

Autonomous Vision-based Control of Aerial Manipulator for Horizontal Pipe Structure Tracking with Continuous Contact

Naisarg Pandya

School of Mechanical and
Materials Engineering

Indian Institute of Technology Mandi
Email: npandya831@gmail.com

Amit shukla

Centre for Artificial Intelligence
and Robotics

Email:amitshukla@iitmandi.ac.in

Pushker Kumar

School of Mechanical and
Materials Engineering

Indian Institute of Technology Mandi
s22038@students.iitmandi.ac.in

Ankit Mehra

School of Mechanical and
Materials Engineering

Indian Institute of Technology Mandi
s22020@students.iitmandi.ac.in

Abstract—Continuous health monitoring of industrial pipelines, railway tracks, and powerline-like horizontal structures is crucial for ensuring human safety. Recently, the aerial manipulator concept is getting more attention for the contact inspection of industrial structures. This paper proposed the autonomous vision-based control of an aerial manipulator for tracking of horizontal pipe structure while maintaining continuous contact by a manipulator. The perception module of the proposed approach includes a deep learning technique for pipe identification in the image, a classical vision for feature extraction, and a Kalman filter to resolve the data latency problem in real time. The integration of LIDAR and the camera sensor has been used to extract the cartesian coordinate of the approximate contact point. The manipulator is designed and developed along with inverse kinematics to maintain continuous contact with the pipe. Image-based sliding mode controller is used for lateral and yaw orientation control of the aerial platform. The altitude control is done using a LIDAR sensor. The novel forward velocity function is introduced which serves the purpose of smooth tracking along with maintaining a pipe in the reachable space of the manipulator. The fully autonomous operation strategy has been designed to organize subtasks sequentially with feedback. During the tracking of a pipeline, the lateral position and altitude position with respect to manipulator's base did not deviate beyond $\pm 0.2\text{m}$. Fully autonomous vision-based control of aerial manipulator has been validated experimentally on a 10-meter long pipeine.

I. INTRODUCTION

The health monitoring of the industrial and commercial infrastructure is paramount to preventing hazardous conditions. Pipelines are one of the most commonly used infrastructures in industries for fluid transportation. Pipelines are usually subjected to fluid friction, high internal pressure, and extreme environmental conditions which damage the pipeline's interior as well as exterior layers. The damaged pipeline surface layers lead to explosion and leakage of the pipelines. The continuous health monitoring of the pipelines can prevent

leakage and emission as well as help to predict the requirement for maintenance and repair. The health monitoring of pipelines includes non-contact and contact inspection. Usually, non-contact inspection methods like visual inspection include leakage monitoring and exterior pipe layer defect detection. Contact inspection such as Ultrasonic Testing (UT) and electromagnetic Acoustic Technology (EMAT) are used for leakage detection, micro-crack detection on an exterior layer, and internal structure defect detection. Manual contact inspection of long-range pipelines is not feasible and costs substantial resources like time, cost, and manpower. In this scenario, the field of robotics can provide a solution that will be accurate, time-efficient, and cost-effective with no human error.

To perform a contact inspection, a robot must have the capability to interact with the environment and maneuverability over the pipelines. The wheeled or crawling robot can have maneuverability and ability to interact with the environment but the maneuverability is limited to the material of the pipeline's surface and these robots have limited speed [1, 2]. The aerial manipulator encapsulates both the capability to interact with the environment irrespective of the material of the pipeline as well as omnidirectional maneuverability with high speed, so the aerial manipulator can perform the contact inspection in a very effective manner.

This paper proposed the methodology of fully autonomous vision-based control of aerial manipulator for tracking the pipe structure along with the end effector will be in continuous contact with the pipe. The manipulator with 3 controllable joints and 2 compliant joints has been used for uninterrupted contact. The end effector of the aerial manipulator consists of a servo-operated gripper with a compliant mechanism to hold the pipe for uninterrupted contact.

II. RELATED WORK

Continuous contact inspection by aerial manipulators involves the accurate identification, localization of objects of interest in real-time, precise motion control of the aerial platform, and continuous manipulator position control. Nowadays, The Object identification task is done using deep learning techniques. Image segmentation is one of the image-processing techniques that precisely separates an object from its background and can be used for object identification. There are some standard image segmentation models like U-Net, Seg-Net, SqueezeNet, DeepLabv3, and DeepLabv3+ are available [3–8]. These models have good segmentation performance but generally, the model size is large so, these models need to be modified and reduced a size in such a way that they can be deployed on edge devices for robotics applications. In the case of a fast-moving object in the image frame, the real-time detection and continuous tracking of the object are difficult using image processing techniques. So, in such cases, a Kalman filter can be used for object localization and tracking [9], [10].

Precise tracking of horizontal structures by aerial platform is one of the most important subtasks in continuous contact inspection by aerial manipulators. Low-altitude tracking is a challenging task because of the small field of view of the vision sensor and the high chance of losing objects from the visual range. In [11, 12] horizontal structure tracking by aerial platform has been described but the interaction with the environment was not involved with the tracking of the structure.

