IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

Event-triggered image moments predictive control for tracking evolving features using UAVs

Sotirios N. Aspragkathos[1], George C. Karras[2] and Kostas J. Kyriakopoulos[3]

***Abstract*—This paper presents a novel approach for tracking deformable contour targets using Unmanned Aerial Vehicles (UAVs). The proposed scheme combines image moments descriptor and Event-Triggered (ET) Nonlinear Model Predictive Control (NMPC) for efficient and accurate tracking. The deformable contour model allows adaptation to the evolving target's shape, while the proposed event-triggered scheme achieves improved computational efficiency and extended flight duration while generating new control sequences for the UAV. Real-world experimental validation as well as a comparative simulation performance analysis validate the scheme, showcasing its robustness in handling complex scenarios. This approach holds promise for various applications, such as surveillance and autonomous navigation.**

***Index Terms*—Visual Servoing, Aerial Systems: Perception and Autonomy, Autonomous Agents**

I. Introduction

IN RECENT years, Unmanned Aerial Vehicles (UAVs), have gained widespread use across various industries, such as agriculture, construction, film production, and environmental monitoring, due to their ability to access remote or challenging terrains. They play a crucial role in inspecting critical infrastructure, enhancing maintenance and surveillance processes. In the field of robotics, particularly with camera-equipped UAVs, challenges arise in tracking dynamically changing targets. Ongoing research aims to develop innovative techniques for real-world surveillance scenarios, improving UAV performance in tracking and stabilizing evolving subjects.

A. Related Work

Visual servoing, combining robotics and computer vision, facilitates accurate control and manipulation tasks based on visual feedback. Image-based Visual Servoing (IBVS), Position-based Visual Servoing (PBVS), 2-1/2-D, and Direct visual servoing are fundamental approaches [1], [2], [3], [4]. IBVS is preferred for its robustness in target tracking, particularly in UAV stabilization and tracking control applications [5]. Image moments provide a stable alternative to point feature-based control laws, offering stability and invariance to camera frame rotation [6], [7], [8], [9], [10].

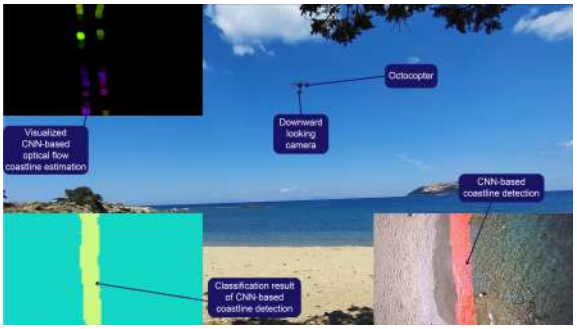


Fig. 1: A multirotor aerial vehicle executing autonomous coastline (deformable target in the nature) surveillance. A CNN classifies the contour in real-time based on the camera images while the prediction is fed on another CNN outputting the optical flow estimation of the detected coastline. The UAV uses an Event-triggered Image moments-based NMPC scheme to navigate along the coastline, detect the target, estimate its motion, and follow it.

NMPC is a powerful method for visual servoing, used in UAV target tracking in combination with various control schemes such as adaptive sliding mode control, prescribed performance control, and deep neural networks [11], [12], [13], [14], [15], [16], [17], [18], [19]. Combining image moments with NMPC addresses singularity cases and incorporates visibility and mechanical constraints [20], [21].

The emerging concept of event-triggered NMPC recommends optimization triggering and less frequent sampling, enhancing system efficiency [22], [23]. This approach has been applied to Position and Image-based Visual Servoing MPC, enabling asynchronous control inputs for autonomous robots. [24] introduces an event-triggering observer-based control scheme for UAVs, maximizing network efficiency. Other studies propose event-triggering controllers for trajectory tracking, swarm formation control, and UAV network control [25], [26], [27], [28].

B. Motivation

This work aims to combine image moments with model predictive control strategies for UAV deformable target tracking in outdoor environments [29]. Previous works utilized traditional and hybrid IBVS schemes [30], [31], [32], effective but

Manuscript received: July, 24, 2023; Revised November, 3, 2023; Accepted November, 27, 2023.

This paper was recommended for publication by Editor Pascal Vasseur upon evaluation of the Associate Editor and Reviewers' comments. comments. This work was supported by the European Union's Horizon 2020 Research and Innovation Program PATHOCERT - Pathogen Contamination Emergency Response Technologies under Grant 883484. (Corresponding author: Sotirios N. Aspragkathos.)

[1] Sotirios N. Aspragkathos is with the School of Mechanical Engineering, Control Systems Lab, National Technical University of Athens, 15780 Athens, Greece saspragkathos@mail.ntua.gr

[2]George C. Karras is with the Department of Informatics and Telecommunications, University of Thessaly, 35100 Lamia, Greece gkarras@uth.gr

[3]Kostas J. Kyriakopoulos is with the Engineering Division and Center for AI & Robotics (CAIR), New York University, Abu Dhabi 128188, UAE kkyria@ny.edu

Digital Object Identifier (DOI): see top of this page.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

amenable to improvement. The proposed method extends the NMPC formulation of [31] into the image-moment paradigm.

C. Proposed Method

The study introduces an aperiodic vision-based NMPC with a triggering mechanism for autonomous visual surveillance, reducing computational effort and energy consumption. The dynamic model, extracted from image moments, tracks contour-based areas with evolving features. Real-time outdoor experiments demonstrate efficacy in achieving convergence and reducing triggering instances. The proposed work combines image moments target dynamics formulation with a vision-based NMPC scheme using an event-triggering mechanism, a novel contribution in UAV deformable target tracking.

D. Preliminary definitions

Some notations and fundamental mathematical definitions are listed with the aim of facilitating the reader's understanding of some notations or specific concepts.

- $diag()$ is the mathematical notation denoting a diagonal matrix whose non-diagonal entries are zero [33].
- Lipschitz function f is any function with a bounded first derivative. An f such that $|f(x) - f(y)| \leq C|x - y|$ for all x and y ,where C is a constant independent of x and y called the Lipschitz constant. [34]
- Positively invariant set: Considering a dynamical system of state variable x, derivative $\dot{x} = f(x)$, trajectory $x(t, x_0)$ and initial point x_0, then a set $\mathcal{O}$ can be defined as positively invariant if the condition $x_0 \in \mathcal{O}$ implies that $x(t, x_0) \in \mathcal{O} \ \forall \ t \geq 0$ [35].

E. Outline

The paper is organized as follows: Sections II and III discuss problem formulation and visual servoing preliminaries. Sections IV and V detail contour target perception, motion estimation, and dynamic model formulation. Section VI presents the event-triggered NMPC scheme, while Section VII provides a comparative, experimental study and a simulation performance analysis. Section VIII concludes the paper.

II. Problem Statement

Target tracking in robotics demands precision, posing challenges for conventional methods when dealing with evolving features like deformable contour-based regions. The dynamic characteristics of targets, such as changes in size, shape, or appearance, create complexities for existing tracking techniques. Consequently, the development of an autonomous vision-based target-tracking system capable of adapting to changing features in contour-based areas becomes imperative.

