-W270-R Dynamixel servo motors at the joints M1, M2, M4 and XM430-W350-R Dynamixel servo motors at the joints M0, M3, M5. It uses two electromagnets (Magma-4230035) installed at its feet to attach to metal pipes, with two 3D-printed ABS foot caps and silicone rings for improved stability and adhesion.

The robot is equipped with 16 input signals in total. Seven signals provide proprioceptive information. One roll signal from an inertial measurement unit (IMU) is used for locomotion with adaptive balancing, and six motor position signals for low-level motor position control (Sections III-B and III-C). Another nine signals provide exteroceptive information obtained from seven IR sensors and two inductive sensors. Three frontal IR sensors (IR_{0-2}) detect obstacles located in the front (Fig. 2(b)). Thus, they are used to classify operating pipe conditions and activate corresponding locomotion modes (Section III-A). Apart from that, four IR sensors (IR_{3-6}) at the feet and two inductive sensors (IND_{0-1}) beneath the feet (Fig. 2(c)) are responsible for foot alignment and surface detection. All these sensory signals are used to generate inchworm-inspired adaptive actions (Section III-C).

Raspberry Pi 4 functions as the main controller to implement the IAMNC (Section III) for locomotion control with an update frequency of 30 Hz. The main controller interfaces with the sensors via an Arduino Nano, the motors via a U2D2, and the electromagnets via MOSFETs. The robot is supplied with 24 V power from a power supply or a battery pack, subsequently converted to 12 V for the actuators and 5 V for other electronic components by the Pololu step-down voltage regulator. Taking into account all components, the robot weighs 4.5 kg and measures 15 cm × 46 cm × 29 cm (W × L × H).

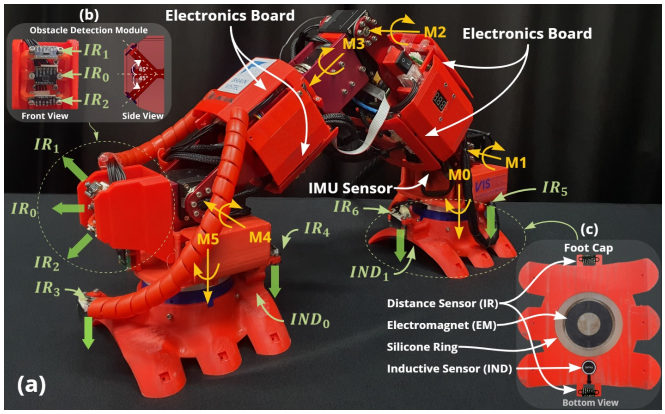

Fig. 2. (a) Physical AVIS-II bipedal robot, where the motor/joint rotation axes are indicated by yellow arrows, the direction of IR-based distance sensors (IR_{0-6}) and inductive sensors (IND_{0-1}) indicated by green arrows, and the components (i.e., electronic components, sensors, and foot caps) indicated by white arrows. (b) Front view of the obstacle detection module. (c) Bottom view of the foot cap.

III. Inchworm-inspired Adaptive Multimodal Neural Control (IAMNC)

To generate autonomous multimodal locomotion and gait transitions of the bipedal robot, the Inchworm-inspired Adaptive Multimodal Neural Control (IAMNC) is developed (Fig. 3). The IAMNC consists of three modules: Mode Selection (MoSe) for autonomous mode selection, Locomotion Control (LoCo) for adaptive crawling, and Transition Control (TraCo) for adaptive gait transitions.

A. Mode Selection Module (MoSe)

To classify different pipe conditions, three frontal IR sensors[1] (Fig. 2(b)) pointing forward (ir_0), upward (ir_1), and downward (ir_2) are employed using a rule-based technique, which allows robust handling of slight variations of pipe conditions. Their signals are smoothed by a low-pass filter.

Initially, the physical robot is placed under various conditions (i.e., obstacle-free pipe, a flange, and -45° to 135° transitions) as shown in Fig. 4(a), to acquire the data for designing the classification rules. Since the recorded data exhibit clear separation, the classification rules are then chosen empirically, as depicted in Fig. 4(b). First, the low-pass filtered forward signal ($ir_0[t]$) is mainly used to detect obstacles and the upward pipe transitions (i.e., 90°–180° transitions). Second, the low-passed filtered upward signal ($ir_1[t]$) is mainly used to distinguish between the obstacle and transitions above 90°. Third, the low-pass filtered downward signal (ir_2[t]) is mainly used to identify downward transitions (i.e., -90°–0°). For each classification condition, a specific operation mode (OM) is set as illustrated in Fig. 4(b) and subsequently used to activate the corresponding modules. For example, when the OM is set to 1, the LoCo module

[1]The IR sensor signal is zero when no object is detected.

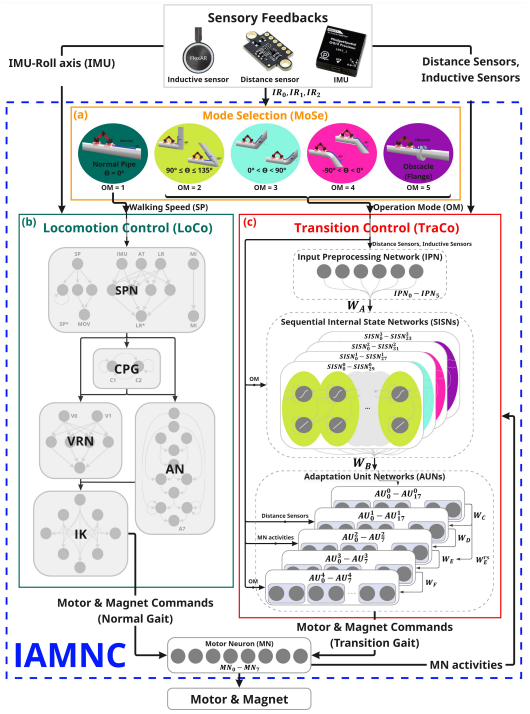

Fig. 3. An overview of the Inchworm-inspired Adaptive Multimodal Neural Control (IAMNC) comprising three modules. (a) Mode Selection (MoSe) presented within the orange box. (b) Locomotion Control (LoCo) presented in the dark green box. (c) Transition Control (TraCo) presented in the red box. The colors of the circles within the MoSe module correspond to those of the SISN network in the TraCo module and the LoCo module to signify the specific network utilized for different pipe connection types.

