QP-based Visual Servoing Under Motion Blur-Free Constraint

Maxime Robic, Renaud Fraisse, Eric Marchand, François Chaumette

To cite this version:

Maxime Robic, Renaud Fraisse, Eric Marchand, François Chaumette. QP-based Visual Servoing Under Motion Blur-Free Constraint. IEEE Robotics and Automation Letters, 2024, 9 (10), pp.8738-8745. 10.1109/LRA.2024.3451394 . hal-04678481

QP-based Visual Servoing Under Motion Blur-Free Constraint

Maxime Robic[1], Renaud Fraisse[2], Eric Marchand[3], and François Chaumette[1]

***Abstract*—This work proposes a QP-based visual servoing scheme for limiting motion blur during the achievement of a visual task. Unlike traditional image restoration approaches, we want to avoid any deconvolution step by keeping the image sequence acquired by the camera as sharp as possible. To do so, we select the norm of the image gradient as sharpness metric, from which we design a velocity constraint that is injected in a QP controller. Our system is evaluated for an Earth observation satellite. Simulation and experimental results show the effectiveness of our approach.**

***Index Terms*—Visual servoing, space robotics and automation, visual tracking.**

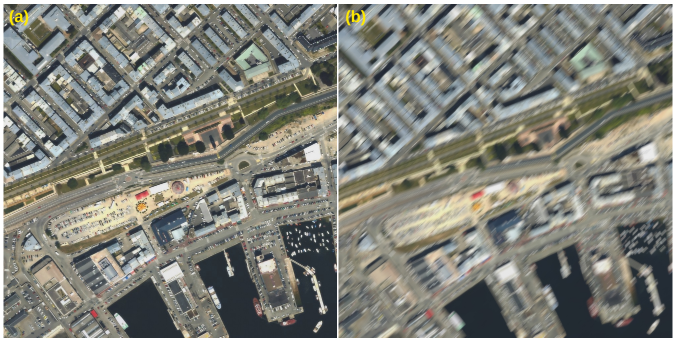

Fig 1. Satellite image of Brest (a) Sharp (b) Motion-blurred (20px,45°)

I. INTRODUCTION

VISUAL Servoing (VS) [6] is a vision-based control approach that relies on visual features extracted from image data. The reliability of the system is highly linked to the feature extraction step, which can be degraded if the image is too blurred. However, high camera velocity or fast-moving objects can create motion blur in the image. Indeed, motion blur is due to the camera and/or objects motion during the integration time for acquiring an image [10]. In VS, it induces a degradation of the image quality, which may perturb the feature extraction step and eventually make the control fail. In this work, we aim to control the orientation of an Earth observation satellite equipped with a snapshot matrix camera, such as [28], in order to center a potentially moving target in the image. This system can be seen as a pinhole camera in a pan-tilt configuration subject to external motions induced by satellite orbit. Pan-tilt motion control using classical VS [30], [12], [18] may lead to motion-blurred images (see Fig. 1). We propose in this paper to constrain the camera rotational velocity using a Quadratic Programming (QP) scheme for limiting motion blur as much as possible while the satellite performs its task.

Motion blur is characterized by its *point spread function* (PSF), or kernel, that can take various forms depending on the cause of the blur, and part of the literature is focused on how to implement a given PSF in image rendering [29], [26], either to enhance the perceived motion in a single snapshot or to increase the realism in movies or video games [24]. Another major stake is to estimate the PSF of motion-blurred images and to recover sharp images through deconvolution. The parameters estimation is usually processed on a single-blurred image, where the length and the orientation of the blur are evaluated through image gradients and spectral approaches [39], [35]. While these studies are often focused on linear motion blur, relevant works have also been made to handle more general motion types [15], such as nonlinear, rotational, or space-variant [10], [8], [27]. Then, motion-blurred image restoration has been extensively tackled with impressive results: for instance, while [32] presents a probabilistic model to deblur a single image, [7] slightly enhances the computation time by introducing a prediction step in the PSF estimation and [38] proposes a robust framework refining the estimation process together with noise suppression during the image restoration step. Finally, learning-based approaches have also been proposed to deblur an image without any estimation of the PSF [25], [20]. Although all these studies show restored images of high quality, they need high computational power and their current computational time is not compatible with real-time processing. Since VS is a closed-loop control scheme based on measurements extracted on each image acquired, and since our ultimate goal is to embed all computations on-board, which induces low-computational process in case of satellites, the previous approaches cannot be employed. So, our method is not based on deblurring the image but uses a measure of the sharpness directly in the control scheme. Several sharpness metrics have been studied in the literature [37]: norm of the image gradient [19], [33], norm of the image Laplacian [19], mutual information [11],

---

Manuscript received: April, 10, 2024; Revised July, 4, 2024; Accepted August, 12, 2024.

