

Self Modeling and Gait Control of Quadruped Robot Using Q-Learning Based Particle Swarm Optimization

Syed Irfan Ali Meerza and Md. Mohiuddin Uzzal

Abstract—In the realm of the living creature human and animals create their body schema by the learning they gather while they interact with the real world. They can also remodel the schema if they have any uncertain changes in their body. This kind of robustness is still not achieved by any machine or artificial system. Researchers are trying to build the machines resilient so that machines can explore the unknown space. In this paper, we used Particle Swarm Optimization (PSO) which a population based algorithm to allow a quadruped robot to learn its body schema using a gyroscopic sensor and real world interaction. We added Q- Value based learning (Q-Learning),s an actor-critic scheme to aid PSO to learn faster and avoid being trap in local optima. Robot creates an imaginary model of its own body which include imaginary gaits using a very little prior knowledge. The robot aims to use the gaits to achieve stability and predictive movements. I can also detect changes in its body and adopt the changes, which leads to a damage diagnosis system. We tested the algorithm using graphics simulator and verified using a 3D printed quadruped robot with 12 actuators.

Keywords—particle swarm optimization, q-learning, evolutionary computation, quadruped, gait control, self-modeling

I. INTRODUCTION

Robots are one of the most common and essential parts of modern technology. Robots are assisting human to accomplish the works which are considered more sophisticated or dangerous for the human to get involved. Multi-legged robots like the quadruped which has four mechanical legs for walking provides human an alternative to exploring the areas where direct human involvement is difficult or considered hazardous. Multiple legged robots provide advantages over the wheeled robots like the greater mobility they provide in uneven and disturbed terrain [1] and stability they provide while walking [2]. Quadruped robots are a rather simpler form of the legged robot which provides both the stability and easier model to work with.

Syed Irfan Ali Meerza

Electrical and Electronics Engineering Department
American International University Bangladesh
Dhaka, Bangladesh
irfanmeerza00185@gmail.com

Md. Mohiuddin Uzzal

Professor
Electrical and Electronics Engineering Department
American International University Bangladesh
Dhaka, Bangladesh
drmohiuddin@aiub.edu

Advantage of the quadruped is that; they can move in two degrees of maneuverability on any terrain. The challenge with the legged robots is the control of the legs. Each leg must give a movement in such a manner that they provide the required thrust and support. All the legs must move in coordination so that they can provide constant torso stability and also move in the required direction.

The main drawback of the robots is, they do not provide robustness when it comes to their physical condition adaptation like the insects and other animals shows [3]. But researchers are trying to solve this problem so that robot can learn their physical models as the other living beings do. Artificial Intelligence (AI) is one of the solutions as the majority of the AI algorithms are inspired by nature. Evolutionary algorithms are best suited for this kind of problem. Particle Swarm Optimization algorithm is a stochastic method inspired by the biological swarms present in nature [4]. It is a population-based algorithm where the system reaches to the optima by the sharing their experience among the swarm of particle and evolution of generations.

PSO alone is a sufficient algorithm, but sometimes it stuck at local optima rather than reaching the global optima which lead to an increase in possible actions hence increases complexity and computation. Q-Learning is an action selection and reward based algorithm which selects the required actions and discards unnecessary actions based on the reward it gets for that specific action [5]. It maintains a Q-Table which contains the reward of a specific action. Q-Learning combined with PSO leads to a limited action policy where actions are done based on swarm learning and selected by the reward it gets.

This paper proposes to use the Q-learning based PSO for the quadruped to learn the gait configuration of its own. Our method is continuous, and it models the gait configuration whenever it faces difficulty on stability or the movement. It forms a loop of learning, validation, and testing for continuous operation.

Rest of the paper is organized as follows, in section II other related works on autonomous gait configuration or the self-modeling of a robot is discussed. In section III the proposed algorithm with the description of the models which are used to verify the algorithm is discussed. While section IV reports the different findings of the algorithm and the model generation in the physical robot. Finally, in section V we concluded the paper and also discussed some factors related to the results we achieved.