(described in III-B) is enabled, with the locomotion speed (SP) set to 0.15, while the entire TraCo module is disabled. In contrast, when the OM is set to 2, the LoCo module is disabled (locomotion speed (SP) set to 0.0), and the first SISN in the TraCo module ($SISN^0$ submodule, described in Section III-C2, for $90° \leq \theta \leq 135°$ transitions) is enabled, where θ is the transition angle.

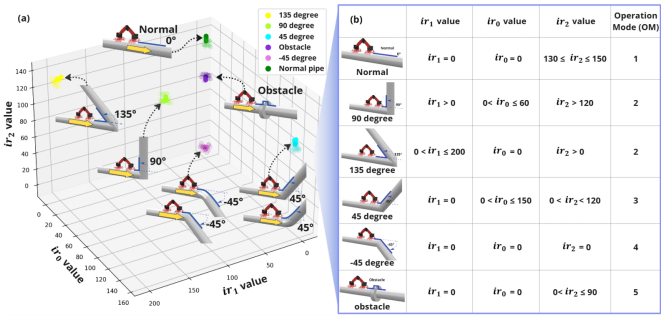

Fig. 4. (a) Six areas corresponding to six classification conditions: one for an obstacle-free pipe or normal pipe (dark color), one for an obstacle (purple) and the remaining four for different pipe configurations and transition angles, namely 45° transition (blue), 135° transition (yellow), 90° transition (light green) and -45° transition (pink). Each of these areas shows sensor data points from three IR sensors (ir_{0-2}) at the robot leg, pointing forward, upward, and downward. (b) Illustrates the five operation modes (OMs) based on the three sensor signals.

B. Locomotion Control Module (LoCo)

To produce an adaptable rhythmic locomotion command, the neural control network in [2] is employed as LoCo. It consists of five sub-neural modules/networks (Fig. 3(b)). First, the Signal Preprocessing Network (SPN) preprocesses sensory feedback (roll angle from the IMU) and the speed commands (SP, from the MoSe module) before transitioning to the following sub-modules. Second, a Central Pattern Generator network (CPG) generates two periodic signals that mimic the inchworm's crawling motion, i.e., (C1) bending and (C2) extending poses. One signal drives body stretching, while the other signal is used to activate the electromagnets. Third, a Velocity Regulating Network (VRN) amplifies or deamplifies the stretching pattern according to the variable multiplication gain, i.e., preprocessed speed command selected as 1 cm/s to match the robot's sensing response. Thus, a higher value leads to faster speed. Fourth, an Adaptation Network (AN) produces sideways adaptation signals to correct robot orientation or retain balance. Fifth, an Inverse Kinematics network (IK) combines the stretching pattern from the VRN with the adaptation pattern from the AN using IK for accurate adaptation. Lastly, the output is fed into the motor neurons (MN). Further details can be found in [2].

C. Transition Control Module (TraCo)

To achieve agile inchworm-inspired transition strategies (Fig. 1(c)) for stable transitions between pipe segments, the control should obtain both adaptive sequential state patterns and adaptive actions based on sensory feedback. The TraCo developed here uses a combination of three types of sub-modules, as shown in Fig. 3(c). First, the Input Preprocessing Network (IPN) preprocesses input sensory feedback signals (e.g., IR sensors and inductive sensors), which are then used as trigger signals to change the corresponding internal states in Sequential Internal State Networks (SISNs). Second, the SISNs generate sequential state patterns, which were tuned to resemble inchworm transition sequence (T1–T4) shown in Fig. 1(c) and are then fed into Adaptation Unit Networks (AUNs). Third, the AUNs are responsible for generating action patterns and adaptive behaviors, e.g., inchworm-like contact search and alignment (T2). The AUNs utilize five layers of AUs. Each internal state output neuron in the SISN connects to multiple neurons corresponding to the first layer (AU^0), which serve as primitive action patterns. The other layers (AU^{1-4}) incorporate sensory feedback (e.g., IR sensors) for adaptive movements (e.g., foot alignment before touching the pipe, see M4 in Fig. 7(a)), motor feedback (e.g., MN activities) for adaptive reflexes (e.g., extensor or flexor reflex signals, see au^3_{4+} signal in Fig. 7(c)), recurrent neural feedback for memory, and OM for switching gait transitions, respectively.