In response to these challenges, we propose a control strategy (see Fig. 2) utilizing image moments as a valuable object descriptor for achieving decoupled visual servo control. This strategy addresses constraints in vision-based NMPC, presenting an aperiodic vision-based deformable target-tracking technique with a triggering mechanism. By focusing on image moment-like parameters (centroid, area, orientation), the approach efficiently tracks objects with complex shapes and evolving features while reducing computational effort and energy consumption. Additionally, a hybrid MB/DD target motion state estimator is integrated into the control strategy.

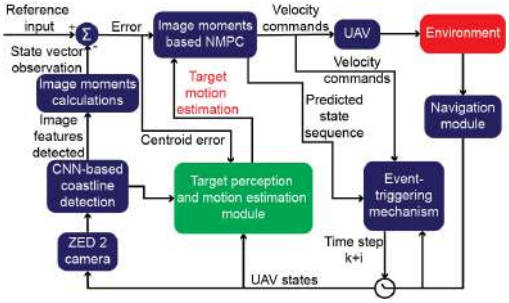


Fig. 2: Proposed control method presented in a block diagram form.

The target observation is initially obtained from the vehicle-mounted camera using a trained Deep Neural Network (DNN) algorithm. This information is processed to form the state error, and the hybrid MB/DD estimator produces an estimate. If the triggering mechanism is violated, this estimate is incorporated into the NMPC optimization problem. The resulting control sequence guides the vehicle in continuous target following and surveillance. However, if control inputs are used throughout the prediction horizon without triggering mechanism violation, the optimization is re-run to generate new velocity commands.

III. Preliminaries on Image Moments Visual Servoing

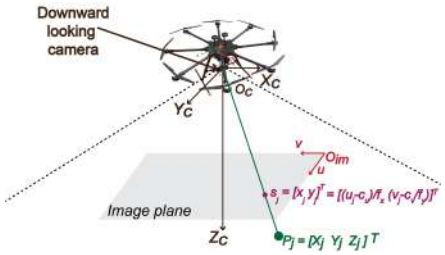


Fig. 3: Geometric representation of a downward-looking camera mounted on a UAV, during autonomous contour surveillance flight [31].

Addressing the contour tracking problem, we draw upon visual servoing concepts [36]. The goal is to track a 3D shape within the camera's environment, considering the shape as a whole. We assume a camera $\mathbf{C}$ with frame $[X_c, Y_c, Z_c]^T$ at center O_C (Fig. 3).

The image frame $\mathbf{I}_{im}$ has coordinates $[u, v]^T$ in pixels, with $O_{I_{im}}$ representing the upper-left corner. By utilizing the camera's geometric model, we project N points ($\mathbf{P_j} = [X_j, Y_j, Z_j]^T$, $j = 1, \ldots, N$) in the camera frame onto the normalized image plane as 2-D points $s_j = [x_j, y_j]^T$, $j = 1, \ldots, N$:

$$s_j = \left[x_j, y_j\right]^{\mathrm{T}} = \left[\frac{X_C}{Z_C}, \frac{Y_C}{Z_C}\right]^{\mathrm{T}} = \left[\frac{u_j - c_u}{f_x}, \frac{v_j - c_v}{f_y}\right]^{\mathrm{T}}, j = 1, \cdots, N \quad (1)$$

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

where c_u, c_v are the pixel coordinates of the primary point, f_x, f_y are the focal lengths, in pixels, for each image axis, and u_j, v_j are the pixel coordinates of the j-th feature. Assuming that any such points move on the physical, three-dimensional space along a path, then the following Ordinary Differential Equation (ODE) models the motion of any feature on the image plane [37]:

$$\dot{s}_j = L^j \nu + \frac{ds_j}{dt} \;\; j = 1, \cdots, N, \tag{2}$$

where $\nu = \left[\nu_x, \nu_y, \nu_z, \omega_x, \omega_y, \omega_z\right]$ denotes the velocity of the camera and L the overall interaction matrix. The function $\frac{ds_j}{dt}$ denotes (with abuse of notation) the time derivative of the motion of each feature on the image plane that is imparted owing to the motion of the features on the physical space.

The 3D shape target on a continuous surface assumes no depth-discontinuity, allowing the depth Z of a target's point as a continuous function of its image plane coordinates. For a deformable target with N point features and (x, y) coordinates from (1), image moments m_{ij} and centered moments μ_{ij} (order $(i + j)$) serve as statistical descriptors [6].

Considering a polygonal contour on the image plane with vertices $s \triangleq \left[s_1^{\mathrm{T}}, s_2^{\mathrm{T}}, \cdots, s_N^{\mathrm{T}}\right] \in \mathbb{R}^{2N}$ and each vertex $s_j = \left[x_j, y_j\right]^{\mathrm{T}}, j = 1, \cdots, N$, we denote the state vector:

$$\tilde{x} = [\bar{x}, \bar{y}, \bar{\sigma}, \bar{\alpha}]^{\mathrm{T}}, \in \mathbb{R}^4, \tag{3}$$

where $\bar{x} = \frac{m_{10}}{m_{00}}$, $\bar{y} = \frac{m_{01}}{m_{00}}$ represent the coordinates of the gravity center, $\bar{\sigma} = m_{00}$ the area of the target under surveillance, where $\bar{\sigma} \in \mathbb{R}_+$ and $\bar{\alpha}$ is a reference angle for the polygon's orientation $\bar{\alpha} = \frac{1}{2} arctan(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}})$.

After extracting analytical expressions over the chosen state vector components, we can now formulate the corresponding dynamics allowing us to implement an NMPC approach for solving the contour tracking task by using image moments as an analytical descriptor. The following Sections present the contour perception and motion estimation module and the formulation of the dynamics used for the control development of the even-triggered image moments-based NMPC scheme for autonomous surveillance.

IV. Target Perception and Motion Estimation Module

In our control approach, we employ a hybrid MB/DD approach for target perception and motion estimation [32]. Focusing on deformable contour-based areas with evolving features, the target detection module uses a Keras CNN trained on real UAV data [31], [38]. The CNN output is processed with FlowNet 2 [39] for optical flow estimation.

For target motion, we use KalmanNet [40], combining CNN-based coastline detection, CNN-based optical flow, and an EKF in a NN-aided framework. This allows motion estimation with partial knowledge of the target's dynamics. The output integrates into our control strategy, detailed in Sections V and VI.

V. Dynamic Model Formulation

In this section, we focus on formulating the dynamics of the state vector components followed by the proposed control development. The calculations presented herein pertain solely to the state vector's motion relative to the camera's movement, excluding the motion of features in physical space. According to the selected state vector and the utilization of the image moments [6], the full dynamics assume the following form:

$$\dot{\tilde{x}} = g(\tilde{x}; s)\nu, \tag{4}$$

where g has the form of the overall interaction matrix of image moments for finite horizon propagation of the NMPC [6], [21], [41] using the state vector established in (3) with the corresponding variables.