For contact inspection, the novel design of multicopter was proposed in [13], which includes a study on the ceiling effect on the stability of multicopter during an interaction. However, during contact interaction, inspection over a long surface cannot be easily achieved by this methodology. In the beginning, the aerial manipulator idea was proposed for the pick and place, valve turning, sensor installation, and other manipulation tasks in [14–17]. Recently, the aerial manipulator came into the application of contact inspection due to its ability to interact with the environment and omnidirectional mobility. In [18], the under-bridge inspection using an aerial manipulator has been described with force control and a novel manipulator design with a compliant mechanism. In [19], the hammering test of the bridge has been performed by an aerial manipulator. In [20], the force control has been done using image-based visual servoing with impedance control. The force control was performed in both static and sliding conditions but the trajectory of sliding of end-effector was not controlled. The predefined sliding trajectory of end-effector has been controlled in [21]. Using a predefined way-point and motion planner, the indoor experiment has been performed to track the end-effector desired trajectory. However, the predefined trajectory is not feasible for long outdoor infrastructures. In [22], the novel aerial manipulator design for NDT contact inspection has been proposed with eight tilted rotor configurations. The contact inspection and sliding of the end

effector has been validated in the outdoor conditions but it's a semi-autonomous operation. First the contact of the manipulator's end effector must be established manually and maintaining the contact is autonomous in [22]. The continuous contact interaction along with the pipe structure tracking is still an unexplored concept which is very useful for performing continuous contact inspection along the length of the pipe. The combination of the autonomous tracking of structure by aerial platform and continuous manipulator position control can lead us to perform contact inspection on the desired trajectory in real-time effectively.

In this paper, we have proposed autonomous vision-based control of aerial manipulator for tracking pipe structure along with continuous uninterrupted contact by a manipulator. The major contribution of this work is as follows:

- The design and develop a perception module that contains a custom-made CNN for semantic segmentation, feature extraction by classical vision technique, and Kalman filter for data latency problem.
- The autonomous vision-based control of all DOF of multicopter has been designed with a sliding mode controller and validated experimentally.
- The development of a manipulator and its pose control to maintain uninterrupted contact with pipe using cartesian coordinates of approximate contact points which has been derived from the image feature and LIDAR map.
- A fully autonomous operation strategy has been developed and experimentally validated for organizing the sequence of sub-tasks to perform autonomous tracking of the pipe along with maintaining continuous contact of an end effector with the pipe.

III. OVERARCHING APPROACH

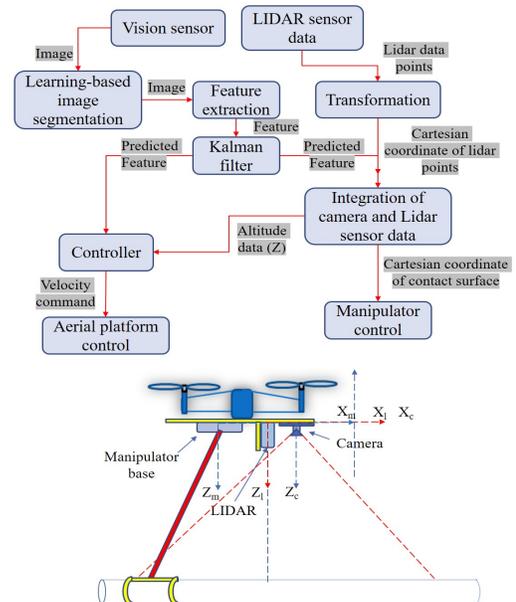


Fig. 1: Overarching methodology.

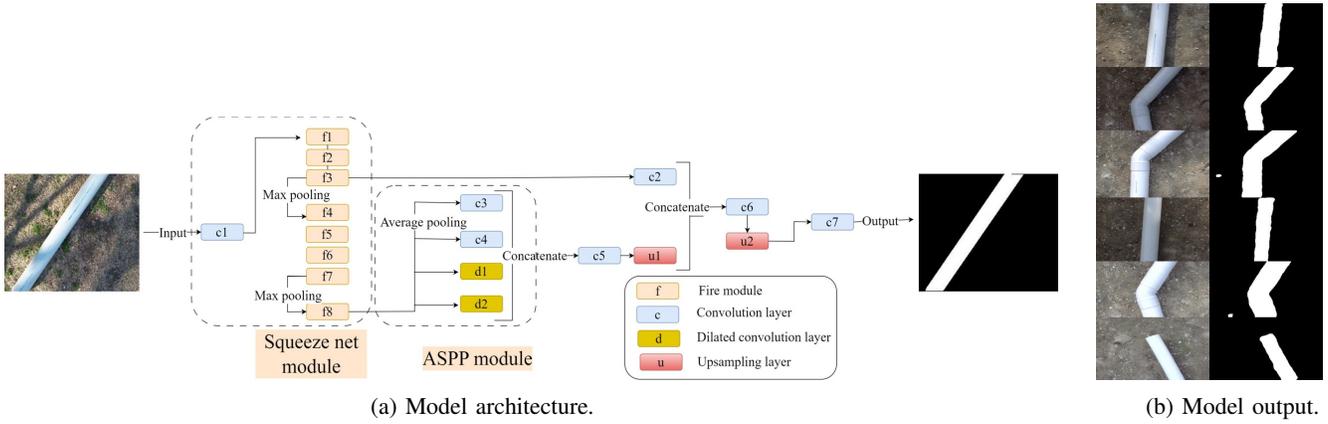


Fig. 2: Convolution neural network architecture.