*1) **Input Preprocessing Network (IPN)**:* The IPN functions as a sensory feedback preprocessing layer, rescaling the sensory feedback inputs and setting their thresholds before feeding them to change the internal state of SISN sub-modules. The IPN comprises six neurons (IPN_{0-5}), receiving

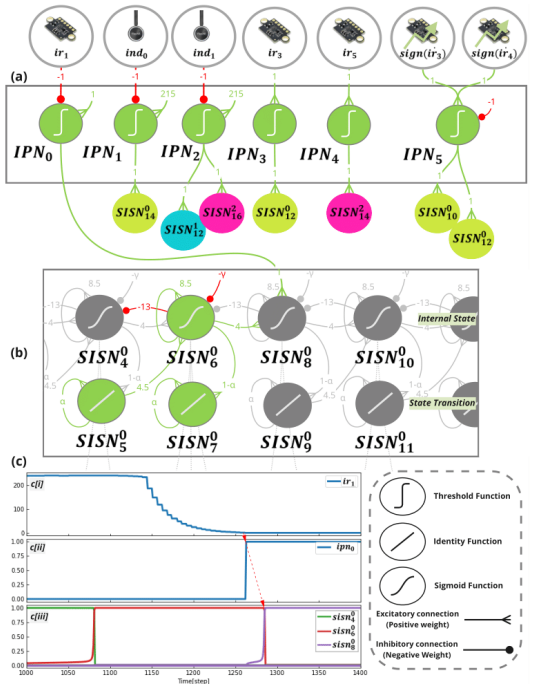

Fig. 5. (a) Neural circuit of the Input Preprocessing Network (IPN). (b) Neural visualization of part of the first Sequential Internal State Network (SISN^0). (c) The output signals of the SISN^0 network with respect to the IPN outputs and sensory feedback (e.g., $ir_1[t]$) as shown in c[i] to c[iii].

seven dimensions of sensory feedback (**fb[t]**): five IR sensor signals ($ir_{1,3,5}[t]$ and sign$(ir_{3,4})[t]$), and two inductive sensor signals ($ind_{0,1}[t]$), as shown in Fig. 5(a). The activities of IPN neurons are governed by: $\mathbf{ipn[t]} = \tau\left(\mathbf{W_{fb}^{ipn}}\mathbf{fb[t]} + \boldsymbol{\nu}\right)$, where **ipn[t]** denotes the activity vector of IPN neurons at timestep t ($\mathbf{ipn[t]} = [ipn_0[t]\, ipn_1[t] \ldots ipn_5[t]]^T$) and **fb[t]** denotes the activity vector of the sensory feedback. $\tau()$ denotes the threshold function. The parameter $\mathbf{W_{fb}^{ipn}}$ denotes the connection weight matrix from the sensory feedback to the IPN neurons, while $\boldsymbol{\nu}$ denotes the bias vector of the IPN neurons. These parameters are determined empirically: the biases are set as thresholds to determine their activations, while the weights are set to rescale the sensor values, as illustrated in Fig. 5(a). The IPN network outputs are then fed to the SISNs as shown in Fig. 5(a) through the mapping matrix $\mathbf{W_A}$ as shown in Fig. 3(c). Taking the IPN$_0$ neuron shown in Fig. 5(c) as an example, when the sensor value ($ir_1[t]$, see Fig. 5(c[i])) reaches a defined threshold, the neuron is then activated, subsequently changing the activation of the internal neural state from $sisn_6^0[t]$ (inactive) to $sisn_8^0[t]$ (active) (Fig. 5(c[iii])) in the SISN network.

*2) **Sequential Internal State Networks (SISNs):*** Receiving the triggering signals from the IPN, the SISNs produce sequential internal states and their outputs, which are later used for producing sequential actions. The SISNs contain four sub-networks (SISN$^{0-3}$), which are designed to deal with four pipe conditions: (1) SISN$^0_{0-29}$ (30 neurons) for $90° \leq \theta \leq 135°$ transitions, (2) SISN$^1_{0-27}$ (28 neurons) for $0° < \theta < 90°$ transitions, (3) SISN$^2_{0-31}$ (32 neurons) for $-90° < \theta < 0°$ transitions and (4) SISN$^3_{0-23}$ (24 neurons) for obstacle crossing. Based on our previous work [2], all sub-networks share the same basic structure. However, they produce different action sequences. Each sub-network is further divided into two neural groups: the internal state neurons in even numbers (i.e., SISN$^{\text{i}}_{2\text{n}}$) and the state transition neurons in odd numbers (i.e., SISN$^{\text{i}}_{2\text{n+1}}$), as shown in Fig. 5(b). The internal state neurons produce a sequential state transition according to: $n_i[t] = \sigma\left(\lambda n_{\text{trig}}[t] + \beta n_{i-2}[t] + \gamma n_i[t-1] - \eta n_{i+2}[t] - \zeta\right)$.

To fulfill the desired behavioral constraints, the network parameters β, η, γ, and ζ are empirically chosen as 4.0, 13.0, 8.5, and 8.0, respectively detailed in [2]. $\sigma()$ denotes the sigmoid activation function. $n_{\text{trig}}[t]$ can be either an internal state transition (e.g., SISN$^0_5$ in Fig. 5(b)), where λ is set to 4.5, or external feedback (e.g., IPN$_0$ in Fig. 5(b)), where λ is set to 1. The connection weights from the IPN neurons to SISN$^{\text{i}}$ neurons are shown in the $\mathbf{W_A}$ table[2]. The internal state activities $n_{\text{even}}[t]$ are passed to the corresponding state output neurons $n_{\text{odd}}[t]$, where a low-pass filter-like function is applied to smooth the output signals, as described in: $n_i[t] = \sigma\big(\alpha n_i[t-1] + (1-\alpha) n_{i-1}[t]\big)$, where $n_i[t]$ indicates the activity of neuron n_i at time step t, with the transition speed gain α set to a low value of 0.1 to ensure slow and stable transitions. These neuron activities are subsequently used as autonomous signals to trigger the activation of the preceding layer ($n_{\text{trig}}[t]$) and also fed into the corresponding pattern neurons in the AUN to produce sequential motor patterns. In addition, to reset the robot states to normal walking and unblock the MoSe module, the activities of the last two internal state neurons are used as the OM resetting signal, while the last state serves as an initial state waiting to re-enable the network for the upcoming transition sequences.