This indicated formulation leads to the complete dynamics in (4). While considering the dynamics of the vertices, this approach allows for computing the relevant dynamics at any time instant. In the subsequent NMPC scheme, the computations associated with the propagation of vertex dynamics are computationally efficient, resulting in a manageable scheme. Hence, the control development employs the following coupled dynamics model:

$$\begin{aligned} \dot{s} &= L\nu + \tfrac{ds}{dt} \\ \dot{\tilde{x}} &= g(\tilde{x}; s)\nu + \nabla_s^{\mathrm{T}} \tilde{x} \tfrac{ds}{dt} \end{aligned} \tag{5}$$

While the above equation could be evidently expressed in the form of a single system of Ordinary Differential Equations ODEs, the above distinction is maintained in the following NMPC formulation for clarity, both notational and in scope, as the two preceding subsystems are employed in an entirely different manner. Additionally, $\nabla_s^{\mathrm{T}} \tilde{x}$ expresses the derivative of the state vector $\tilde{x}$ with respect to the vertices comprising the contour-based area under tracking.

VI. Control Design

The control methodology based on (4) is being presented through the application of the Newton-Euler method. The discrete-time form of (5) at $k + 1$ with $k = 1, \ldots, n$, for a sampling time Δt is given:

$$\begin{aligned} s_{k+1} &= s_k + L_k \nu_k \Delta t + \tfrac{ds}{dt}\Big|_{t_k} \Delta t \\ \tilde{x}_{k+1} &= \tilde{x}_k + g(\tilde{x}_k; s_k)\nu_k \Delta t + \nabla_s^{\mathrm{T}} \tilde{x} \tfrac{ds}{dt}\Big|_{t_k} \Delta t \end{aligned} \tag{6}$$

where $s_k = \left[s_k^1, \ldots, s_k^N\right]^T$ is the vector of the features comprising the polygon and $\tilde{x}_k = [\bar{x}_k, \bar{y}_k, \bar{\sigma}_k, \bar{a}_k]^T$ is the discrete-time form of the proposed state vector at each time-step k. For the interaction matrix and camera velocity at sampling time k. The stacked version of nominal model (6) can be expressed with $\tilde{x}_{k+1} = f(\tilde{x}_k; s_k, \nu_k)$, where the camera velocities belong to $\nu_k \in \nu_{set} \subset \mathbb{R}^4$ and are bounded through the input constraints such as $\|\nu_x^k\| \leq \overline{\nu}_x, \|\nu_y^k\| \leq \overline{\nu}_y, \|\nu_z^k\| \leq \overline{\nu}_z, \|\omega_x^k\| \leq \overline{\omega}_x, \|\omega_y^k\| \leq \overline{\omega}_y, \|\omega_z^k\| \leq \overline{\omega}_z$. The upper bound of system's control input is denoted through $(\bar{\cdot})$. Furthermore, we assume that the system is affected by disturbances (i.e. change of lighting conditions, weather effects, dynamic background, shadows and occlusions, reflection and glare or even wildlife and insects in the Field of View (FoV)) caused during

the visual tracking algorithm, hence the stacked version of the system can be expressed in a more general form as:

$$\tilde{x}_{k+1} = f(\tilde{x}_k; s_k, \nu_k) + \xi_k \tag{7}$$

where ξ_k a vector representing these disturbances expressed at each time step as $\xi_k = \left[\xi_{\bar{x}_k}, \xi_{\bar{y}_k}, \xi_{\bar{\sigma}_k}, \xi_{\bar{a}_k}\right]^T$ where the disturbances are considered to belong to the set Ξ_{set} where $\xi_k \in \Xi_{set} \subset \mathbb{R}^4$. Meanwhile, the system must comply to the constraints of the state variables, which are formulated through the limits of the camera FoV resulting to the state constraint set $\tilde{X}_{set}$, which is defined as $\tilde{x}_k \in \tilde{X}_{set} \subset \mathbb{R}^4$:

$$\begin{aligned} \bar{x}^{min} \leq \bar{x}_k \leq \bar{x}^{max} \\ \bar{y}^{min} \leq \bar{y}_k \leq \bar{y}^{max} \\ \bar{\sigma}^{min} \leq \bar{\sigma}_k \leq \bar{\sigma}^{max} \\ \bar{a}^{min} \leq \bar{a}_k \leq \bar{a}^{max} \end{aligned} \tag{8}$$

where, $\bar{x}_k^{max}, \bar{x}_k^{min}$ and $\bar{y}_k^{max}, \bar{y}_k^{min}$ are the maximum and minimum bounds for the position of the center of gravity in the image plane. $\bar{\sigma}_k^{min}$ and $\bar{\sigma}_k^{max}$ are the maximum and minimum bounds of the value of the target's area defined from $\bar{\sigma} = m_{00}$, while $\bar{a}_k^{min}$ and $\bar{a}_k^{max}$ the maximum and minimum bounds of the α angle of the target's orientation. These bounds are defined by the nature of the application and the visual data specifications we need to acquire during the autonomous control procedure. For example, in the case of monitoring an oil spill, it is required to occupy a size of the image plane, and to be recorded from some angles, by extension, defining the boundaries of $\bar{\sigma}_k^{min}$, $\bar{\sigma}_k^{max}$ and $\bar{a}_k^{min}$ and $\bar{a}_k^{max}$.

A NMPC scheme aims to drive (7) to a desired state $\tilde{x}_k^{des} = [\bar{x}_k^{des}, \bar{y}_k^{des}, \bar{\sigma}_k^{des}, \bar{a}_k^{des})]^T \in \tilde{X}_{set} \subset \mathbb{R}^4$ through constant solving the optimization problem over a receding finite horizon of size n. The desired state values are determined while trying to maintain the target's center of gravity close to the center image plane, while the same stands for $\bar{\sigma}$ and $\bar{\alpha}$. According to the nature of the application and the visual data specification we may need to acquire during the autonomous control procedure, similar to the visual constraint bounds determination. The solution of the optimization problem is taking place online based on the constant observation of the states at the time step k.