The overarching approach and data flow in the system for autonomous control of the aerial manipulator is shown in Fig.1. The aerial manipulator control has been divided into aerial platform control and manipulator control. The vision sensor (camera) will feed the real-time images into semantic segmentation which gives pipe segmented binary image. The binary image is used for feature extraction. The kalman filter will use the extracted feature and estimate the feature to increase the data frequency for vision based control. The LIDAR will provide LIDAR map that will be superimposed onto binary image. The camera and LIDAR sensor data has been fused to extract the cartesian coordinate of approximate contact point.

The aerial platform's lateral and yaw motion has been controlled based on the kalman filter data and altitude control uses the LIDAR data. The manipulator joint parameters controlled based on the cartesian coordinate of approximate contact point.

IV. PERCEPTION MODULE

A. Deep-learning based image segmentation

The identification of the pipe in the image frame is performed using a Convolution Neural Network (CNN). The SqueezeNet and DeepLabv3+ are popular CNN architectures for image segmentation. The DeepLabv3+ included Astrous Special Pyramid Pooling (ASPP) in the encoder to capture the multiscale features of the object and special decoder structure [7],[8]. SqueezeNet contains the fire module which helps to reduce the number of hyperparameters in the model in an effective manner [6]. We combined both architectures to have a small model size with reasonable performance of pipe segmentation. The model architecture is shown in the Fig. 2a.

The CNN model output will be the binary image in which the pipe will be segmented as 1 and the background as 0. The model output is shown in Fig. 2b. The CNN model is followed by the classical vision technique for feature extraction.

The Pipe localization in the image frame is performed by locating the lateral point and yaw point as shown in Fig. 3. To locate these points on the segmented image, the image has been sliced into sub-images. On each sub-image, contour

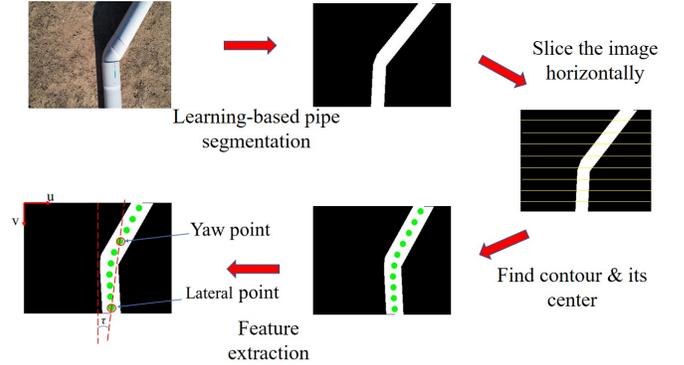


Fig. 3: Feature extraction.

detection has been implemented to locate the center of each segment of the pipe.

B. Latency problem in real-time data

The latency problem has inherently come when we have deployed the image processing with CNN and all the ROS networks on the edge device (NVIDIA Jetson Nano). To resolve this issue, we have used semantic segmentation with 3hz frequency and to increase the data frequency for the controller, the Kalman filter has been used. The Kalman filter will estimate the image feature when the image feature data has not been received from the perception module (CNN and classical vision technique). The Kalman filter algorithm is implemented with modifications in estimation terms as stated in the equ 4. We have set the 15hz frequency for the Kalman filter output that will be used by the vision-based controller.

$$X_{kp} = AX_{(k-1)} \quad (1)$$

$$P_{kp} = AP_{(k-1)}A^T + Q \quad (2)$$

$$K = \frac{P_{kp} H^T}{(HP_{kp} H^T + R)} \quad (3)$$

$$X_k = X_{kp} + K(Z_k - HX_{kp}) + g(X_{(k-1)} - Z_k) \quad (4)$$

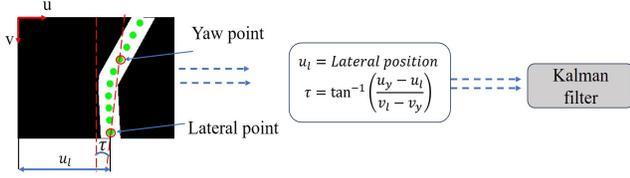


Fig. 4: Implementation of Kalman filter for feature estimation.

$$P_k = (I - KH)P_{k,p} \quad (5)$$

We have modified the estimation term because of latency in measurement terms. The Kalman filter will help to reduce the noise in data and it will also avoid the sudden changes in the image feature.