*3) **Adaptation Unit Networks (AUNs):*** To mimic the adaptive actions of inchworms, AUNs are developed. Each network consists of multiple general units, referred to as Adaptation Units (AUs), each of which is designed to incorporate adaptation into actions, as illustrated in Fig. 6. AN$^{\text{i}}_{\text{j}}$ refers to the j^{th} AU located at layer i.

Each AU consists of two neurons: positive AU ($au^{\text{i}}_{\text{j+}}[t]$) and negative AU ($au^{\text{i}}_{\text{j-}}[t]$), represented together as the activity vector, $\mathbf{au^i_j[t]} = [au^{\text{i}}_{\text{j+}}[t]\, au^{\text{i}}_{\text{j-}}[t]]^T$. Besides, each AU receives three inputs: two feedback inputs ($fb^{\text{i}}_{\text{j0}}[t]$ and $fb^{\text{i}}_{\text{j1}}[t]$) and one input ($ai^{\text{i}}_{\text{j}}[t]$) feeding to both neurons. They are represented together as another activity vector, $\mathbf{ai^i_j[t]} = [fb^{\text{i}}_{\text{j0}}[t]\, ai^{\text{i}}_{\text{j}}[t]\, fb^{\text{i}}_{\text{j1}}[t]]^T$. Thus, each AU is modeled as:

$$\mathbf{au^i_j[t+1]} = f^{\text{i}}_{\text{j}}\left(\alpha^i_j \mathbf{au^i_j[t]} + \mathbf{W}^{\mathbf{au^i_j}}_{\mathbf{ai^i_j}}\mathbf{ai^i_j[t]} + \beta^{\text{i}}_{\text{j}}\right), \quad (1)$$

$$\mathbf{W}^{\mathbf{au^i_j}}_{\mathbf{ai^i_j}} = \begin{bmatrix} w_+^{fb_0} & w_+^{ai} & w^{fb} \\ w^{fb} & w_-^{ai} & w_-^{fb_1} \end{bmatrix}, \quad (2)$$

[2]All mapping weight connections and code can be accessed from https://github.com/wonner163/AVIS-II.

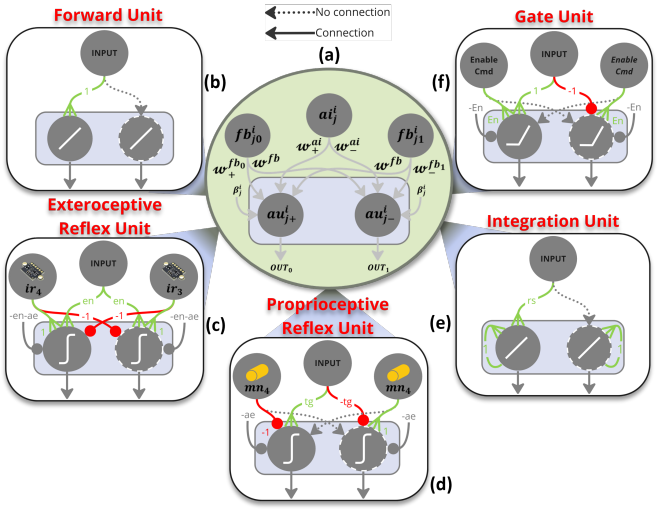

Fig. 6. (a) Neural circuit of an Adaptation Unit (AU), which can be configured to achieve five different functional units: (b) Forward Unit, (c) Exteroceptive Reflex Unit, (d) Proprioceptive Reflex Unit, (e) Integration Unit, and (f) Gate Unit.

where, $f^{i}_{j}()$ denotes the activation function used to control the output range (e.g., threshold function for discrete output or rectified linear function (ReLU) for continuous output); α^{i}_{j} denotes a scalar recurrent weight used to hold activities, functioning as memory; β^{i}_{j} denotes a scalar bias used for pre-inhibiting or pre-activating the unit; and w^{ai}_{+}, w^{ai}_{-}, $w^{fb_0}_{+}$, $w^{fb_1}_{-}$, and w^{fb} denote the parameters of the weight matrices $\mathbf{W}^{\mathbf{au}^{i}_{j}}_{\mathbf{ai}^{i}_{j}}$ mapping from $\mathbf{ai}^{i}_{j}\mathbf{[t]}$ to $\mathbf{au}^{i}_{j}\mathbf{[t]}$.

Thus, each AU is parameterized by the adaptation input vector $\mathbf{ai}^{i}_{j}\mathbf{[t]}$, activation function $f^{i}_{j}()$, and the seven parameters: $\{w^{ai}_{+}, w^{ai}_{-}, w^{fb_0}_{+}, w^{fb_1}_{-}, w^{fb}, \alpha^{i}_{j}, \beta^{i}_{j}\}$. With different sets of these parameters, the AU operates in different modes/functions, as shown in Fig. 6.

In this work, the AUN consists of five layers: AU^{0-4}. Each receives shared inputs from its previous layer ($\mathbf{ai^{i}[t]} = \mathbf{W ai^{i-1}[t]}$, where $\mathbf{ai^{i}[t]} = [ai^{i}_{0}[t]\, ai^{i}_{1}[t] ..., ai^{i}_{j}[t]]^{T}$ and $\mathbf{W}$ denotes the mapping matrix) and feeds its outputs to those of the next layer.