This action requires the utilization of the dynamics established in Section V for the four state variables, $\bar{x}_k, \bar{y}_k, \bar{\sigma}_k, \bar{a}_k$. After each optimization run a n-sized control vector $\nu_{\mathrm{F}}(k) = [\nu(k|k), \nu(k+1|k, \ldots, \nu(k+n-1|k))]$ is generated which minimizes the cost function while simultaneously satisfying all the aforementioned state and input constraints of the system while the optimization problem of the NMPC scheme is formulated such as:

$$\begin{aligned} &\min_{\nu_{\mathrm{F}}(k)} \mathcal{J}_n(\tilde{x}_k; s_k, \nu_{\mathrm{F}}(k)) = \\ &\min_{\nu_{\mathrm{F}}(k)} \textstyle\sum_{i=0}^{n-1} \mathrm{F}(\hat{\tilde{x}}(k+i|k), \nu(k+i|k)) + \mathrm{E}(\hat{\tilde{x}}(k+N|k)) \\ &\text{subject to:} \\ &\hat{\tilde{x}}(k+i|k) \in \tilde{X}_{set} \quad \forall i = \{1, \ldots, n-1\} \\ &\nu(k+i|k) \in \nu_{set} \quad \forall i = \{1, \ldots, n-1\} \\ &\hat{\tilde{x}}(k+n|k) \in \mathrm{E}_f \end{aligned} \tag{9}$$

where n the prediction horizon, $\hat{\tilde{x}}$ the states error w.r.t. the desired state vector values, $\hat{\tilde{x}} = \tilde{x} - \tilde{x}^{des}$, $\mathrm{F}(\hat{\tilde{x}}, \nu) = \hat{\tilde{x}}^T Q \hat{\tilde{x}} + \nu^T R \nu$ the stage cost, while $\mathrm{E}(\hat{\tilde{x}}) = \hat{\tilde{x}}^T P \hat{\tilde{x}}$ is the terminal cost. All the positive definite matrices incorporated in (9) are defined such as: $Q = diag(q_1, \ldots, q_4)$, $R = diag(r_1, \ldots, r_4)$ and $P = diag(p_1, \ldots, p_4)$. To fully define the control problem and formulate the framework for NMPC, we introduce the following lemmas and assumptions.

Assumption 1: The nominal model has the following characteristics: the origin is a steady state, $f(\tilde{x}, \nu)$ is locally Lipshitz in $\tilde{x}$ in the domain $\tilde{X}_{set} \times \nu_{set}$, while finally there exists a $0 < \mathcal{L}_f < \infty$ such that for all $\tilde{x}_1, \tilde{x}_2 \in \tilde{X}_{set}$ $\forall$ $\nu \in \nu_{set}$,

$$\|f(\tilde{x}_1, \nu) - f(\tilde{x}_2, \nu)\|_j \leq \mathcal{L}_f \|\tilde{x}_1 - \tilde{x}_2\|_j \tag{10}$$

where the $\mathcal{L}_f$ constant denotes the Lipschitz constant in some j norm, i.e. $j = 1, 2, \ldots, \infty$.

Lemma 1: For the aforementioned system the error between the predicted and the actual system state observation at each $k+i$ should be limited within a certain range provided with an input vector:

$$|\hat{\tilde{x}}_{k+i} - \hat{\tilde{x}}(k+i|k-1)| \leq \textstyle\sum_{j=0}^{i-1} (\mathcal{L}_f)^j \xi \tag{11}$$

where $\sum_{j=0}^{i-1} (\mathcal{L}_f)^j = \frac{\mathcal{L}_f^i - 1}{\mathcal{L}_f - 1}$ and $\mathcal{L}_f$ is the Lipschitz constant of the nominal system.

Lemma 2: Considering the input and FoV constraints, we can determine its Lipschitz continuity of the nominal system (7) in $\tilde{X}_{set}$. The value of $\mathcal{L}_f$ is defined by the equation:

$$\mathcal{L}_f = [2 \max(4(1 + \tfrac{\bar{\nu}_z}{z} dt)^2, 4(\bar{\omega}_z dt)^2), 1]^{\frac{1}{2}} \tag{12}$$

Lemma 3: Let $\mathcal{L}_{\mathrm{F}}$ a Lipschitz constant for the stage cost. The stage cost is Lipschitz continuous while the constant stands for $\tilde{X}_{set}$ and $\mathcal{L}_{\mathrm{F}_\nu}$ for ν_{set}, while both constants are non-negative real numbers.

$$\mathcal{L}_{\mathrm{F}} = 2(|\bar{x}|^2 + |\bar{y}|^2 + |\bar{\sigma}|^2 + |\bar{\alpha}|^2)^{\frac{1}{2}} \sigma_{max}(Q) \tag{13}$$

where $\sigma_{max}(Q)$ represents the biggest singular value of matrix Q.

Lemma 4: We have to fill in that a K_∞-function provides a lower bound for the stage cost $\mathrm{F}(\tilde{x}; s, \nu)$:

$$\mathrm{F}(\tilde{x}; s, \nu) \geq \underline{\mathrm{F}}(\tilde{x}; s, \nu) = \min(q_1, \ldots, q_4, r_1, \ldots, r_4) |\tilde{x}|^2 \tag{14}$$

Assumption 2: Supposing that there exists a positively invariant set E_N, as defined in [35], within the set $\tilde{X}_{set}$ for the system described in equation (7). Then this set satisfies $\mathrm{E}_f \subset \mathrm{E}_N$, where E_N is defined as $\tilde{x} \in \tilde{X}_{set} : |\tilde{x}| \leq \epsilon_0$ and ϵ_0 is a positive value.

Assumption 3: Additionally, an existing controller denoted as $\nu_k = h(\tilde{x}_k) \in \nu_{set}$ that stabilizes the system within the set E_N. Subsequently, this controller is accompanied by a Lyapunov function, denoted as V, satisfying the inequality $\mathrm{V}(f(\tilde{x}_k, h(\tilde{x}_k))) - \mathrm{V}(\tilde{x}_k) \leq -\mathrm{F}(\tilde{x}_k, h(\tilde{x}_k))$ for all $\tilde{x}$ in E_N. Moreover, for any $\tilde{x}$ in E_N, there are positive constants $\mathcal{L}_h$ and $\mathcal{L}_{f_h}$ such that the following conditions hold: $|h(\tilde{x})| \leq \mathcal{L}_h |\tilde{x}|$ and $|f(\tilde{x}, h(\tilde{x}))| \leq \mathcal{L}_{f_h} |\tilde{x}|$.

Assumption 4: Lipschitz continuity of the terminal cost associated Lyapunov function $\mathrm{V}(\tilde{x})$ with respect to E_N is ensured by a Lipschitz constant $\mathcal{L}_{\mathrm{E}} = 2\epsilon_0 \sigma_{max}(P)$ for all $\tilde{x} \in E_N$:

$$\|\mathrm{V}(\tilde{x}_1) - \mathrm{V}(\tilde{x}_2)\| \leq \mathcal{L}_{\mathrm{E}} \|\tilde{x}_1 - \tilde{x}_2\| \tag{15}$$

Assumption 5: A local stabilizing controller $h(\tilde{x}) \in \nu_{set}$, exists such that there is a positive constant $\mathcal{L}_h > 0$ for all $\tilde{x} \in \mathrm{E}_N$, satisfying $|h(\tilde{x})| \leq \mathcal{L}_h|\tilde{x}|$. Consequently, it can be inferred that for any positive constant $\mathcal{L}_{f_h} > 0$ and all $\tilde{x} \in \mathrm{E}_N$, the condition $|f(\tilde{x}, h(\tilde{x}))| \leq \mathcal{L}_h|\tilde{x}|$ holds true.