This paper was recommended for publication by Giuseppe Loianno upon evaluation of the Associate Editor and Reviewers' comments.

This work was supported by BPI France Lichie project. Experiments presented in this paper were carried out thanks to a platform of the Robotex 2.0 French research infrastructure.

[1]M. Robic and F. Chaumette are with Inria, Univ Rennes, CNRS, IRISA, Rennes, France `Firstname.Name@inria.fr`

[2]R. Fraisse is with Airbus Defense and Space, Toulouse, France `renaud.fraisse@airbus.com`

[3]E. Marchand is with Univ Rennes, Inria, CNRS, IRISA, Rennes, France `eric.marchand@irisa.fr`

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

eigenvalues of the image covariance matrix [37], and other statistical functions like normalized variance, auto-correlation, or standard-deviation-based correlation [34]. In this study, we have selected the norm of the image gradient that has already been considered in visual servoing [21][22], especially for autofocusing [9]. In these works, a single objective is considered that consists in maximizing the sharpness, while we also have to simultaneously consider the achievement of the primary visual task, that is, centering and tracking the target. As already said, this is achieved by limiting motion blur during the camera motion by considering it as a constraint in the control scheme.

Simple constraints in visual servoing such as limiting the camera velocity and acceleration can be dealt using ad-hoc techniques [23], [30]. More difficult constraints, such as avoiding joint limits and ensuring the visibility of the target, have to be considered with more elaborated methods such as adding features [17] or using the task-priority framework in case the system has some redundancy [16]. Optimal control, such as Quadratic Programming (QP) and Model-Predictive Control (MPC), is also an appealing general approach effective when multiple constraints have to be ensured simultaneously. QP has already been applied to VS [1], the same for MPC [2], [14], [31]. In [13] and [36], an MPC has been designed to limit motion blur while optimizing the trajectory of a quadrotor with respect to perception features. In both papers, the control tackles motion blur by considering the reduction of the projected velocity of a landmark in the cost function. The main differences with our approach are that we explicitly consider motion blur through a sharpness metric measured in the current image, and motion blur limitation is expressed as a constraint instead of a supplementary term in the cost function, which is impossible in our case due to the lack of redundancy in pan-tilt control. Finally, QP has been preferred to MPC due to the low computing resources on a satellite.

To sum up, the contributions of this paper are to propose a QP controller where motion blur limitation is expressed as a constraint to be handled simultaneously with a centering and target tracking task. It is based on the norm of the image gradients whose Jacobian with respect to the control inputs is determined so that a velocity constraint is injected into the controller. A predictive step is also introduced to improve the behavior of the system.

The paper is ordered as follows: Section II presents the QP-based VS controller and the modeling of the centering task with a pan-tilt configuration. Section III discusses the parameters affecting our sharpness metric, from which its Jacobian is computed. In Section IV we express the sharpness constraint through a control barrier function for several configurations of the system: fixed pan-tilt, pan-tilt subject to external motions, and target tracking. Experimental results based on a satellite path simulator and Earth images are presented in Section V to validate our approach for these different use cases. Finally, Section VI concludes on the results of our work.

II. CONSTRAINED VISUAL SERVOING

In this section, we present a simple formulation to design an IBVS controller of a pan-tilt camera under constraints. As this system has only 2 degrees of freedom, the visual feature chosen is simply an image point $\mathbf{x} = (x, y)$ belonging to the target, typically its center of gravity, and its desired position at the image center is denoted $\mathbf{x}^*$. Thus, the visual servoing control law aims to regulate to 0 a visual error $\mathbf{e}$ defined by

$$\mathbf{e} = \mathbf{x} - \mathbf{x}^* \tag{1}$$

The dynamic equation that links the variation of the visual error with the camera pan-tilt velocity $\boldsymbol{\omega}_c = (\omega_x, \omega_y)$ is then:

$$\dot{\mathbf{e}} = \mathbf{L}_{\boldsymbol{\omega}} \boldsymbol{\omega}_c \tag{2}$$

where the pan-tilt part $\mathbf{L}_{\boldsymbol{\omega}}$ of the interaction matrix related to $\mathbf{x}$ is given by [6]:

$$\mathbf{L}_{\boldsymbol{\omega}} = \begin{pmatrix} xy & -(1+x^2) \\ 1+y^2 & -xy \end{pmatrix} \tag{3}$$