II. RELATED WORKS

In neuroscience, there is an established theory that, a higher being must develop a predictive model of their own body as the biological sensors are too slow to provide feedback in case of fast movement [6]. Body model must predict the movement related to specific muscle action without any feedback from the sensors [7]. But achieving this robustness in machines or robots is a difficult task and researchers tried different ways to allow robots to create a model of their own body. In [8] they used a genetic algorithm to develop a controller for hexapod robot, they simulated it and later [9] they applied it into an insect type hexapod robot and successfully evolved full locomotion controllers both with sensors and without and generated a pattern of leg movements known as the tripod gait. In [10]

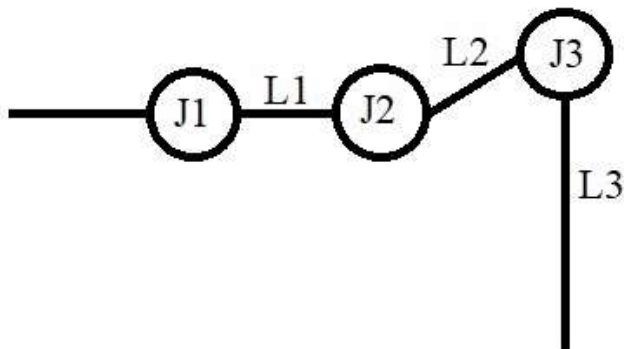


Fig. 1. Joint configuration of a typical quadruped leg. Here J1 is coxal joint, J2 is tibia joint and J3 is the femur joint.

Spencer used genetic programming to generate a program for hexapod gaits using less prior knowledge and simulated it in a robot simulator. He was successful in generating required activation and stable forward movement. In [11] they proposed a multi-sensor based four-legged robot with decision making capability. The robot had the learning capabilities and monitoring of the environment and their own parameters. Mahdavi and Bentley used a genetic algorithm to detect damages in a snake robot which uses shape memory alloy as muscle [12].

III. PROPOSED SYSTEM

Self-modeling of a multiple lagged robot stands for the robot to discover its own gait configuration. A combination of reinforcement learning and population-based search method is used here. To test the proposed method, a target model is used in this work. The system contains two parts: firstly, characterization of the target to be identified and the algorithm to be fitted for the model.

A. Characterization of the proposed model

The target system is a quadrupedal robot with twelve degrees of freedom. We chose the quadruped robot as it shows a simple design with stability. Legged robots are well fitted for this particular work as the wheel robots have a fewer parameter to be controlled and the self-modeling is not an easier task for a wheeled robot. Our robot has a rectangular body with four legs connected with the body using servo motors. Each leg contains three servo motors to representing three joints; Coxal, Tibia, and Femur joints. Fig.

1 shows the joint configuration of the leg. Table I gives the overall dimension of the robot. All the servo motors are controlled by a microcontroller. These servo drives are capable of producing 1.8 kg-meter of torque and 60 degrees per second of speed. The servos are actuated within a range to prevent unrecoverable movement. Table II summarizes the range of all joints of the robot. The robot is equipped with different types of sensors. A MEMS-based accelerometer, gyroscope, magnetometer combo sensor is used at the center of the robot for detecting its left/right and forward/back tilts. Four infrared sensors are attached to four legs which gives four binary value whether the legs are touching the ground or not. Another infrared sensor is attached at the belly of the robot. It gets the information of belly touching the ground. Fig. 2 shows the physical quadruped robot that is used as the target model for this piece of work.