Assumption 6: The Lyapunov function associated with the set E_N can be expressed as follows: $\mathrm{V}(\tilde{x}) = \tilde{x}^T P \tilde{x}$. This function is bounded and satisfies the condition $\mathrm{V}(\tilde{x}) \leq a_\epsilon$, where a_ϵ is a positive constant defined as $a_\epsilon = \max\{p_1, \ldots, p_4\}\epsilon_0^2 > 0$. Additionally, if we assume that E_N is defined as $\tilde{x} \in \tilde{X}_{set}(n-1) : h(\tilde{x}) \in \nu_{set}$, and we choose a positive parameter $a_{\epsilon f}$ such that $0 < a_{\epsilon f} < a_\epsilon$, we can state that the terminal set E_f is defined as $\tilde{x} \in \mathbb{R}^4 : \mathrm{V}(\tilde{x}) \leq a_\epsilon$. Moreover, for any $\tilde{x} \in \mathrm{E}_N$, the function $f(\tilde{x}, h(\tilde{x}))$ belongs to E_f.

In time-triggered NMPC, only the first control sequence element is implemented. The event-based policy triggers optimization when the error surpasses a threshold. As long as the error stays below this limit, the computed control sequence based on the current system state is applied in an open-loop manner. The problem aims to achieve two objectives: (i) develop a feedback control scheme ensuring Input to State Stability (ISS) for the nominal system under NMPC constraints, and (ii) determine the event-based condition for initiating the optimization problem solution, accounting for constraints and control sequence updates.

A. Feasibility & Convergence Analysis of the Event-triggering NMPC scheme

Let at $k_i \triangleq k-i$ the NMPC optimization problem is solved for the first time, generating an optimal control velocity vector $\nu_{\mathrm{F}}^*(k-1) \triangleq [\nu_{\mathrm{F}}^*(k-1|k-1), \ldots, \nu_{\mathrm{F}}^*(k+n-2|k-1)]$, for which the $*$ notation is used. Assuming that the state vectors are all contained in $\tilde{X}_{\mathrm{F}}$, where a feasible generated velocity vector satisfies all the optimization problem constraints. In this case, any $\bar{\nu}_{\mathrm{F}}(k+m)$ at a subsequent $k+m$ with $m = 0, \ldots, n-1$ is described by:

$$\bar{\nu}_{\mathrm{F}}(k+m) = \bar{\nu}(k+i|k+m) = \begin{cases} \nu^*(k+i|k-1) & for\ i = m, \ldots, n-2 \\ h(\hat{\tilde{x}}(k+i|k+m)) & for\ i = n-1, \ldots, n+m-1 \end{cases} \tag{16}$$

where a part of the optimal velocity vector generated at $k-1$ is $\nu^*(k+i|k-1)$ for $i = m, \ldots, n-2$. The generated at $k-1$, optimal control vector $\nu_{\mathrm{F}}^*(k-1)$ denotes that at all subsequent time steps $m = 0, \ldots, n-1$ can also belong to the ν_{set}, i.e. $\bar{\nu}(k+i|k+m) \in \nu_{set}$. To improve recursive feasibility, it is necessary to prove that $\hat{\tilde{x}}(k+n-1|k+m) \in \mathrm{E}_f$ for all time steps. Utilizing (11), we can derive:

$$\|\hat{\tilde{x}}(k+n-1|k+m) - \hat{\tilde{x}}(k+n-1|k-1)\| \leq \mathcal{L}_f^{(n-1)-m} \textstyle\sum_{i=0}^{m} \mathcal{L}_f^i \xi \tag{17}$$

Utilizing the terminal cost function Lipschitz constant from (15), we can derive:

$$\begin{aligned} \mathrm{E}(\hat{\tilde{x}}(k+n-1|k+m)) - \mathrm{E}(\hat{\tilde{x}}(k+n-1|k-1)) \leq \\ \mathcal{L}_{\mathrm{E}}\|\hat{\tilde{x}}(k+n-1|k+m) - \hat{\tilde{x}}(k+n-1|k-1)\| \leq \\ \mathcal{L}_{\mathrm{E}}\mathcal{L}_f^{(n-1)-m} \textstyle\sum_{i=0}^{m} \mathcal{L}_f^i \xi \end{aligned} \tag{18}$$

Exploiting $\hat{\tilde{x}}(k+n-1|k-1) \in \mathrm{E}_f$ as initial feasibility, we may utilize Assumption 6 to derive that $\mathrm{E}(\hat{\tilde{x}}(k+n-1|k-1)) \leq a_{\epsilon f}$. While the aim is the proof of $\hat{\tilde{x}}(k+n-1|k+m) \in \mathrm{E}_f$, Assmption 6 proves that:

$$\begin{aligned} \mathrm{E}(\hat{\tilde{x}}(k+n-1|k+m)) \leq a_\epsilon \Longrightarrow \\ a_{\epsilon f} + \mathcal{L}_{\mathrm{E}}\mathcal{L}_f^{(n-1)-m} \textstyle\sum_{i=0}^{m} L_f^i \xi \leq a_\epsilon \Longrightarrow \\ \xi \leq \frac{a_\epsilon - a_{\epsilon f}}{\mathcal{L}_{\mathrm{E}}} \mathcal{L}_f^{(n-1)-m} \textstyle\sum_{i=0}^{m} \mathcal{L}_f^i \end{aligned} \tag{19}$$

The aforementioned mathematical expression states that as long as (19) bounds the disturbances of the system then a robust positively invariant set $\tilde{X}_{\mathrm{F}}$ and a local stabilizing controller $h(\tilde{x})$ ensure that $\hat{\tilde{x}}(k+n-1|k+m) \in \mathrm{E}_f$ for all $m = 0, \ldots, n-1$.

For the proof of ISS stability under an event triggered NMPC scheme, the cost function of the optimization problem under the optimal velocity vector ν_{F}^* is chosen as a Lyapunov function. Formulating $\Delta\mathcal{J}_m^* = \mathcal{J}_n^*(k+m) - \mathcal{J}_n^*(k-1)$ as the optimal cost difference between the initial triggering instant $k-1$ and a subsequent step $k+m$ we can extract the following Lemma:

Lemma 5: Let the (7) satisfy all the state, input constraints, and all the aforementioned conditions and assumptions. In this case, we can denote that the difference between the optimal cost value at $k-1$ and all subsequent steps $k+m$, with $m = 0, \ldots, n-1$, is bounded as described by:

$$\Delta\mathcal{J}_m^* \leq \mathcal{L}_{z_m}|\hat{\tilde{x}}_{k+m} - \hat{\tilde{x}}(k+m|k-1)| - \textstyle\sum_{i=0}^{m} \underline{\mathrm{F}}(\|\tilde{x}_{k-1-i+m}\|) \tag{20}$$

where the $\mathcal{L}_{z_m}$ is defined by: $\mathcal{L}_{z_m} = \mathcal{L}_{\mathrm{E}}\mathcal{L}_f^{(n-1)-m} + \mathcal{L}_{\mathrm{F}}\frac{\mathcal{L}_f^{(n-1)-m}-1}{\mathcal{L}_f-1}$.