A. QP-formulation

In classical IBVS, the control scheme is directly designed through inverse kinematics. In QP-based approach, it is expressed as a quadratic cost function to be minimized wrt. the control variables, here $\boldsymbol{\omega}_c$. Additionally, every constraint has to be expressed linearly with respect to these variables. From (2) and considering velocity, acceleration and motion blur-free constraints, we obtain the following optimization system [16]:

$$\begin{aligned} \boldsymbol{\omega}_c^* = \arg\min_{\boldsymbol{\omega}_c} \quad & \| \mathbf{L}_{\boldsymbol{\omega}} \boldsymbol{\omega}_c - \dot{\mathbf{e}}_d \|^2 \\ \text{s.t.} \quad & -\boldsymbol{\omega}_{max} \leq \boldsymbol{\omega}_c \leq \boldsymbol{\omega}_{max} \\ & -\boldsymbol{\gamma}_{max} \leq \boldsymbol{\gamma}_c \leq \boldsymbol{\gamma}_{max} \\ & \mathbf{c}_S^T \, \boldsymbol{\omega}_c \geq c_s \end{aligned} \tag{4}$$

The desired evolution $\dot{\mathbf{e}}_d$ chosen for the visual error $\mathbf{e}$ is evaluated at each iteration and described in the next section. Furthermore, $\boldsymbol{\gamma}_c$ denotes the camera acceleration while $\boldsymbol{\omega}_{max}$ and $\boldsymbol{\gamma}_{max}$ are respectively the velocity and the acceleration bounds. In order to express linearly the acceleration constraint with respect to $\boldsymbol{\omega}_c$, it is implemented by discretizing $\boldsymbol{\gamma}_c = (\boldsymbol{\omega}_c(k+1) - \boldsymbol{\omega}_c(k))/\Delta t$, $\boldsymbol{\omega}_c(k)$ being the result obtained at the previous step. Finally, $\mathbf{c}_S$ and c_s are respectively the sharpness constraint vector and a scalar involved for ensuring the motion blur-free constraint. The complete definition of the couple ($\mathbf{c}_S$, c_s), together with a stability and performance analysis of the system, is provided in Section IV.

B. Visual error model

In the nominal case of a static pan-tilt unit observing a motionless object, the classical model is an exponential decay of the error vector:

$$\dot{\mathbf{e}}_d = -\lambda \mathbf{e} \tag{5}$$

with $\lambda > 0$ an adaptive gain defined in [30]. This case will be studied as a proof of concept, as it is widely used in the visual servoing community.

In case the camera is subject to known external motions, which is our use-case since the satellite has its own orbital speed, the visual error model is refined for compensating these external motions. More precisely, for a satellite observing a static terrestrial target, we set [30]:

$$\dot{\mathbf{e}}_d = -\lambda \mathbf{e} - \mathbf{L}_{\boldsymbol{v}} {}^{c}\boldsymbol{v}_{ext} \tag{6}$$

where ${}^{c}\boldsymbol{v}_{ext} = {}^{c}\boldsymbol{v}_s - {}^{c}\boldsymbol{v}_{o_E}$ is the known external velocity applied to the camera, with ${}^{c}\boldsymbol{v}_s$ and ${}^{c}\boldsymbol{v}_{o_E}$ respectively the translational velocity of the satellite and the translational velocity induced by Earth's rotation, both expressed in the camera frame. Then, $\mathbf{L}_{\boldsymbol{v}}$ is the translational part of the interaction matrix of $\mathbf{x}$, given by [6]:

$$\mathbf{L}_{\boldsymbol{v}} = \begin{pmatrix} -1/Z & 0 & x/Z \\ 0 & -1/Z & y/Z \end{pmatrix} \tag{7}$$

with Z the depth of the target, which is obtained knowing the orbital trajectory.

Finally, in case the satellite observes a mobile target without any knowledge about its velocity, it is possible to estimate the effect in the image $\frac{\partial \mathbf{x}_t}{\partial t}$ due to this unknown motion and to compensate for it [6], leading to:

$$\dot{\mathbf{e}}_d = -\lambda \mathbf{e} - \mathbf{L}_{\boldsymbol{v}} {}^{c}\boldsymbol{v}_{ext} - \widehat{\frac{\partial \mathbf{x}_t}{\partial t}} \tag{8}$$

with $\widehat{\frac{\partial \mathbf{x}_t}{\partial t}}$ the estimation of $\frac{\partial \mathbf{x}_t}{\partial t}$.