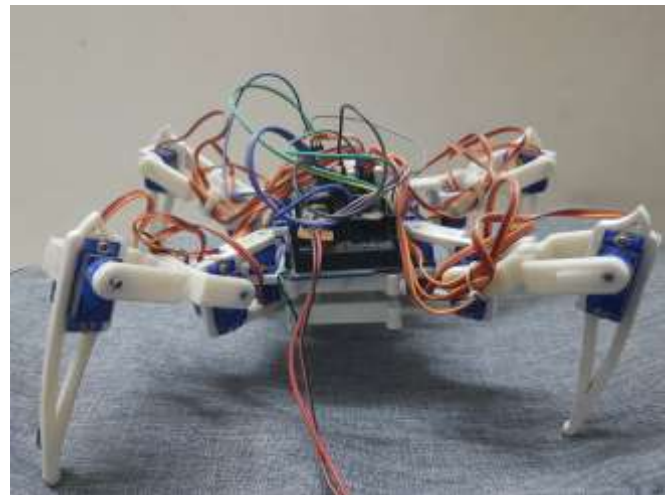


Fig. 2. Quadruped robot physical model.

B. Characterization of 3D model

Models are the three-dimensional representation of the robot using a graphics library and IDE as a simulation. The simulation only considers the probable model of the physical robot, not its movement. Only known parameters here are the size of each part of the legs. The simulation starts with the planner configuration, and as the robot moves, each part start to connect and create a model based on the data collected and algorithmic estimation. Models are encoded as vectors, and the data collected from different sensors are used to simulate the model vector and create a possible articulated robot. Each leg part contains two positions of joint. The Q table contains the probability of each part of the leg for a specific position of the torso and also contains the probability of one leg part connection with other. These probabilities construct a matrix, which represents the joints of the robot.

C. Proposed Algorithm

The proposed algorithm is a combination of Q-Learning and Particle Swarm Optimization (PSO). Here the actions are selected using the Q-learning algorithm. Q-Learning assigns reward point against an action based on the outcomes of the

TABLE I. PHYSICAL DIMENSION OF THE ROBOT

Parameter	Dimension (mm)
Length of the body	150

Parameter	Dimension (mm)
Width of the body	80
Height of the body	55
Coax length	35
Femur length	60
Tibia length	90

TABLE II. JOINT PROPERTIES OF THE ROBOT

Joints	Lower Range (Degree)	Upper Range (Degree)
Coxal joint	-30	30
Femur joint	-30	30
Tibia joint	0	90

action. Here actions are the angular movement of servo motors in the leg. Servos are allowed to move only 10 degrees a time. One action contains the movement of three servos in a group. Rewards are given on the basis of the level of stability achieved while having all the legs on the ground and belly clear of the ground. In our algorithm, each leg works as a particle and PSO is used to optimize the leaning. Here learning from each leg's action and reward table after one iteration information is shared between the legs and on the basis of the best action, all other actions are optimized. This leads to faster learning. This process runs continuously and until the robot has flat torso with a minimum distance from the ground and all the legs are touching the ground. At the time of movement, the robot finds the best way to move in a certain direction while having the torso flat and clear off the ground. If the robot reaches the stability goal the learning algorithm stops until it detects any changes in the sensor data. Fig. 3 shows the algorithm flow chart.

D. Characterization of the controller

The main aim of the controller to achieve the angle combination of the servo motors that makes the robot to stability goal. The robots start at a planner position with all the angles at zero degrees. The angle of the joints is allowed to choose within a range (-30, 30) with five-degree increments. These constraints are limiting the range of motions,