The first layer (AU^0) functions as a high-level action pattern module to subsequently activate multiple neurons in the next layer for multiple actions. AU^0 receives the outputs from the SISNs as its inputs to activate the desired action patterns in a sequential state: $\mathbf{ai^0[t]} = \mathbf{W_B[sisn^0[t]\, sisn^1[t]\, sisn^2[t]\, sisn^3[t]]}^{T}$, where $\mathbf{W_B}$ (Fig. 3(c)) is the matrix mapping from 57 state neurons ($\text{SISN}^{0-3}_{2n} = 15, 14, 16$, and 12) to 18 primitive action patterns encoded in 18 AU (AU^{0}_{0-17}, Fig. 6(b)). Finally, the outputs are then fed to the second layer.

The second layer (AU^1) generates a high-level adaptive signal for adaptive action based on IR sensor feedback. AU^1 receives outputs from AU^0 as its share inputs to activate the adaptive action units: $\mathbf{ai^1[t]} = \mathbf{W_C ai^0[t]}$, where $\mathbf{W_C}$ (Fig. 3(c)) is the matrix mapping from 18 preliminary action patterns to 18 action patterns encoded in the 18 AUs in this layer. These 18 AUs are divided into two types: 16 forward units (AU^{1}_{0-15}) for passing the signal to the following layer, and two exteroceptive reflex units (AU^{1}_{16-17}) for reflex action based on exteroceptive feedback, incorporating IR distance feedback (IR_3/IR_5 as positive feedback input and IR_4/IR_6 as negative feedback input, (Fig. 6(c))) to adapt the action. Thus, en is selected as 400 (the maximum feedback value) while ae is 5 (i.e., acceptance error of 5 mm). Finally, the outputs are then fed into the third layer.

The third layer (AU^2) generates an adaptive reflex signal (either extensor or flexor through positive AU ($au^{i}_{j+}[t]$) or negative AU ($au^{i}_{j-}[t]$)'s output). AU^2 receives the outputs from AU^1 as its inputs: $\mathbf{ai^2[t]} = \mathbf{W_D ai^1[t]}$, where $\mathbf{W_D}$ (Fig. 3(c)) is the matrix mapping from 18 action patterns to eight commands (six motor directional commands and two on/off electromagnet commands), encoded in the eight AUs in this layer. These eight AUs are divided into two types: six AUs (AU^{2}_{0-5}) for six motor commands to generate reflex action based on proprioceptive feedback and two forward AUs (AU^{2}_{6-7}) for two electromagnet commands. To compute the desired movement (e.g., extensor or flexor reflex signals), the proprioceptive reflex AUs (Fig. 6(d)) take the motor commands/MNs' activities at the previous timestep ($mn_k[t-1]$) as both feedback inputs and use the mapping matrix $\mathbf{W_D}$ to determine the target (tg) at each stage, while ae is empirically set as 0.1 (the acceptance error of 0.1 rad). Finally, the outputs are then fed into the fourth layer.

The fourth layer (AU^3) functions as a memory unit, storing the final motor position of the adaptive actions at the end of one state action and using it as the initial position in the next state of the sequential action. AU^3 receives the outputs from AU^2 as its shared inputs: $\mathbf{ai^3[t]} = \mathbf{W_E ai^2[t]}$, where $\mathbf{W_E}$ (Fig. 3(c)) is the matrix forwarding eight motor directional commands/electromagnet commands to the corresponding integrated signals, encoded in eight integration AUs (Fig. 6(e)) in this layer. To adjust the slope/integration gain independently of the SISN activities, the rs of all AUs are selected as $\mathbf{W^{rs}_E\ au^0[t]}$ where $\mathbf{W^{rs}_E}$ denotes the matrix mapping from 18 adaptable actions to the speed gains of six motor commands. This allows different trajectories of the motor command to be set independently for each action pattern. Finally, the outputs are then fed into the fifth layer.

The fifth layer (AU^4) functions as a gate unit, blocking or unblocking the signal from the previous layer to the MN based on the operation mode (OM). AU^4 receives the outputs from AU^3 as its share inputs: $\mathbf{ai^4[t]} = \mathbf{W_F ai^3[t]}$, where $\mathbf{W_F}$ is the matrix forwarding eight integrated signals to six motor target position commands and two electromagnet commands, encoded in eight gate AUs (Fig. 6(f)) in this layer. The OM is employed as input feedback to enable/disable the output signals from AU^3 (Fig. 3(c)). The parameter en is selected as π (the maximum input value). Finally, the outputs are combined with the LoCo module at the MNs.

An example of how AU^{0-3} works is provided in Fig. 7(a). AU^0 takes the SISNs signal to enable the units, as shown in

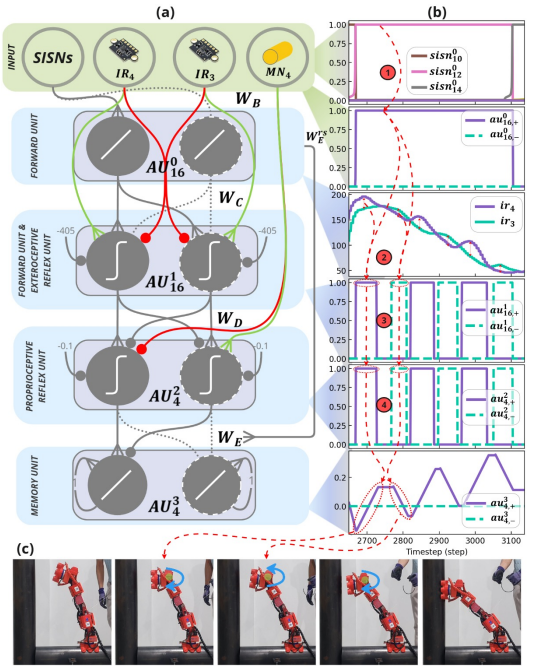

Fig. 7. (a) Example neural circuits of the first to the fourth layer of the Adaptation Unit Network (AUN), including AU^{0-1}_{16} and AU^{2-3}_{4}. (b) Neuron signals of four AUs (AU^{0-1}_{16} and AU^{2-3}_{4}), where the numbers in red indicate the signal sequence. (c) Snapshots of the real robot action mapped to the neural signals, with a blue arrow indicating the direction of M4 motor.