III. MOTION BLUR METRIC

In this section, the norm of the image gradient is evaluated to measure the motion blur, and its Jacobian that will be used in the QP-controller is derived. In our application, since the camera is able to compensate for the translational velocity of the satellite and the velocity involved by Earth's rotation, motion blur occurs in two configurations: as soon as the centering from $\mathbf{x}$ towards $\mathbf{x}^*$ starts, in which case the initial acquired image is sharp, and when the camera tracks a mobile target whose own motion is unknown, in which case the part of the image corresponding to the target is initially blurred.

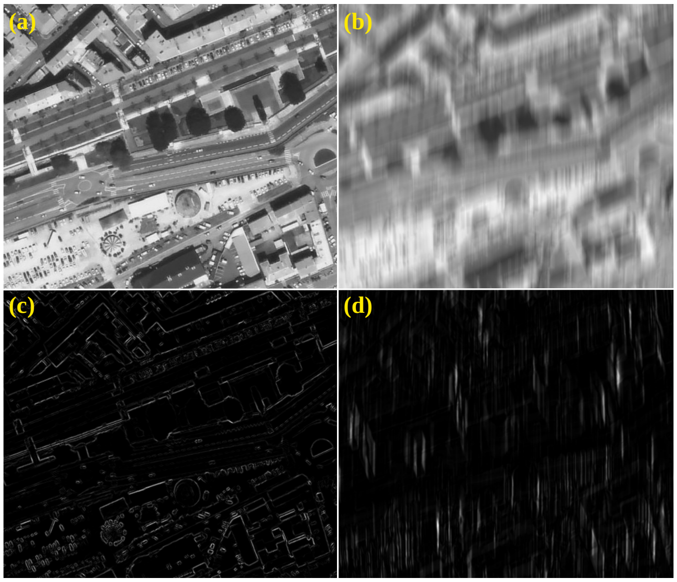

Fig 2. Norm of the image gradient applied to a satellite image of Brest (a) Sharp (b) Vertical blur (c)(d) Resulting norm of the gradient.

A. Norm of the image gradient

As already stated in the introduction, the norm of the image gradient is a good indicator of the image sharpness. This norm is expressed as:

$$S = \frac{1}{N} \sum_{\mathbf{u} \in \text{ROI}} \nabla_u^2 I(u,v) + \nabla_v^2 I(u,v) \tag{9}$$

with N the number of pixels $\mathbf{u} = (u,v)$ in a region of interest (ROI) centered around $\mathbf{x}$, and $\nabla_k I$ the spatial image gradient along axis k. Furthermore, motion blur can be modeled as a convolution of a sharp image with a particular kernel h. More precisely, if $I^*(u,v)$ denotes the intensity of the sharp image at pixel (u,v) and $\circledast$ the convolution operator, the intensity of the blurred image is given by:

$$I(u,v) = I^*(u,v) \circledast h \tag{10}$$

with [35]:

$$h(a,b) = \begin{cases} 1/L & \text{if} \sqrt{a^2+b^2} \leq \frac{L}{2} \text{ and } \frac{b}{a} = -\tan\alpha \\ 0 & \text{else} \end{cases} \tag{11}$$

where the motion amplitude $L = t_k ||\dot{\mathbf{x}}_p||$ is the product of the integration time t_k of the camera with the norm of the projected speed $\dot{\mathbf{x}}_p = (\dot{x}_p, \dot{y}_p)$ in the image, expressed in pixel/s, and $\alpha = \arctan(-\dot{y}_p/\dot{x}_p)$ is the direction of motion. Thanks to this kernel, the sharpness function S has been evaluated on simulated images with different camera velocities (see Fig. 2). As can be seen on Fig. 3, S is maximum when $\dot{\mathbf{x}}_p \approx 0$, which means the image is indeed sharp. We also see local maxima depending on the direction of motion, which are stronger when considering large values of motion amplitude.