V. SENSOR INTEGRATION AND TRACKING CONTROL

A. Sensor integration

The integration of the camera and LIDAR sensor is used to get the cartesian coordinate of the approximate contact point of the pipe. The position of the camera, LIDAR, and manipulator on the aerial platform is shown in Fig.5. The

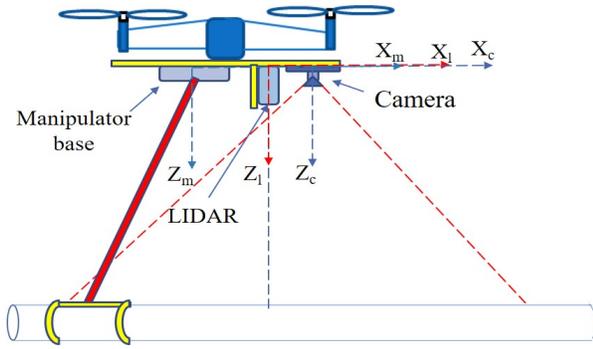


Fig. 5: Autonomous operation task distribution.

$\sum[X_c, Y_c, Z_c], \sum[X_l, Y_l, Z_l], \sum[X_m, Y_m, Z_m]$ represent the coordinate frame of the camera, LIDAR, and manipulator respectively. The coordinate frame of the camera, LIDAR, and manipulator is defined in such a way that the X-axis of each frame will be coincident and the Z-axis will be colinear. The monocular camera can provide 2-D information of the object in the image frame and LIDAR can generate a depth map in one plane. The polar coordinate of an object of interest can be found using a camera and LIDAR as shown in Fig. 6.

$$\alpha = \tan^{-1}\left(\frac{u * \tan(HFOV * 0.5)}{320}\right) \quad (6)$$

The angular position α of the object (pipe) in the camera frame can be found using equ 6. Where u and $HFOV$ represent the lateral position of the object in pixel coordinate and horizontal field of view respectively. Where u is the lateral position of an object (pipe) where the LIDAR map intersects on the image frame and $HFOV$ is horizontal field of view. The LIDAR can provide depth information r at a particular angular position α . We assumed that the pipe distance from the aerial platform

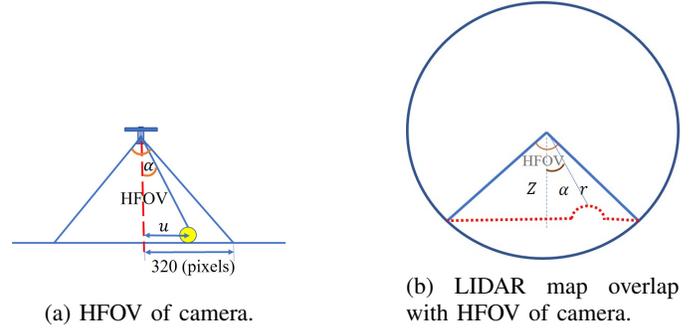


Fig. 6: Polar coordinate using camera & LIDAR sensor.

along the Z_m axis would be the same for all points on the pipe in the image.

We have considered a lateral point as an approximate contact point because it is the closest point where the manipulator will be in contact with the pipe. The cartesian coordinate (z_p, y_p) of the approximate contact point in the manipulator frame can be calculated using equ. 7 and 8 respectively.

$$z_p = r * \cos(\alpha) \quad (7)$$

$$y_p = \frac{u_l * z_p}{f * \rho} \quad (8)$$

Where the f and ρ represent the focal length and pixel density respectively. We have used the camera intrinsic parameter to calculate the y_p because the lateral point will not coincide with the LIDAR plan. The cartesian coordinate (y_p, z_p) will be used for calculating the joint parameter of the manipulator.

B. Tracking control

The vision-based tracking control of the aerial platform (multicopter) has been employed to have smooth tracking of pipe. The tracking needs to be performed in such a way that the pipe will always stay in the reachable space of the manipulator as well as in the field of view of the camera. The

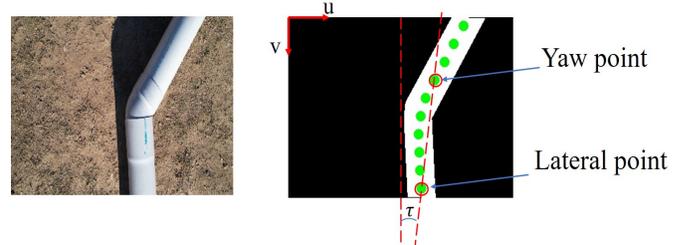


Fig. 7: Lateral and yaw error in the image frame.

lateral error and yaw error can be found from the extracted visual feature as shown in Fig. 7. To get the controlled motion of a multicopter, the sliding mode controller with PI sliding surface and exponential reaching law have been implemented. The error of the lateral correction, yaw correction, and depth correction can be defined as per equ.9, 10, 11.