Utilizing [42] the event-triggering policy ensures operation with ISS while:
(i) there exists a positively invariant set, $\tilde{X}_{\mathrm{F}}$, of the system.
(ii) the existence of the functions $\alpha_1, \alpha_2, \alpha_3 \in K_\infty$ and $\sigma_1 \in K$ satisfies the two conditions following:

$$\begin{aligned} \alpha_1(\|\tilde{x}\|) \leq \mathcal{J}(k) \leq \alpha_2(\|\tilde{x}\|) \\ \mathcal{J}(k+m) - \mathcal{J}(k-1) \leq -\alpha_3(\|\tilde{x}\|) + \sigma_1(\|\hat{\tilde{x}}\|) \end{aligned} \tag{21}$$

The aforementioned Lyapunov function operates under upper and lower bounds that have been defined by (19) and utilized through combination with the optimality of the NMPC OCP solution resulting to:

$$\alpha_1(\|\tilde{x}\|) \leq \mathcal{J}_N^*(k) \leq \alpha_2(\|\tilde{x}\|) \tag{22}$$

where $\alpha_1(\|\tilde{x}\|) = \underline{\mathrm{F}}(\|\tilde{x}_k\|)$ and $\alpha_2(\|\tilde{x}\|) = (\mathcal{L}_{\mathrm{F}} + \mathcal{L}_{\mathrm{F}_\nu}\mathcal{L}_h)\frac{\mathcal{L}_{f_h}^N - 1}{\mathcal{L}_{f_h} - 1}\|\tilde{x}_k\| + \mathcal{L}_{\mathrm{E}}\mathcal{L}_{f_h}\|\tilde{x}_k\|$. Exploiting (20) it is denoted that:

$$\begin{aligned} \sigma_1(\|\hat{\tilde{x}}\|) = \mathcal{L}_{z_m}|\hat{\tilde{x}}_{k+m} - \hat{\tilde{x}}(k+m|k-1)| \\ \alpha_3(\|\tilde{x}\|) = \textstyle\sum_{i=0}^{m} \underline{\mathrm{F}}(\|\tilde{x}_{k-1-i+m}\|) \end{aligned} \tag{23}$$

Hence, the system achieves input to state stability under the event-triggering NMPC method while $\sigma_1(\|\hat{\tilde{x}}\|) \leq \epsilon\alpha_3(\|\tilde{x}\|)$ where ϵ is a positive constant $0 < \epsilon < 1$ and the triggering condition ruling all the above is:

$$\mathcal{L}_{z_m}|\hat{\tilde{x}}_{k+m} - \hat{\tilde{x}}(k+m|k-1)| \leq \epsilon \textstyle\sum_{i=0}^{m} \underline{\mathrm{F}}(\|\tilde{x}_{k-1-i+m}\|) \tag{24}$$

Finally the OCP of the NMPC scheme will be solved again if the triggering condition (24) is violated.

B. Implementation Details

1) Multirotor Under-actuation: The study uses cascaded P/PID controllers for multirotor control. Attitude control stabilizes orientation and angular velocity, while velocity control manages linear and yaw rate velocity commands. UAVs typically use velocity control mode to address underactuation. Effective control requires interaction matrices considering kinematic capabilities and omitting corresponding columns for under-actuation [13].

2) Multirotor low-level control: Cascaded P/PID controllers ensure stability and regulate attitude and velocity in the vehicle. Underactuation is addressed through velocity control mode. The deployed UAV uses Pixhawk Cube Orange with ArduPilot framework for a fully integrated low-level control architecture. Velocities are computed in the camera frame $\mathbf{C}$ and transformed to the vehicle body frame $\mathbf{B}$ [43].

VII. Results

A. Experimental Validation of the Proposed Method

1) Experimental Setup: Two sets of experiments were conducted to evaluate the proposed event-triggered image moments NMPC against a time-triggered counterpart. A custom octocopter equipped with a ZED 2 Stereo Camera and Pixhawk Cube Orange running ArduPilot firmware served as the experimental platform. The Robot Operating System (ROS) [44] facilitated target recognition, optical flow prediction, Extended Kalman Filtering, and vision-based NMPC. MAVROS communication protocol [45] conveyed velocity commands to the Pixhawk microprocessor. The detailed octocopter low-level control is in Section VI-B2.

2) Experimental Results: Experiments focused on dynamic shoreline tracking using the proposed NMPC. Vision-based NMPC operated below 20 m above sea level, assuming constant depth for all target features. Environmental variables remained exogenous and immeasurable. Motion estimates were derived from the target perception module, detailed in Fig. 4.

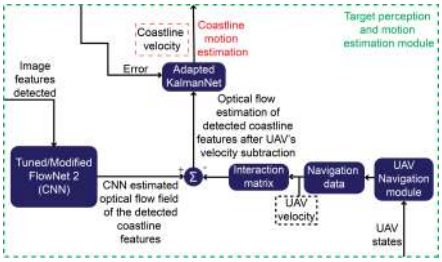
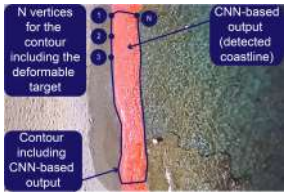


Fig. 4: Detailed view of the green block in Fig. 2, representing the target perception and motion estimation module.

Fig. 5 illustrates the detection of a dynamic shoreline, with subsequent optical flow prediction and motion estimation using a hybrid MB/DD framework [32].

Fig. 6 compares two scenarios: time-triggered NMPC (blue line) and proposed event-triggered NMPC (orange line). The



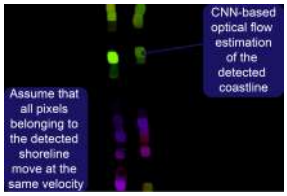


Fig. 5: Results of the trained CNN results for coastline detection (**left**) and optical flow estimation (**right**).

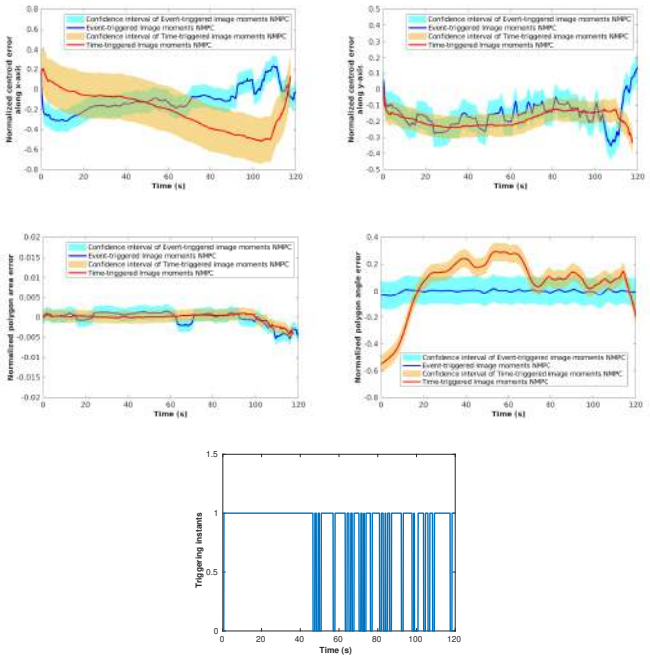


Fig. 6: Experimental validation results for a time- (**blue line**) and an event-triggered (**orange line - proposed strategy**) image moments NMPC: Normalized centroid error along x-axis (**upper left**), y-axis (**upper right**), normalized polygon area error (**middle left**), normalized polygon angle error (**middle right**) with respect to y-axis of the image plane during the experimental scenarios conducted in a beach setting. In all cases the confidence interval calculated over all the experimental trials for each method, is depicted. Triggering instances (**bottom**), each time we get value 1.0, denoting the times that the (24) is violated and the OCP is running again.

latter demonstrates successful coastline tracking while reducing optimization instances by $24-30\%$ across 10 sessions.