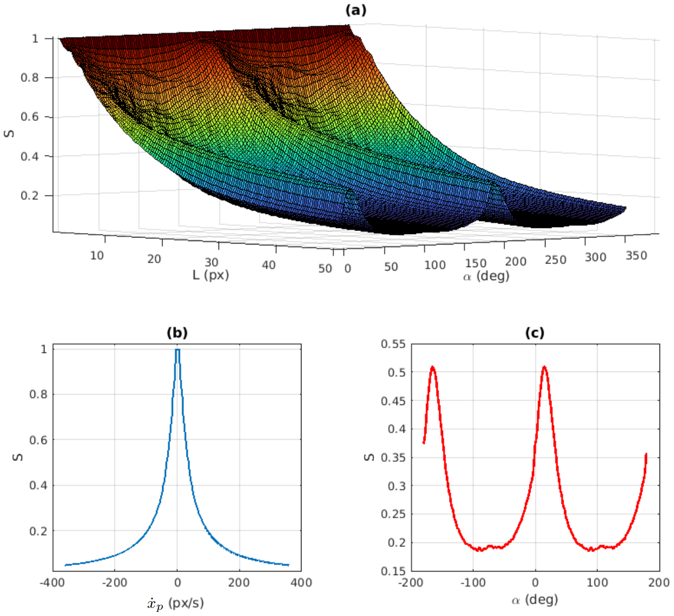

Fig 3. Norm S of the image gradient: (a) S with respect to L and α, (b) S with respect to $\dot{\mathbf{x}}_p = (\dot{x}_p, 0)$px/s, (c) S with respect to α with $\dot{\mathbf{x}}_p = (100, 0)$px/s

B. Jacobian of the sharpness function

As will be detailed in the next section, it is necessary to relate the sharpness function S to the camera pan-tilt velocities for using it in our QP controller. As a first step for that, we determine in this section the Jacobian of S with respect to the parameters L and α involved in kernel h, as well as the variation of these parameters with respect to the acceleration $\ddot{\mathbf{x}}_p = (\ddot{x}_p, \ddot{y}_p)$ in the image. We have:

$$\dot{S} = \frac{\partial S}{\partial L}\dot{L} + \frac{\partial S}{\partial \alpha}\dot{\alpha} \tag{12}$$

1) Motion amplitude: From (9), the variation of S related to L, $J_L = \frac{\partial S}{\partial L}$, is given by:

$$J_L = \frac{1}{N}\sum_{\mathbf{u}\in\text{ROI}} 2(\nabla_u I \frac{\partial \nabla_u I}{\partial L} + \nabla_v I \frac{\partial \nabla_v I}{\partial L}) \tag{13}$$

Then, from (10), we have:

$$\begin{aligned} \frac{\partial \nabla_u I(u,v)}{\partial L} &= \nabla_u (I^*(u,v) \circledast \frac{\partial h}{\partial L}) \\ \frac{\partial \nabla_v I}{\partial (} u,v)L &= \nabla_v (I^*(u,v) \circledast \frac{\partial h}{\partial L}) \end{aligned} \tag{14}$$

Finally, from (11), we obtain:

$$\frac{\partial h}{\partial L} = \begin{cases} -1/L^2 & \text{if} \sqrt{a^2+b^2} \leq \frac{L}{2} \text{ and } \frac{b}{a} = -\tan\alpha \\ 0 & \text{else} \end{cases} \tag{15}$$

which can be rewritten under the simple form:

$$\frac{\partial h}{\partial L} = -\frac{1}{L}h \tag{16}$$

leading to:

$$J_L = -\frac{2}{LN}\sum_{\mathbf{u}\in\text{ROI}} \nabla_u^2 I(u,v) + \nabla_v^2 I(u,v) = -2S/L \tag{17}$$

As for the variation of L, from its definition $L = t_k ||\dot{\mathbf{x}}_p||$, we have:

$$\dot{L} = t_k \frac{d||\dot{\mathbf{x}}_p||}{dt} = t_k \frac{\dot{x}_p \ddot{x}_p + \dot{y}_p \ddot{y}_p}{||\dot{\mathbf{x}}_p||} \tag{18}$$

2) Direction of motion: Similarly, the variation of S related to α, $J_\alpha = \frac{\partial S}{\partial \alpha}$, has the same form as (13):

$$J_\alpha = \frac{1}{N}\sum_{\mathbf{u}\in\text{ROI}} 2(\nabla_u I \frac{\partial \nabla_u I}{\partial \alpha} + \nabla_v I \frac{\partial \nabla_v I}{\partial \alpha}) \tag{19}$$