Fig. 7(b). When $sisn^0_{12}[t]$ is activated, the $au^0_{16,+}[t]$ is then activated.

As a result, the $au^1_{16}[t]$ takes two IR sensors (IR_{3-4}) as feedback and the previous layer ($au^0_{16,+}[t]$) as input to enable the unit. It then produces either a positive or a negative AU signal based on an error value from two IR sensors, as shown in Fig. 7(b). When the $ir_4[t]$ value exceeds $ir_3[t]$, the $au^1_{16,+}[t]$ is then activated. Conversely, if the $ir_3[t]$ value exceeds $ir_4[t]$, the $au^1_{16-}[t]$ is activated.

AU^2_4 takes MN_4 as feedback and $\text{AU}^1_{16+,-}$ as inputs. It then generates either a positive or a negative AU signal based on an error value from the target position (from the previous layer) and current motor feedback, as shown in Fig. 7(b). This results in a switching signal between positive $au^2_{4+}[t]$ and negative $au^2_{4-}[t]$ AU signals (extensor or flexor reflex signals).

AU^3_4 takes AU^2_4 as input, producing the corresponding extensor or flexor reflex signals, as shown in Fig. 7(b) and snapshots of the real robot actions in Fig. 7(c). These mechanisms enable inchworm-like contact searching (T2). Removing these AUs, therefore, disables contact alignment, resulting in unsuccessful transitions.

AU^4_4 takes AU^3_4 as input and the OM signal as an enable command, as illustrated in Fig. 8(a). $au^4_{4,+}[t]$ is activated with a positive signal input, while $au^4_{4,-}[t]$ is activated with a negative signal input, as depicted in Fig. 8(b). Moreover, when the OM is not equal to 1, AU^4 unblocks the signal (transition gait) and blocks it when the OM equals 1 (normal gait) using the LoCo module, as shown in the periodic crawling signal in Fig. 8(c). Without these AUs, the robot crawled before stable contact was formed, leading to falls.

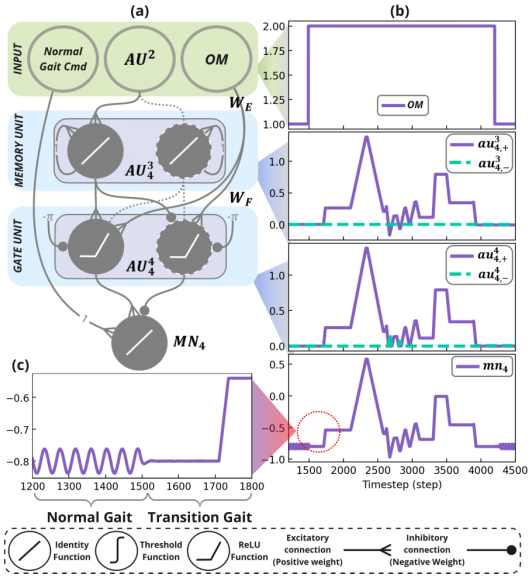

Fig. 8. (a) An example neural circuit of the fourth (AU^3) and fifth layer (AU^4) of the AUN for Motor Neuron 4 (MN_4). (b) The neuron signal between the fourth (AU^3_4), fifth layer (AU^4_4) for MN_4 correlates with the operation mode (OM) and normal gait command signal. (c) Transition of the $mn_4[t]$ signal between the normal gait and transition gait.

IV. Experiments and Results

Two main physical robot experiments were conducted to demonstrate the performance of IAMNC. The first experiment was designed to verify the success rates, transition time, cost of transport (COT), current consumption, and traveling distance of the robot under various transition conditions, namely, -45°, obstacle crossing (0°), 45°, 90°, and 135°. The second experiment demonstrates autonomous locomotion and transitions in a complex pipe environment, as shown in Fig. 1. Note that, throughout all experiments, industrial-standard pipes of six inches in diameter were used and the robot functioned fully autonomously.

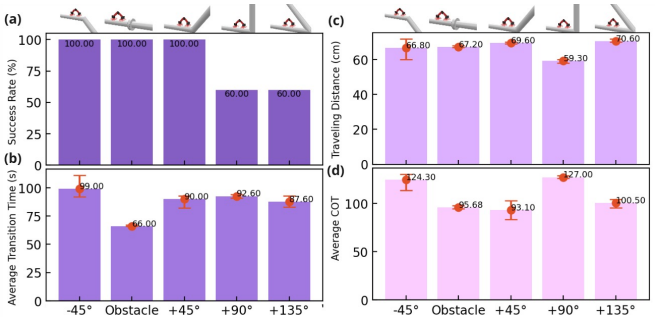

Fig. 9. Performance evaluation of the IAMNC in terms of (a) success rate (sr_a), (b) average transition time (Δt_a), (c) average traveling distance (d_a), and (d) average cost of transport (cot_a) obtained from five transition conditions, with distributions shown by red bars.