$$e_l = u_l - u_i \quad (9)$$

$$\tau = \tan^{-1}\left(\frac{u_y - v_l}{v_l - v_y}\right) \quad (10)$$

$$e_z = z_p - z_d \quad (11)$$

In equ.9, 10, the $(u_l, v_l), (u_y, v_y)$ are the lateral point coordinate and yaw point coordinate. u_i is the image center point coordinate. In equ. 11, the z_p and z_d are the current pipe depth and desired pipe depth with respect to the LIDAR frame. The sliding surface and reaching law can be defined as per equ. 12, and 14 respectively.

$$S = ce + k \int edt \quad (12)$$

$$\dot{S} = c\dot{e} + ke \quad (13)$$

$$\dot{S} = -\epsilon \tanh(S) - k(S) \quad (14)$$

By equating equ. 13 and 14, the sliding mode controller for velocity in lateral, yaw, and altitude can be derived as equ. 15 by incorporating the respective error in the sliding surface.

$$V = -\frac{1}{c}\epsilon \tanh(S) - \frac{1}{c}k(S) - \frac{1}{c}e \quad (15)$$

The novel continuous forward motion approach is introduced in this work to have smooth tracking while considering the lateral, altitude, and yaw position of the multicopter with respect to pipe. The hyperbolic function is used for forward motion control which is dependent on the lateral correction, yaw correction, and altitude correction. The forward velocity function is designed as per equ. 16.

$$V_x = \frac{A}{f(e_l, \tau, e_d)} \quad (16)$$

$$f(e_l, \tau, e_d) = w_1 e_l + w_2 \tau + w_3 e_d + B \quad (17)$$

In equ. 16, and 17, the A, B, w_1 , and w_2 are the heuristically tuned parameter. The max and minimum forward velocity can be set using A and B. The w_1, w_2 , and w_3 are the weights for the lateral error, yaw error, and depth error.

VI. MANIPULATOR'S CONFIGURATION AND INVERSE KINEMATICS

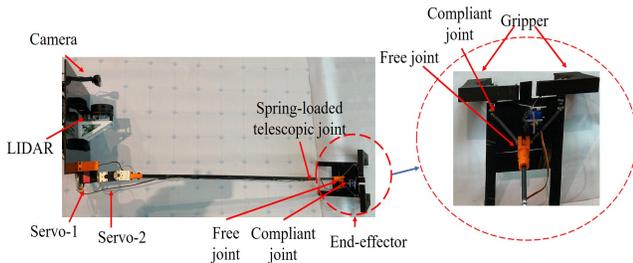


Fig. 8: Autonomous operation task distribution.

We have designed and developed a special-purpose manipulator to maintain continuous contact during pipe tracking as

shown in Fig. 8. The manipulator has to maintain the end-effector's position on the pipe while pipe tracking. There is a possibility that the aerial platform may deviate laterally or in altitude direction so, the manipulator has to compensate this movement for uninterrupted contact. The gripper orientation must be down-facing and aligned with the pipe while holding it.

To maintain the end-effector position on the pipe, we have introduced 2 controllable joints at the base and it will be actuated using the cartesian coordinate feedback from the perception module. The joint parameters can be calculated from the cartesian coordinate according to equ. 18 and 19.

$$\theta_1 = \cos^{-1}\left(\frac{z_p}{l_1}\right) \quad (18)$$

$$\theta_2 = \sin^{-1}\left(\frac{y_p}{l_1 * \sin(\theta_1)}\right) \quad (19)$$

For the down-facing of the gripper, the free joint has been incorporated at the end-effector and one compliant joint has been used for the flexible orientation of the gripper during the movement of the manipulator base along Y-axis and Z-axis. One more spring-loaded telescopic joint has been introduced in the link of a manipulator to absorb shock during pipe tracking. The feed-forward control has been used to control the 3-servo angular position.

VII. AUTONOMOUS OPERATION STRATEGY

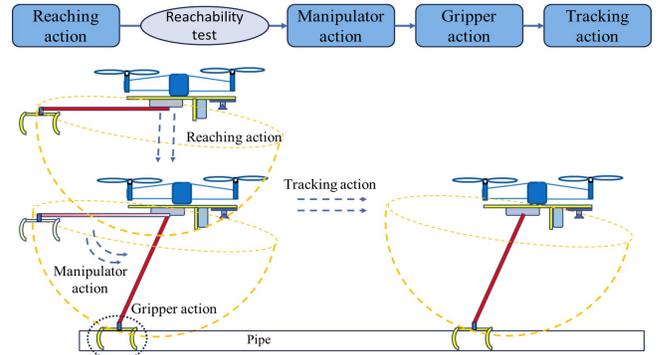


Fig. 9: Autonomous operation task distribution.

The fully autonomous operation strategy is designed to organize the sub-tasks as shown in Fig. 9 to perform a fully autonomous operation of pipe tracking along with maintaining continuous contact on the pipe. The fully autonomous operation has been divided into several actions which have been described as follows.