Errors were computed over 10 sessions for each controller. Fig. 7 displays statistics for state variables, including mean, minimum, maximum values, and standard deviation. In Fig. 7, the event-triggered scheme outperforms the time-triggered one, as seen in VII-A3. This is attributed to unaccounted factors like wind gusts. A comparative study in the next section with [32] and [31] provides conclusive insights.

3) Experimental video: A video demonstrating the aforementioned experimental comparative results of the proposed methodology can be found at the following URL: https://youtu.be/7fE9W2vPsnA

B. Comparative Simulation Results

1) Simulation Setup: To evaluate the proposed strategy in a realistic scenario, a synthetic coastline environment is

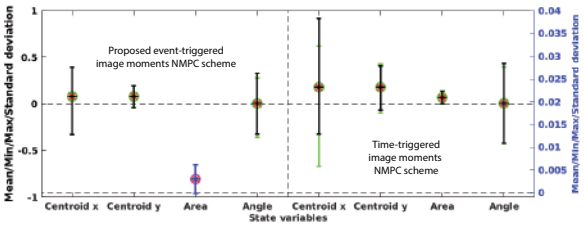


Fig. 7: Average comparative performance analysis between the proposed controller and a time-triggering image moments NMPC for the coastline UAV tracking. The figure is color-coded according to the different axes, black on the left and blue on the right axes.

employed, as described in [32], using Robot Operating System (ROS) [44], Gazebo [46], and MAVROS [45] as SITL communication protocol.

2) Simulation Results: Five comparative contour surveillance scenarios of a synthetic coastline environment are considered:

- Simulation scenario 1: Partitioned Visual Servo (PVS) controller of [32].
- Simulation scenario 2: Time-triggered IBVS-NMPC scheme of [31].
- Simulation scenario 3: Event-triggered IBVS-NMPC scheme of [31].
- Simulation scenario 4: Time-triggered image moments NMPC scheme presented in Section VII-A2.
- Simulation scenario 5: Event-triggered image moments NMPC (**proposed method**) scheme presented in Section VII-A2.

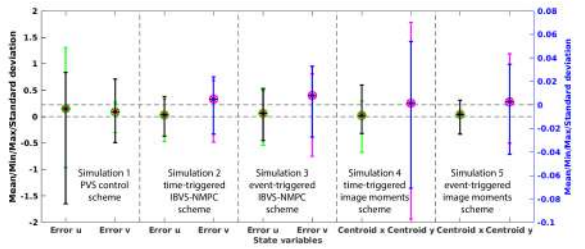


Fig. 8: Average comparative performance analysis between the 5 Simulation scenarios presented: PVS control scheme, time-triggered IBVS-NMPC scheme, event-triggered IBVS-NMPC scheme, time-triggered image moments NMPC scheme, **proposed** event-triggered image moments NMPC scheme. The figure is color-coded according to the different axes, black on the left and blue on the right axes.

Based on 50 sessions executed per scenario, statistical metrics (mean, min, max, standard deviation) for all state variables are calculated and presented in Fig. 8. The performance ranking is as follows:

1) Simulation scenario 4 deploying the time-triggered image moments NMPC of the current work.
2) Simulation scenario 5 deploying the **proposed** event-triggered image moments NMPC scheme of the current work.
3) Simulation scenario 2 deploying the time-triggered IBVS-NMPC scheme presented in [31].
4) Simulation scenario 3 deploying the event-triggered IBVS-NMPC presented in [31], also managing a $15-22\%$ reduction of the triggering instants.
5) Simulation scenario 1 deploying the PVS control scheme of [32].

Results indicate that the **proposed** event-triggered image moments NMPC method, while not having the highest accuracy, offers significant power efficiency with a $30-35\%$ reduction in triggering instants. This makes it suitable for tracking deformable targets using UAVs with enhanced energy endurance.

VIII. Conclusion

This work presents a novel approach for autonomously tracking moving deformable targets comprised of evolving features using UAVs. It utilizes Visual Servoing NMPC and image moments to effectively track evolving features. The proposed scheme was successfully tested in beach settings, tracking a coastline with an octocopter. By incorporating a motion target estimate and an event-triggering framework, computing resources are minimized without sacrificing performance. This architecture enhances autonomy and endurance in small autonomous robotic systems like UAVs.

References

[1] F. Chaumette and S. Hutchinson, "Visual servo control. i. basic approaches," *IEEE Robotics & Automation Magazine*, vol. 13, no. 4, pp. 82–90, 2006.

[2] F. Chaumette and S. Hutchinson, "Visual servo control. ii. advanced approaches [tutorial]," *IEEE Robotics & Automation Magazine*, vol. 14, no. 1, pp. 109–118, 2007.

[3] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE transactions on robotics and automation*, vol. 12, no. 5, pp. 651–670, 1996.

[4] G. Silveira and E. Malis, "Direct visual servoing: Vision-based estimation and control using only nonmetric information," *IEEE Transactions on Robotics*, vol. 28, no. 4, pp. 974–980, 2012.

[5] C. Kanellakis and G. Nikolakopoulos, "Survey on computer vision for uavs: Current developments and trends," *Journal of Intelligent & Robotic Systems*, vol. 87, no. 1, pp. 141–168, 2017.

[6] F. Chaumette, "Image moments: a general and useful set of features for visual servoing," *IEEE Transactions on Robotics*, vol. 20, no. 4, pp. 713–723, 2004.

[7] N. Mansard, A. Remazeilles, and F. Chaumette, "Continuity of varying-feature-set control laws," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2493–2505, 2009.

[8] C. Collewet and F. Chaumette, "Visual servoing on non-planar objects from active vision," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pp. 2446–2451, IEEE, 2007.

[9] O. Bourquardez, R. Mahony, T. Hamel, and F. Chaumette, "Stability and performance of image based visual servo control using first order spherical image moments," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4304–4309, IEEE, 2006.

[10] A. Dani, G. Panahandeh, S.-J. Chung, and S. Hutchinson, "Image moments for higher-level feature based navigation," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 602–609, IEEE, 2013.

[11] F. Allgower, R. Findeisen, Z. K. Nagy, *et al.*, "Nonlinear model predictive control: From theory to application," *Journal-Chinese Institute Of Chemical Engineers*, vol. 35, no. 3, pp. 299–316, 2004.