A. Experiment 1: Performance Evaluation

In the first experiment, five different transition conditions, namely -45°, obstacle crossing (0°), 45°, 90°, and 135°, were tested (Fig. 9). The robot was initially placed on the pipe; thus, it entered the crawling operation mode. Subsequently, the robot crawled forward toward the obstacle/connection before performing its transition. For each condition, this process was repeated five times (i = 1–5). During this process, the travel time (t_i), the travel distance measured along the pipes (d_i), the average current consumption (I_i), and the Boolean success status (s_i) were recorded. The recorded data were used to compute the success rate ($sr_a = 100\Sigma_i(s_i/5)\%$), transition speed ($v_i = d_i/t_i$), average transition time (t_a), average traveling distance (d_a), and average cost of transport ($cot_a = \frac{1}{5}\sum_{i=1}^{5} I_i V_i / W\, v_i$), where V_i denotes the fixed operating voltage of 24 V, and W denotes the total physical load which is 44.15 N.

Figure. 9(a) shows the robot successfully handling -45° transitions, obstacles, and 45–135° transitions. It achieved a 100% success rate for obstacle and ±45° transitions (5/5 trials) but only 60% for +90° and +135° transitions (3/5 trials), due to insufficient holding force during the dragging phase despite being within kinematics range. Figures. 9(b) and (c) show the average transition time and travel distance. The transition time varied from 66 s (obstacle crossing, 10 actions) to 87–99 s (pipe transitions, 12–14 actions). The travel distance differed according to pipe conditions, with +135° transitions covering 70 cm versus 59 cm for +90°, despite the same mode being used. This difference resulted from sensory-based adaptive sequencing and adaptive actions. Figure. 9(d) presents the average cost of transport (cot_a), influenced by transition time, distance, and pipe configuration. COT ranged from 93 to 127. The +90° transition had the highest energy demand (127) due to its shorter travel distance (59 cm). In contrast, the lowest COT (93) was observed with the +45° transition due to minimal energy consumption.

B. Experiment 2: Autonomous Locomotion and Transitions on a Complex Pipeline

The robot was tested on a complex pipeline under five conditions with a total length of 7.5 m.: an obstacle (flange), +45° bevel, -45° bevel, elbow, and +135° bevel connections as shown in Fig. 10(a). It started from the leftmost end and autonomously crawled to the rightmost end, balancing on pipes, stepping across obstacles, transitioning between segments, and climbing a vertical pipe. The operation mode (OM), SISN activities, and motor/electromagnet commands were recorded, as shown in Fig. 10(b). Initially, detecting no obstacles, the robot performed normal crawling (OM = 1), using the LoCo module (periodic patterns in $m_{1,2,4}$) while adapting to roll feedback (m_3). Around timestep 700, detecting a flange via IR_2, it switched to obstacle-crossing mode (OM = 5), using $SISN^3$ to step across (the purple dashed line in Fig. 10(b)). After completing the sequence at timestep 2,500, it resumed crawling. At timestep 3,200,

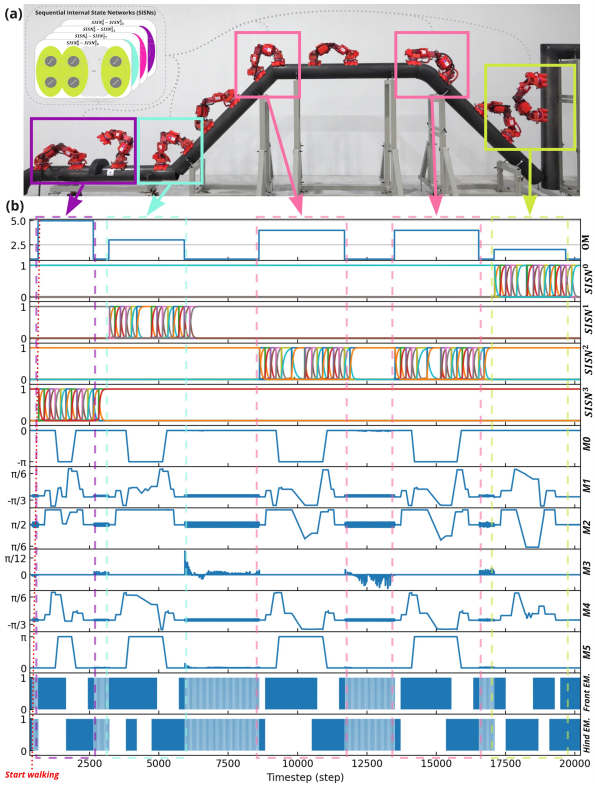

Fig. 10. (a) Snapshots of the robot performing adaptive multimodal locomotion including crawling on horizontal and vertical pipes, crossing an obstacle (flange), and performing gait transitions between different pipe segments. (b) Neural signals, including MoSe outputs (OM), SISN activities ($SISN^{0-3}$), motor commands ($M0 - M5$), and gait diagrams of the front and hind legs (EM). Note that the signals correspond to the diagram in Fig. 3, while the gait diagram represents electromagnet activation (attachment, blue) and deactivation (detachment, white). A video is available at https://youtu.be/TiLU2V0leHs.

detecting a +45° connection via IR_0 and IR_2, it switched to upward transition mode (OM = 3), using $SISN^1$ (the green dashed line in Fig. 10(b)) to move onto the +45° pipe. After completing the sequence at timestep 5,900, it resumed crawling. At timesteps 8,600 and 13,500, detecting a -45° connection (IR_2), it switched to steep upward transition operation mode (OM = 4), using $SISN^2$ (pink highlight in Fig. 10(b)) to step onto horizontal and -45° pipes. After completing transitions at 11,700 and 16,500 timesteps, it resumed crawling. Finally, at timestep 17,100, detecting a +135° connection (IR_1, IR_2), it switched to downward transition mode (OM = 2), using $SISN^0$ (yellow green highlight) to step onto the +135° pipe. After completing the transition at 19,660 timesteps, it resumed crawling upward on the vertical pipe. In this experiment, the robot successfully navigated the entire pipeline within 11 minutes, with an average walking speed of 1.14 cm/s.