- Reaching action: The aerial manipulator will perform position control action to reach toward the pipe and it will maintain an appropriate distance in such a way that the pipe will come in the reachable space of the manipulator.
- Reachability test: The reachability test provides a close loop feedback to execute manipulator action. To hold the pipe, the pipe should be in the reachable space of a manipulator and in the field of view camera. Fig. 10

shows the intersection of the manipulator's work space and the field of view of the camera. For the safety

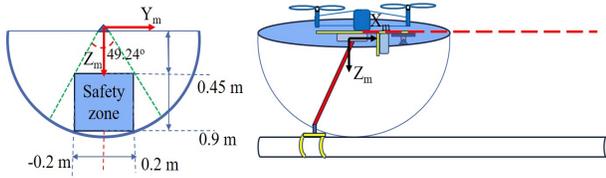


Fig. 10: Reachability test.

of the manipulator as well as aerial platform stability, the safety zone has been defined by heuristically tuned thresholds in y_p and z_p coordinates which are reflected in the experimental results also.

- Manipulator action: The manipulator will execute controlled joint parameters if the pipe is in the reachable space. The manipulator action will maintain the gripper position in such a way that the gripper can hold the pipe.
- Gripper action: The manipulator action will be followed by the gripper action. The gripper will be opened during the reaching action and manipulator action. After the manipulator action, the gripper will hold the pipe.
- Tracking action: To have continuous contact along the pipe, the aerial manipulator will start the tracking action. The aerial manipulator will perform the tracking action along with the continuous manipulator action.

VIII. EXPERIMENTAL RESULTS

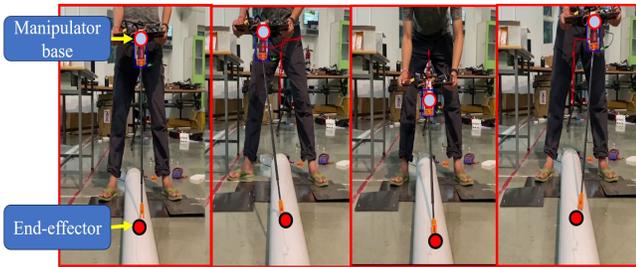


Fig. 11: Validation of vision-based control of manipulator based on sensor integration.

The vision-based control of the manipulator along with the perception module is validated before it deploys on the aerial platform. The manipulator platform was equipped with a Logitech c-270 camera, RPLIDAR-A1 sensor and for onboard computation Nvidia Jetson Nano. As shown in Fig. 11 while the manipulator base was moving along the Y-axis and Z-axis, the manipulator's end effector was maintaining its position on the using feedback from the perception module.

The experimental validation of autonomous vision-based control of aerial manipulator has been performed on a 10-meter long pipe in outdoor conditions. The experiment has executed an autonomous operation strategy as we have described in section . The aerial manipulator is made of a high-payload capacity multicopter (EFT E616P 16L Agricultural



Fig. 12: Fully autonomous operation of aerial manipulator and end-effector trajectory.

Drone Hexacopter) and a custom-made manipulator. Fig. 12 shows the trajectory of the multicopter and end effector. At the beginning of the operation, the multicopter is manually set in such a way that the pipe will come in the field of view of the camera and switch to the autonomous mode. In autonomous mode, according to the autonomous operation strategy first of all the reaching action will take place. In reaching action forward velocity will be zero. When the pipe is in the reachable space and safe zone the manipulator action will be executed and it will start maintaining the end effector position on the pipe. The gripper action will start after the manipulator action with a small time delay. The tracking action will start when the gripper grasps the pipe. During tracking action, the multicopter will start moving forward along with maintaining a lateral and altitude position. Fig. 13 and 14

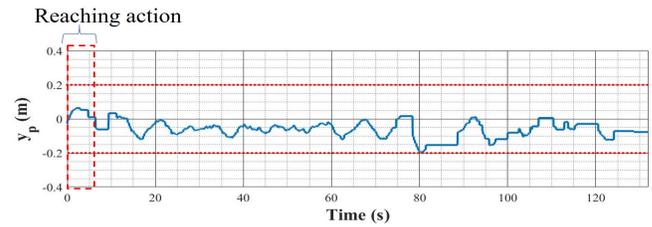


Fig. 13: y_p coordinate of the approximate contact point.

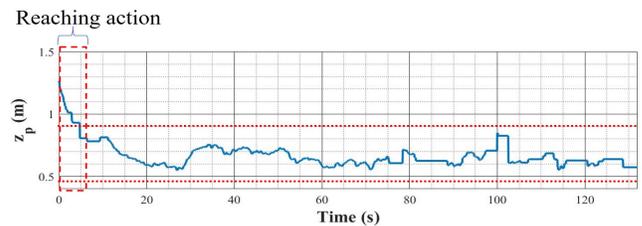


Fig. 14: z_p coordinate of the approximate contact point.

show the y_p and z_p coordinates of the approximate contact point. During the operation lateral position was always in the safe zone. In the beginning, the z_p coordinate was 1.25 m so, the reaching action has been activated. The reaching action was active in the time duration of 0 to 6.88s. In reaching

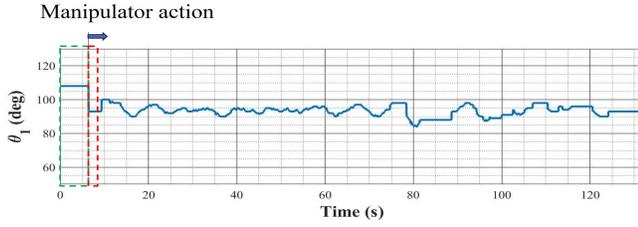


Fig. 15: θ_1 variation for manipulator action.