From (10), we now have:

$$\begin{aligned} \frac{\partial \nabla_u I(u,v)}{\partial \alpha} &= \nabla_u (I^*(u,v) \circledast \frac{\partial h}{\partial \alpha}) \\ \frac{\partial \nabla_v I(u,v)}{\partial \alpha} &= \nabla_v (I^*(u,v) \circledast \frac{\partial h}{\partial \alpha}) \end{aligned} \tag{20}$$

Unfortunately, the α-derivative of h can not be expressed directly since α appears in an edge condition of h. Thus, we propose to approximate h by a Gaussian function that is continuous wrt. α. Let us call this function h_α defined by:

$$h_\alpha(a,b) = \begin{cases} \frac{1}{L}e^{-\pi(a\sin\alpha + b\cos\alpha)^2} & \text{if} \sqrt{a^2+b^2} \leq \frac{L}{2} \\ 0 & \text{else} \end{cases} \tag{21}$$

leading to:

$$\frac{\partial h_\alpha}{\partial \alpha} = -2\pi(a\cos\alpha - b\sin\alpha)(a\sin\alpha + b\cos\alpha)h_\alpha \tag{22}$$

As for the variation of α, we obtain from its definition:

$$\dot{\alpha} = \frac{\dot{y}_p \ddot{x}_p - \dot{x}_p \ddot{y}_p}{||\dot{\mathbf{x}}_p||^2} \tag{23}$$

3) Complete Jacobian: By combining the previous developments, we obtain:

$$\dot{S} = t_k J_L \frac{\dot{x}_p \ddot{x}_p + \dot{y}_p \ddot{y}_p}{||\dot{\mathbf{x}}_p||} + J_\alpha \frac{\dot{y}_p \ddot{x}_p - \dot{x}_p \ddot{y}_p}{||\dot{\mathbf{x}}_p||^2} \tag{24}$$

which, since $L = t_k ||\dot{\mathbf{x}}_p||$, can be rewritten as:

$$\dot{S} = t_k \mathbf{J}_S \ddot{\mathbf{x}}_p \tag{25}$$

where the Jacobian of S wrt. $\dot{\mathbf{x}}_p$ is given by:

$$\mathbf{J}_S = \frac{\dot{\mathbf{x}}_p^T}{||\dot{\mathbf{x}}_p||} \begin{pmatrix} J_L & -J_\alpha/L \\ J_\alpha/L & J_L \end{pmatrix} \tag{26}$$

On one hand, we see from (17) that J_L can be computed directly from S and L. On the other hand, from (20), computing J_α necessitates to know I^* that is not available in case of target tracking. Furthermore, from (22), it is clear that supplementary processing are necessary to evaluate J_α. Finally, from the results depicted on Fig. 3, there is a clear preponderance in the variations of S due to L wrt. α. That is why we consider only the parameter L in the following, which is easily implemented by setting $J_\alpha = 0$ in (26).

IV. SHARPNESS CONSTRAINT

In this part, we aim to express a constraint expressed linearly wrt. the camera pan-tilt velocities for limiting as much as possible the motion blur and ensuring sharpness.

A. Sharpness as an angular velocity constraint

A basic idea to achieve this goal would be to state that S should not decrease, leading to the inequality constraint $\dot{S} \geq 0$. Ensuring this condition would force the control to improve the sharpness, or at least to maintain it steady. However, if the image is initially sharp, camera pan-tilt motions for centering the target will necessary decrease the sharpness, which is not compatible with the above constraint. A better strategy is to impose S to be higher than an acceptable value for maintaining a certain level of sharpness in the image. This can be expressed through the following condition:

$$S \geq \tau\, S_{max} \tag{27}$$

where $\tau \in [0, 1)$ is the level of desired sharpness (see Fig. 4 for some examples) and S_{max} is the norm of the image gradients for the sharpest image available I_{max}. In case the camera observes a static target, the initial ROI is sharp, from which $I_{max} = I^*$ is directly obtained, and so is S_{max}. When tracking a mobile target, I_{max} and S_{max} are again first selected and computed from the initial ROI, but since the initial ROI is blurred and the control is aimed to decrease this blur, it is possible to update both I_{max} and S_{max} such that:

$$\text{if } S(t) \geq S_{max} \text{ then } I_{max} = I(t) \text{ and } S_{max} = S(t) \tag{28}$$

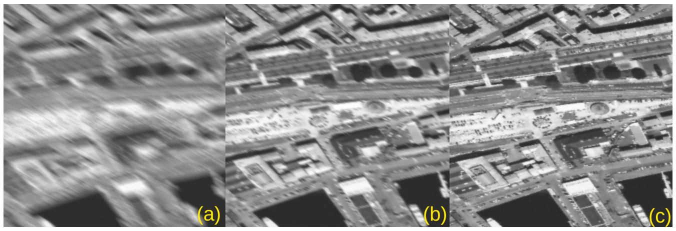

Fig 4. Normalized sharpness S/S^*: (a) 0.2, (b) 0.6, (c) 0.9.