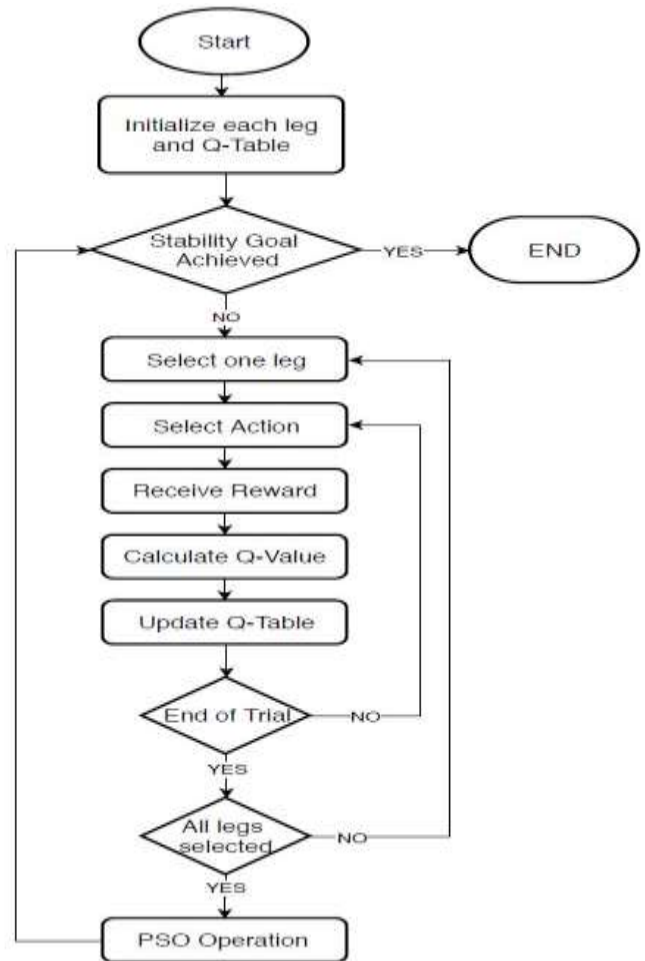


Fig. 3. Flowchart of the algorithm. Main focus of this algorithm is to achieve the stability. When the stability achieved the algorithm reaches the end until then it tries to iterate through the Q-learning and PSO operation.

but they are used to prevent any unexpected movement that can damage the robot and also eliminate any complex movement that may mislead the algorithm. This limits came from the limit in the range of motion of a spider. After initialization, a random motor program is generated and initialized the controller. Then each legs q-value is calculated for a trial of 5. After all trials of all legs the PSO is performed to found the global best position of the motors and on this basis PSO the best angle for the joints of the robot is determined, and all the legs joints are converted to that angles. In case of movements like walking forward, backward, left or right the algorithms run to make the robot move forward by a series of joint movement. Rewards are given on the degree of movement achieved while keeping the stability conditions intact.

IV. RESULT ANALYSIS

Our test results are divided into two sections one is parametrical results which consist of different parameters in the robot modeling and movement; another one is topological results which is the outcome of the algorithm.

A. Parametrical Results

In this section only the parametric data are discussed, where robots learning through the self-modeling is focused. In first set experiments, the robot has only one task to

perform which is to stand on its feet. Evaluation is divided into four generations where different random motor models are used in each generation. Each generation has 10 trials which means the robot has a total of 40 trials. In each trial, the subjective error is calculated where subjective error stands for the difference in the present sensor data and estimated sensor data. Estimated sensor data are predetermined and manually entered to the robot's program. Another error is calculated which is the model accuracy error. Here the model generated from the sensor data in graphics IDE is compared with the robot's actual model and the error is calculated from the accuracy of modeling. Fig. 4, Fig. 5 and Fig. 6 reports the result from a typical run of the robot.

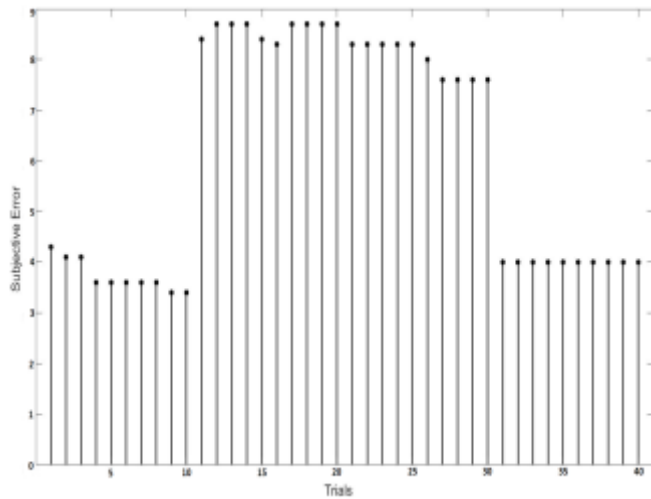


Fig. 4. Parametric identification for subjective error. We can see that the first run has a lower error as the robot started from the planer configuration. But as it start to move the error rises and it again falls down at the fourth generation after 30 trails. And in the fourth generation, the error is stable as the robot reaches the stable position. This error generated due to the difference in manually estimated stable position and the calculated stable position by the algorithm.