V. Conclusion and Discussion

Inspired by inchworms, which use minimal exteroceptive sensors [10, 11], an IAMNC with nine exteroceptive signals (seven IR and two inductive sensors) is developed in

TABLE I
COMPARISON OF FIVE STATE-OF-THE-ART ROBOTS IN TERMS OF TRANSITIONS AND OBSTACLE CROSSING

Robot	Control	IK	FI		Transitions‡	
			Type	Dims	Type	COT
[3]	AUTO	✓	DCAM DIS IMU	240p 4 3	OBS +90°	n/a
[4]	AUTO	✓	Map	$>100/m^2$	+90°	n/a
[5]	AUTO	✓	DCAM	240p	+90°	n/a
[2]	SEMI	✓	OC	5	OBS +90°	n/a n/a
AVIS-II	AUTO	✗	DIS IND	7 2	-45° OBS +45° +90° +135°	124.20 95.68 93.10 127 100.5

IK: inverse kinematics required; **FI**: feedback information; **AUTO**: autonomous control; **SEMI**: semi-autonomous control; **OC**: operator command; **DCAM**: depth camera; **DIS**: distance sensor; **IND**: inductive sensor; **OBS**: obstacle; ‡ obtained/estimated from real-world demonstration as reported/presented in the original paper/video;

this work. It integrates three modules: MoSe for rule-based environment classification, LoCo for adaptive crawling, and the newly proposed TraCo for inchworm-inspired transitions. Each module is designed neuron-by-neuron for transparency and interpretability [12]. The TraCo module enables adaptive transitions using SISNs for sequential actions and AUs for adaptability. Additionally, this work introduces AU as an interpretable module performing five key functions: forward, exteroceptive reflex, proprioceptive reflex, integration, and gating. When performing as an AUN network, the AU can execute complex adaptive sequences, such as those required for autonomous gait transitions. Using only nine exteroceptive signals, the AVIS-II robot successfully navigated -45° to +135° pipe transitions, including obstacle crossing, outperforming previous works that achieved at most +90° transitions [2–7] or required over 76,800-dimensional depth images [3–5] (see Table I). Extending from current 2D planar pipeline, future work will be validated with complex geometries (e.g., 3D pipelines) by adding IR sensors to detect the extra dimension, enabling MoSe to classify more pipe configurations. In the current setup, a 180° rotation in OM 3–5 is enforced to ensure rear-foot placement on clear pipe surfaces rather than connectors. To remove this constraint, we will implement inchworm-like soft gripping feet [13] to improve adhesion, success rates, and adaptability to various pipe diameters and materials.

ACKNOWLEDGMENTS

This research was supported by Fundamental Fund, Thailand Science Research and Innovation (TSRI, Grant No. FRB690039/0457) and by AI and Robotics Ventures (ARV) under the AdVanced Pipe Inspection Robot System (AVIS) project.

REFERENCES

[1] C. Viegas and M. Tavakoli, "A single DOF arm for transition of climbing robots between perpendicular planes," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 2867–2872.

[2] A. Srisuchinnawong, K. Phongaksorn, W. Ausrivong, and P. Manoonpong, "Adaptive bipedal robot walking on industrial pipes under neural multimodal locomotion control: Toward robotic out-pipe inspection," *IEEE/ASME Transactions on Mechatronics*, 2023.

[3] H.-D. Bui, S. Nguyen, U. H. Billah, C. Le, A. Tavakkoli, and H. M. La, "Control framework for a hybrid-steel bridge inspection robot," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.

[4] W. Chen, S. Gu, L. Zhu, H. Zhang, H. Zhu, and Y. Guan, "Representation of truss-style structures for autonomous climbing of biped pole-climbing robots," *Robotics and Autonomous Systems*, vol. 101, pp. 126–137, 2018.

[5] P. Ward, P. Manamperi, P. Brooks, P. Mann, W. Kaluarachchi, L. Matkovic, G. Paul, C. Yang, P. Quin, D. Pagano *et al.*, "Climbing robot for steel bridge inspection: Design challenges," in *Austroads Bridge Conference*. ARRB Group, 2014.

[6] J.-H. Kim, J.-C. Lee, and Y.-R. Choi, "PiROB: Vision-based pipe-climbing robot for spray-pipe inspection in nuclear plants," *International Journal of Advanced Robotic Systems*, vol. 15, no. 6, 2018.

[7] M. Tavakoli, L. Marques, and A. de Almeida, "3DCLIMBER: Climbing and manipulation over 3D structures," *Mechatronics*, vol. 21, pp. 48–62, 02 2011.

[8] M. Brandao, G. Canal, S. Krivić, and D. Magazzeni, "Towards providing explanations for robot motion planning," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021.

[9] R. H. Plaut, "Mathematical model of inchworm locomotion," *International Journal of Non-Linear Mechanics*, vol. 76, pp. 56–63, 2015.

[10] L. I. N. N. van Griethuijsen, "Behavioral responses to mechano-sensory information in a soft-bodied terrestrial animal," Ph.D. dissertation, Tufts University, 2012.

[11] P. K. Ward, "Design of a biologically inspired climbing robot and an adhesion mechanism for reliable and versatile climbing in complex steel structures," Ph.D. dissertation, 2016.

[12] C. Glanois, P. Weng, M. Zimmer, D. Li, T. Yang, J. Hao, and W. Liu, "A survey on interpretable reinforcement learning," *Machine Learning*, 2024.

[13] Q. Zhao, Z. Jiang, and H. K. Chu, "A soft-rigid air-propelled pipe-climbing robot," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, May 2021, pp. 11 850–11 855.