In order to express this constraint linearly wrt. to camera pan-tilt velocities, we propose to use a control barrier function [3] from (27), defined by:

$$\dot{S} \geq \sigma \text{ with } \sigma = \kappa(\tau S_{max} - S) \tag{29}$$

where $\kappa > 0$ is a scalar gain. We now have to express (29) under the form given in (4), i.e.:

$$\mathbf{c}_S^T \boldsymbol{\omega}_c \geq c_s \tag{30}$$

For that, by considering a forward Euler discretization of (25), taking arbitrarily $\Delta t = t_k$, we obtain:

$$\dot{S} = \mathbf{J}_S(k)(\dot{\mathbf{x}}_p(k+1) - \dot{\mathbf{x}}_p(k)) \tag{31}$$

The projected speed in the image $\dot{\mathbf{x}}_p$ (expressed in pixel/s) is then given in the general case by:

$$\dot{\mathbf{x}}_p = p\,(\mathbf{L}_{\boldsymbol{\omega}}\boldsymbol{\omega}_c + \mathbf{L}_{\boldsymbol{\upsilon}}\boldsymbol{\upsilon}_{ext} + \widehat{\frac{\partial \mathbf{x}_t}{\partial t}}) \tag{32}$$

where p is the ratio between the focal length and the size of a pixel for converting meters into pixels. We can now substitute $\dot{\mathbf{x}}_p(k+1)$ in (31) to obtain:

$$\begin{aligned}\dot{S} = \mathbf{J}_S(k)\big[&p\mathbf{L}_{\boldsymbol{\omega}}(k)\boldsymbol{\omega}_c(k+1)\\ &+ p\mathbf{L}_{\boldsymbol{\upsilon}}(k)\boldsymbol{\upsilon}_{ext}(k+1) + p\widehat{\frac{\partial \mathbf{x}_t}{\partial t}}(k+1) - \dot{\mathbf{x}}_p(k)\big]\end{aligned} \tag{33}$$

Note that $\mathbf{L}_{\boldsymbol{\omega}}$ and $\mathbf{L}_{\boldsymbol{\upsilon}}$ are evaluated at instant k since their value at instant $k+1$ is not available yet. The sharpness constraint (29) can thus be written as (30) with:

$$\mathbf{c}_S^T = p\,\mathbf{J}_S(k)\mathbf{L}_{\boldsymbol{\omega}}(k) \tag{34}$$

$$\begin{aligned}c_s = \sigma - &\mathbf{J}_S(k)\big[p\mathbf{L}_{\boldsymbol{\upsilon}}(k)\boldsymbol{\upsilon}_{ext}(k+1) + p\widehat{\frac{\partial \mathbf{x}_t}{\partial t}}(k+1)\big]\\ &+ \mathbf{J}_S(k)\dot{\mathbf{x}}_p(k)\end{aligned} \tag{35}$$

This inequality constraint is linear through $\boldsymbol{\omega}_c(k+1)$ and can be handled by our QP scheme. Note that if the target to track is static ($\widehat{\frac{\partial \mathbf{x}_t}{\partial t}} = \mathbf{0}$), and/or if there are no external motions applied to the camera (${}^c\boldsymbol{\upsilon}_{ext} = \mathbf{0}$), $\dot{\mathbf{x}}_p$ in (32) is simplified, and so is c_s in (35).