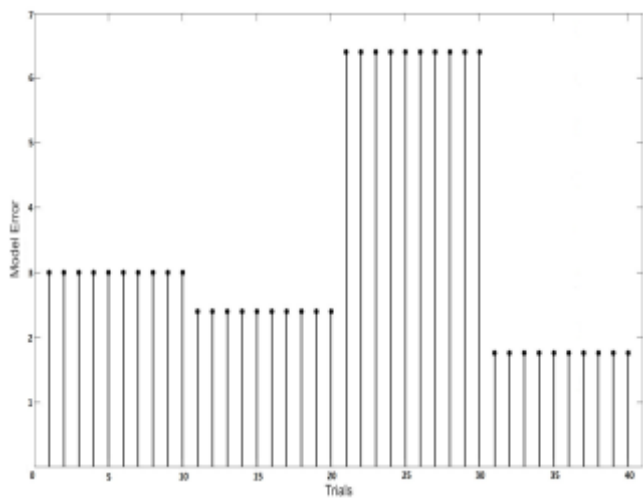


Fig. 5. Parametric identification for model error. This error stands for the accuracy of the prediction. Model error is the same for all the runs as the error is calculated after each generation. At the end of the fourth generation, the error found is 1.76 which means accuracy found is 98.24%.

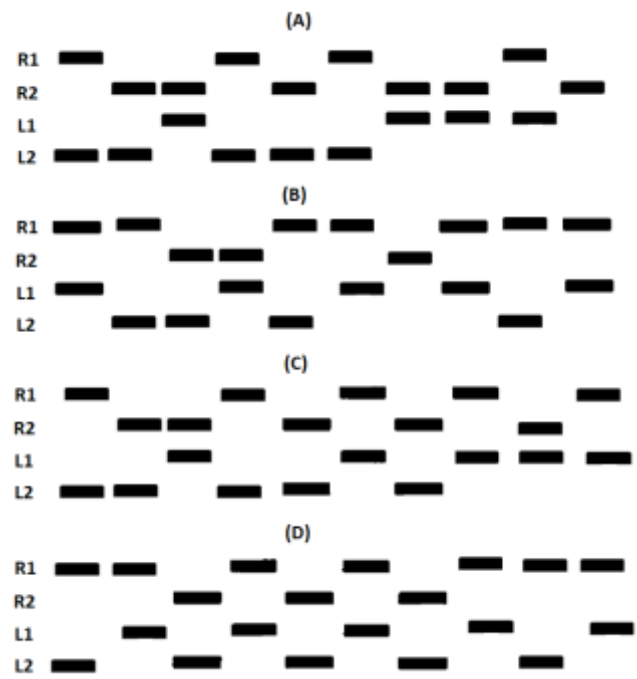


Fig. 6. Behavior of a locomotion controller for four generations from A to D where A is the first generation and D is the fourth generation. The legs are labeled L for left and R for right and numbered from 1 to 2 starting from the front of the insect. Black bars denote the swing phase of a leg, and the space between bars represents a stance phase.

B. Topological Results

In topological results, only the robot's self-modeling capability is discussed. As the modeling requires more information, so we increased the number of trails to 100 in each generation. We removed one tibia from one leg and tested if the robot can remodel itself. We also checked the standing configuration of the robot and checked the subjective error and model error. Fig 7 shows the subjective error after removing the tibia. And Fig. 8 shows the model obtained at the end of the second, third and fourth generation, respectively.