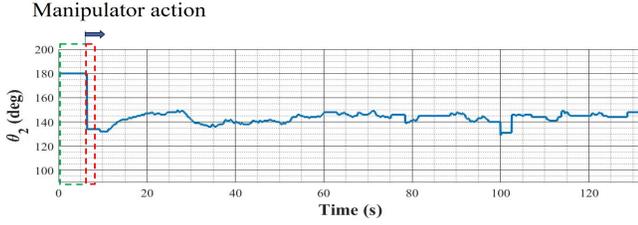


Fig. 16: θ_2 variation for the manipulator action.

action, the θ_1 , θ_2 , θ_3 were at their ideal positions and V_x was zero. The reaching action is followed by the manipulator action when the pipe comes into the safe zone. The continuous manipulator action started from 6.88s onwards as shown in Fig 15 and 16. During the reaching action and manipulator action

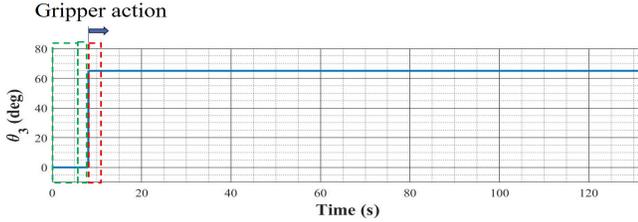


Fig. 17: θ_3 variation indicating the gripper action.

the gripper was open from 0 to 8.11s as shown in Fig. 17. The manipulator action is followed by the gripper action. The gripper was closed during the tracking action as shown in the Fig. 17.

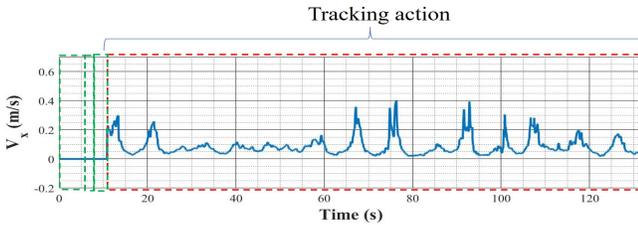


Fig. 18: Forward velocity of aerial platform.

The aerial platform's forward velocity and lateral velocity are shown in Fig. 18, and 19. During the reaching action, at the beginning of the manipulator action and the gripper action the forward velocity was zero from 0s to 10.8s. After the gripper action, the aerial platform started moving forward according to equ. 16. The lateral velocity varies according to the pipe

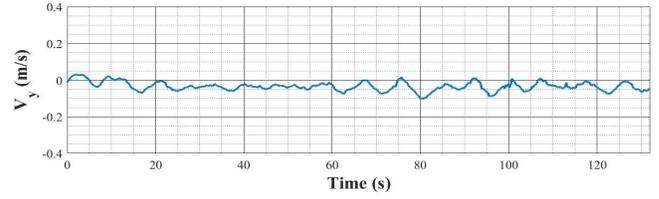


Fig. 19: Lateral velocity of aerial platform.

position in the image frame due to external disturbance as shown in Fig. 19. The depth velocity varies according to

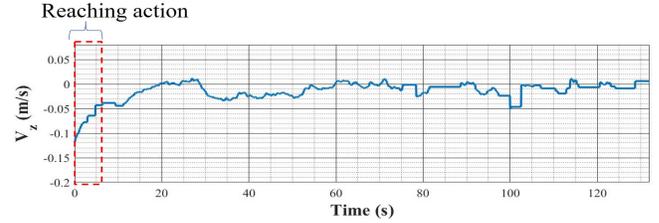


Fig. 20: Depth velocity variation.

the z_p coordinate of the approximate contact point. In the beginning, the Aerial platform height was 1.25 meters so, the depth velocity was high in the negative direction during the reaching action. The experiment shows that the reaching action, manipulator action, gripper action, and tracking action have been executed sequentially according to the autonomous operation strategy.

IX. CONCLUSION

This paper presents the autonomous vision-based control of an aerial manipulator for tracking horizontal pipe structures along with maintaining continuous contact with the pipe. We have designed a perception module that consists of custom-made CNN with classical vision technique for pipe identification along with feature and Kalman filter for increasing data frequency. The sensor fusion of the camera and LIDAR is used to extract the cartesian coordinate of an approximate contact point. The new manipulator has been fabricated along with the development of its feed-forward position control. The image feature-based tracking control of the aerial platform is developed with a sliding mode controller. The autonomous operation strategy to organize the sub-tasks is developed and experimentally validated.