[12] D. Lee, T. Ryan, and H. J. Kim, "Autonomous landing of a vtol uav on a moving platform using image-based visual servoing," in *2012 IEEE international conference on robotics and automation*, pp. 971–976, IEEE, 2012.

[13] G. C. Karras, C. P. Bechlioulis, G. K. Fourlas, and K. J. Kyriakopoulos, "Target tracking with multi-rotor aerial vehicles based on a robust visual servo controller with prescribed performance," in *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 480–487, IEEE, 2020.

[14] P. Vlantis, P. Marantos, C. P. Bechlioulis, and K. J. Kyriakopoulos, "Quadrotor landing on an inclined platform of a moving ground vehicle," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2202–2207, IEEE, 2015.

[15] H. Jabbari Asl, G. Oriolo, and H. Bolandi, "Output feedback image-based visual servoing control of an underactuated unmanned aerial vehicle," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 228, no. 7, pp. 435–448, 2014.

[16] H. J. Asl and H. Bolandi, "Robust vision-based control of an underactuated flying robot tracking a moving target," *Transactions of the Institute of Measurement and Control*, vol. 36, no. 3, pp. 411–424, 2014.

[17] M. A. Kassab, A. Maher, F. Elkazzaz, and Z. Baochang, "Uav target tracking by detection via deep neural networks," in *2019 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 139–144, IEEE, 2019.

[18] C. Sampedro, A. Rodriguez-Ramos, I. Gil, L. Mejias, and P. Campoy, "Image-based visual servoing controller for multirotor aerial robots using deep reinforcement learning," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 979–986, IEEE, 2018.

[19] S. Lee, T. Shim, S. Kim, J. Park, K. Hong, and H. Bang, "Vision-based autonomous landing of a multi-copter unmanned aerial vehicle using reinforcement learning," in *2018 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 108–114, IEEE, 2018.

[20] C. Copot, C. Lazar, and A. Burlacu, "Predictive control of nonlinear visual servoing systems using image moments," *IET control theory & applications*, vol. 6, no. 10, pp. 1486–1496, 2012.

[21] A. Burlacu, C. Copot, and C. Lazar, "Predictive control architecture for real-time image moments based servoing of robot manipulators," *Journal of Intelligent Manufacturing*, vol. 25, pp. 1125–1134, 2014.

[22] S. Heshmati-Alamdari, A. Eqtami, G. C. Karras, D. V. Dimarogonas, and K. J. Kyriakopoulos, "A self-triggered visual servoing model predictive control scheme for under-actuated underwater robotic vehicles," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3826–3831, IEEE, 2014.

[23] S. Heshmati-Alamdari, G. C. Karras, A. Eqtami, and K. J. Kyriakopoulos, "A robust self triggered image based visual servoing model predictive control scheme for small autonomous robots," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5492–5497, IEEE, 2015.

[24] X. Lou and Z. Jiang, "Event-triggered control of spatially distributed processes via unmanned aerial vehicle," *International Journal of Advanced Robotic Systems*, vol. 13, no. 6, p. 1729881416675138, 2016.

[25] J.-F. Guerrero-Castellanos, A. Vega-Alonzo, N. Marchand, S. Durand, J. Linares-Flores, and G. Mino-Aguilar, "Real-time event-based formation control of a group of vtol-uavs," in *2017 3rd International Conference on Event-Based Control, Communication and Signal Processing (EBCCSP)*, pp. 1–8, IEEE, 2017.

[26] Z. Xuan-Zhi, P. Casau, and C. Silvestre, "Global trajectory tracking for a quadrotor through event-triggered control: Synthesis, simulations, and experiments," in *2020 American Control Conference (ACC)*, pp. 3852–3857, IEEE, 2020.

[27] Z. Cai, H. Zhou, J. Zhao, K. Wu, and Y. Wang, "Formation control of multiple unmanned aerial vehicles by event-triggered distributed model predictive control," *IEEE Access*, vol. 6, pp. 55614–55627, 2018.

[28] D. Jang, C. Y. Son, J. Yoo, H. J. Kim, and K. H. Johansson, "Efficient networked uav control using event-triggered predictive control," *IFAC-PapersOnLine*, vol. 52, no. 15, pp. 412–417, 2019.

[29] K. Zhang, Y. Shi, and H. Sheng, "Robust nonlinear model predictive control based visual servoing of quadrotor uavs," *IEEE/ASME Transactions on Mechatronics*, vol. 26, no. 2, pp. 700–708, 2021.

[30] S. N. Aspragkathos, G. C. Karras, and K. J. Kyriakopoulos, "A visual servoing strategy for coastline tracking using an unmanned aerial vehicle," in *2022 30th Mediterranean Conference on Control and Automation (MED)*, pp. 375–381, IEEE, 2022.

[31] S. N. Aspragkathos, M. Sinani, G. C. Karras, F. Panetsos, and K. J. Kyriakopoulos, "An event-triggered visual servoing predictive control strategy for the surveillance of contour-based areas using multirotor aerial vehicles," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2203–2210, IEEE, 2022.

[32] S. N. Aspragkathos, G. C. Karras, and K. J. Kyriakopoulos, "A hybrid model and data-driven vision-based framework for the detection, tracking and surveillance of dynamic coastlines using a multirotor uav," *Drones*, vol. 6, no. 6, p. 146, 2022.

[33] J. Nearing, *Mathematical tools for physics*. 2003.

[34] M. O'Searcoid, *Metric spaces*. Springer Science & Business Media, 2006.

[35] A. Benzaouia, *Saturated switching systems*, vol. 426. Springer Science & Business Media, 2012.

[36] É. Marchand and F. Chaumette, "Feature tracking for visual servoing purposes," *Robotics and Autonomous Systems*, vol. 52, no. 1, pp. 53–70, 2005.

[37] F. Chaumette, "Potential problems of stability and convergence in image-based and position-based visual servoing," in *The confluence of vision and control*, pp. 66–78, Springer, 2007.

[38] D. Gupta, "A beginner's guide to deep learning based semantic segmentation using keras," 2019.

[39] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, "Flownet 2.0: Evolution of optical flow estimation with deep networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2462–2470, 2017.

[40] G. Revach, N. Shlezinger, X. Ni, A. L. Escoriza, R. J. Van Sloun, and Y. C. Eldar, "Kalmannet: Neural network aided kalman filtering for partially known dynamics," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1532–1547, 2022.

[41] O. Tahri and F. Chaumette, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Transactions on Robotics*, vol. 21, no. 6, pp. 1116–1127, 2005.

[42] M. Lazar, W. M. H. Heemels, and A. R. Teel, "Further input-to-state stability subtleties for discrete-time systems," 2010.

[43] A. P. group, "Ardupilot documentation," 2016.

[44] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng, *et al.*, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, p. 5, Kobe, Japan, 2009.

[45] M. M. A. V. Protocol, "Mavlink to ros gateway with proxy for ground control station," 2013.

[46] N. Koenig and A. Howard, "Design and use paradigms for gazebo, an open-source multi-robot simulator," in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*, vol. 3, pp. 2149–2154, IEEE, 2004.