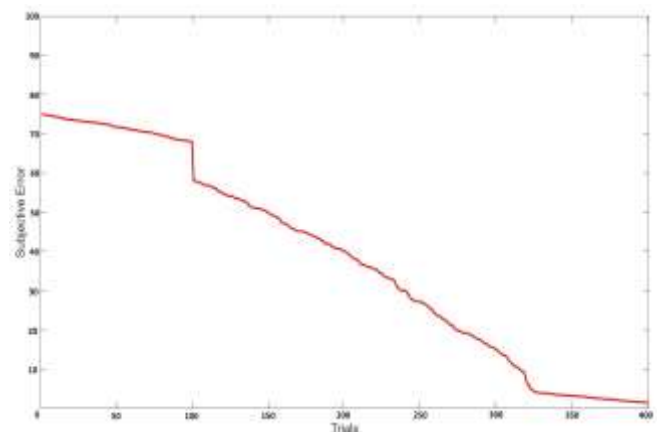


Fig. 7. Subjective error for 100 trails when the tibia is removed. It shows that the error is too high when the first generation is started because the robot is not able to stabilize itself and keeps falling in each trial. After first generation, robot has quite a good amount of data and start to learn the gait configuration which leads to a lower error. The error reaches 1.53 at the end of the fourth generation.

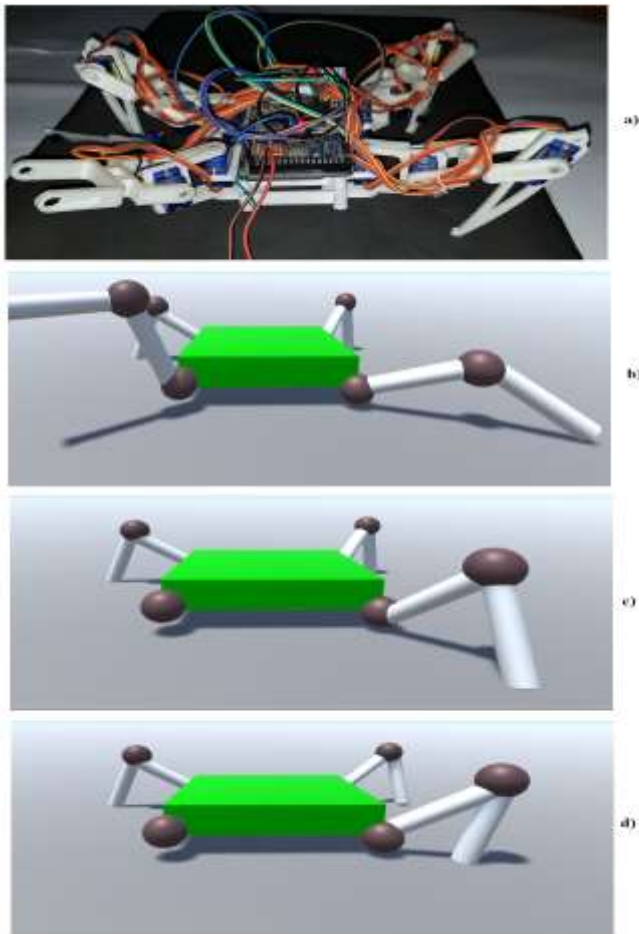


Fig. 8. a) The standing configuration after fourth generation. b) predicted gait configuration after second generation in the 3D model. c) predicted gait configuration after third generation. d) final gait configuration after all the trials. The accuracy of the predicted model is 98.47%.

V. CONCLUSION

Here we have reported the successful implementation of proposed Q-learning based PSO algorithm for automated synthesis of robot models based on a physical robot's embodied and situated interactions with its environment. Specifically, we have demonstrated successful parametric identification, in which robot tried to achieve standing pose with no knowledge of its legs and gaits and calculated the deviation from the best possible pose. We also demonstrated topological identification, in which robot tries to find out its physical model built up by combining disparate model building blocks (in this work, leg parts) together in the right way. We tested the modeling capability by modifying the robot's physical appearance and let it find the new model. This ability enables the self-diagnosis in the robotic system.