The fully autonomous operation has been experimentally validated in outdoor conditions on the 10-meter-long pipe. The tracking of the pipe by the aerial platform has been achieved in a way that the y_p -coordinate and z_p -coordinate did not deviate over $\pm 0.2m$. During tracking action continuous contact has been achieved over the pipe which has been validated through the experimental data. In future work, the force feedback to ensure the contact and force control to perform contact-based inspection will be carried out to make the system more useful and reliable in industrial applications.

REFERENCES

- [1] Y. Mano, R. Ishikawa, Y. Yamada, and T. Nakamura, "Development of high-speed type peristaltic crawling robot for long-distance and complex-line sewer pipe inspection," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 8177–8183, 2018.
- [2] I. N. Ismail, A. Anuar, K. S. M. Sahari, M. Z. Baharudin, M. Fairuz, A. Jalal, and J. M. Saad, "Development of in-pipe inspection robot: A review," in *2012 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (STUDENT)*, pp. 310–315, 2012.
- [3] N. O'Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan, and J. Walsh, "Deep learning vs. traditional computer vision," in *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 I*, pp. 128–144, Springer, 2020.
- [4] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-net and its variants for medical image segmentation: A review of theory and applications," *Ieee Access*, vol. 9, pp. 82031–82057, 2021.
- [5] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [6] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.
- [7] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.
- [8] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.
- [9] S. Chen, "Kalman filter for robot vision: a survey," *IEEE Transactions on industrial electronics*, vol. 59, no. 11, pp. 4409–4420, 2011.
- [10] P. R. Gunjal, B. R. Gunjal, H. A. Shinde, S. M. Vanam, and S. S. Aher, "Moving object tracking using kalman filter," in *2018 International Conference On Advances in Communication and Computing Technology (ICACCT)*, pp. 544–547, IEEE, 2018.
- [11] S. Rathinam, Z. Kim, A. Soghikian, and R. Sengupta, "Vision based following of locally linear structures using an unmanned aerial vehicle," in *Proceedings of the 44th IEEE Conference on Decision and Control*, pp. 6085–6090, 2005.
- [12] A. Kumar, A. Shukla, A. Gupta, and A. K. Shivratri, "Vision-based horizontal structure tracking and inspection via unmanned aerial vehicle," in *ASME International Mechanical Engineering Congress and Exposition*, vol. 86656, p. V003T04A031, American Society of Mechanical Engineers, 2022.
- [13] A. E. Jimenez-Cano, P. J. Sanchez-Cuevas, P. Grau, A. Ollero, and G. Heredia, "Contact-based bridge inspection multirotors: Design, modeling, and control considering the ceiling effect," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3561–3568, 2019.
- [14] S. Kim, S. Choi, and H. J. Kim, "Aerial manipulation using a quadrotor with a two dof robotic arm," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4990–4995, 2013.
- [15] G. Garimella and M. Kobilarov, "Towards model-predictive control for aerial pick-and-place," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4692–4697, 2015.
- [16] A. Suarez, H. Romero, R. Salmoral, J. A. Acosta, J. Zambrano, and A. Ollero, "Experimental evaluation of aerial manipulation robot for the installation of clip type bird diverters: Outdoor flight tests," in *2021 Aerial Robotic Systems Physically Interacting with the Environment (AIRPHARO)*, pp. 1–7, IEEE, 2021.
- [17] M. Orsag, C. Korpela, S. Bogdan, and P. Oh, "Valve turning using a dual-arm aerial manipulator," in *2014 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 836–841, 2014.
- [18] A. Jimenez-Cano, G. Heredia, and A. Ollero, "Aerial manipulator with a compliant arm for bridge inspection," in *2017 international conference on unmanned aircraft systems (ICUAS)*, pp. 1217–1222, IEEE, 2017.
- [19] T. Ikeda, S. Yasui, M. Fujihara, K. Ohara, S. Ashizawa, A. Ichikawa, A. Okino, T. Oomichi, and T. Fukuda, "Wall contact by octo-rotor uav with one dof manipulator for bridge inspection," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5122–5127, IEEE, 2017.
- [20] M. Xu, A. Hu, and H. Wang, "Image-based visual impedance force control for contact aerial manipulation," *IEEE Transactions on Automation Science and Engineering*, vol. 20, no. 1, pp. 518–527, 2022.
- [21] M. Tognon, H. A. T. Chávez, E. Gasparin, Q. Sablé, D. Bicego, A. Mallet, M. Lany, G. Santi, B. Revaz, J. Cortés, and A. Franchi, "A truly-redundant aerial manipulator system with application to push-and-slide inspection in industrial plants," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1846–1851, 2019.
- [22] M. Á. Trujillo, J. R. Martínez-de Dios, C. Martín, A. Viguria, and A. Ollero, "Novel aerial manipulator for accurate and robust industrial ndt contact inspection: A new tool for the oil and gas inspection industry," *Sensors*, vol. 19, no. 6, p. 1305, 2019.