REFERENCES

- [1] F. Tedeschi and G. Carbone, "Design Issues for Hexapod Walking Robots," *robotics*, Jun. 2014.
- [2] "Robot Platform | Knowledge | Wheeled Robots." [Online]. Available: http://www.robotplatform.com/knowledge/Classification_of_Robots/legged_robots.html. [Accessed: 24-Jun-2016].

- [3] A. Cully, J. Clune, D. Tarapore and J. Mouret, "Robots that can adapt like animals", Macmillan Publishers Limited, Vol. 521, pp 503-515, May 2015.
- [4] S. I. A. Meerza, M. Islam, M. M. Uzzal, "Optimal Path Planning Algorithm for Swarm of Robots Using Particle Swarm Optimization Technique", in *Proceedings 3rd International Conference on Information Technology, Information System and Electrical Engineering*, 2018, pp. 331-335.
- [5] S. I. A. Meerza, M. Islam, M. M. Uzzal, "Q-Learning Based Particle Swarm Optimization Algorithm for Optimal Path Planning of Swarm of Mobile Robots", in *Proceedings 1st International Conference on Advances in Science, Engineering and Robotics Technology*, 2019, pp. 858-862.
- [6] D. Wolpert, R. C. Miall, M. K. (1998). Internal models of the cerebellum. *Trends in Cognitive Sciences*, 2:2381–2395.
- [7] Llinas, R. R. (2001). *The i of the Vortex*. Cambridge, MA: MIT Press.
- [8] R.D. Beer and J.C. Gallagher, "Evolving dynamical neural networks for adaptive behavior", *Adapt. Behav.*, 1, pp. 91–122, 1992.
- [9] J.C. Gallagher and R.D. Beer, "Application of evolved locomotion controllers to a hexapod robot", Technical Report CES-94-7, Department of Computer Engineering and Science, Case Western Reserve University, 1994.
- [10] G. Spencer, "Automatic generation of programs for crawling and walking", in *Advances in Genetic Programming*, K. Kinneer Jr., Ed., Cambridge, MA: MIT Press, 1994, pp. 335–353.
- [11] A. Mahajan and F. Figueroa, "Four-legged intelligent mobile autonomous robot," *Robotics and Computer-Integrated Manufacturing* 13, 51-61 (1997).
- [12] S.H. Mahdavi and P.J. Bentley. "An evolutionary approach to damage recovery of robot motion with muscles". In *17th European Conference on Artificial Life (ECAL03)*, pp. 248-255, Spinger, Berlin, 2003.
- [13] P. A. Kumar and Y. S. Narayan, "Design of a quadruped robot and its inverse kinematics", in *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)*, Vol. 7, Issue 4, pp. 241-252, Aug 2017.



Syed Irfan Ali Meerza received B.Sc. degree in Electronics and Communication Engineering from KUET, Bangladesh in 2015 and M.Sc. degree in Electrical and Electronic Engineering from American International University Bangladesh (AIUB) in 2019. His research interest is Machine Learning, Control System and Robotics. He is currently working as a Graduate Teaching Assistant (GTA) at, University of Louisville (UofL), USA.

Mohammad Mohiuddin Uzzal received B.Sc. degree in Electrical and Electronic Engineering from BUET, Dhaka, Bangladesh in 2002 and M.S.E. degree in Electrical Engineering from Arizona State University (ASU), USA in 2006. He had earned his Ph.D. from Electrical and Computer Engineering Department of University of New Mexico, USA. While pursuing, Ph.D. degree his research focus was mixed signal CMOS circuits for Bio-sensing application specially A-ISFET based DNA Sequencing Chip. From 2004 to 2012, he was with American International University - Bangladesh as a Full-time Faculty in Electrical and Electronic Engineering Department. From May to Dec. 2014, he was with Intel Corporation, Santa Clara, CA



and worked on next generation high speed (10-16Gbps) I/O design and simulation. His research interest includes analog/RF and mixed signal circuit design, low cost bio-sensors and Nano-electronics. He

holds a professional MBA degree from Institute of Business Administration, University of Dhaka, in 2008. He is currently ranked as a Professor.