B. Predictive sharpness constraint

The main issue with (35) is that the constraint is based on current measurements. It means that if the image is sharp at instant k, the constraint is not activated for computing $\boldsymbol{\omega}_c\,(k+1)$, so if blur is created at instant $k+1$ due to $\boldsymbol{\omega}_c(k+1)$, the QP will consider it only at the next step. Therefore, we propose to consider a one-step predictive constraint designed as follows:

- Predict the next velocity based on inverse kinematics through the dynamic model $\widehat{\boldsymbol{\omega}_c}(k+1) = \mathbf{L}_{\boldsymbol{\omega}}^{-1}\dot{\mathbf{e}}_d$
- Blur the ROI I_{max} obtained in (28) using (10) according to $\widehat{\boldsymbol{\omega}_c}(k+1)$, and compute $\widehat{\mathbf{J}_S}(k+1)$,
- Fill the constraint for the QP at instant k by using:

$$\mathbf{c}_S^T = p\,\mathbf{J}_S(k+1)\mathbf{L}_{\boldsymbol{\omega}}(k) \tag{36}$$

$$\begin{aligned}c_s = \sigma - &\mathbf{J}_S(k+1)\big[p\mathbf{L}_{\boldsymbol{\upsilon}}(k)\boldsymbol{\upsilon}_{ext}(k+1) + p\widehat{\frac{\partial \mathbf{x}_t}{\partial t}}(k+1)\big]\\ &+ \mathbf{J}_S(k+1)\dot{\mathbf{x}}_p(k)\end{aligned} \tag{37}$$

C. Stability and performance analysis

In case the target speed is correctly estimated, the quadratic part of the system (4) is stable since it corresponds exactly to the classic control scheme that is globally asymptotically stable [30]. However, the sharpness and mechanical constraints may slow down the system, i.e., they may decrease the time-to-convergence of the centering task. This can lead to a failure of the system for a moving target if it does not move toward the center of the image and if its speed is too fast. In all other cases, the stability of the system is preserved. In terms of performances, ensuring the sharpness constraint induces a switching behavior:

- while $S < \tau S_{max}$, it is clear from (29) that $\sigma > 0$, which means constraint (29) $\dot{S} \geq \sigma$ forces the system to increase S until S reaches τS_{max}, which decreases the time-to-convergence of the centering task.
- if $S \geq \tau S_{max}$ then $\sigma \leq 0$, which means constraint (29) $\dot{S} \geq \sigma$ allows S to decrease with a decreasing speed till S reaches τS_{max}. This has less impact on the time-to-convergence of the centering task.

S_{max} is computed according to (28), while $\tau \in [0, 1)$ is determined by the user. In practical terms, the higher τ, the sharper image, but the centering task will converge more slowly. Conversely, a lower τ allows for a faster convergence of the centering task with a reasonable loss of sharpness. The case $\tau = 1$ is naturally excluded, because impossible to ensure when the image is initially sharp, as any camera motion induces even small motion blur while the centering task has not started yet. A study of parameter τ is presented in the next section.

V. SIMULATION AND EXPERIMENTAL RESULTS

In this section, we evaluate our system for several configurations: fixed pan-tilt tracking a static object, and pan-tilt embedded on a satellite tracking a terrestrial target, motionless or not. These use cases are validated either in simulations or with experiments on a robotic platform.

The satellite path simulator is implemented as in [30], a camera is fixed to a satellite that moves in a sun-synchronous circular orbit of 500 km and acquires generated images with a 50-cm resolution (GSD) of Brest and an airplane above water moving at 800 km/h, set in their real location on the Earth's surface. The velocity and acceleration bounds of the satellite are $\boldsymbol{\omega}_{max} = (3, 3)$ deg/s and $\boldsymbol{\gamma}_{max} = (0.6, 0.6)$ deg/s^2, which corresponds to dynamic constraints provided by Airbus Defence & Space.

The experimental setup involves a Cartesian robot whose end-effector is an Intel D405 camera mounted on a pan-tilt head. The robot operates a flyby of an aerial view of Rennes with a downscaled satellite motion, while the pan-tilt head is servoed by our controller. The mechanical constraints are adjusted regarding the depth of the scene, i.e., multiplied by a factor 1500. Finally, the robot armature is covered by curtains, and a light source illuminates uniformly the scene to recreate actual optical satellite conditions. The setup is shown in the accompanying video.

For both simulations and experiments, an SSD inverse compositional template tracker [23], [4] is considered to track the ROI, whose centroid forms the target point $\mathbf{x}$. The camera frequency F has been set to 5 Hz, which corresponds to the frequency of a typical space sensor, and we have selected $t_k = 1/F$ for the integration time of an image. This high value usually generates large motion blur. To solve the QP, we use ProxQP [5] a QP-solver with augmented Lagrangian.

In the following, we highlight the influences on the control of the sharpness constraint defined in (29) tuned with different desired level of sharpness using $\tau \in [0, 1)$ while κ is set to the highest value that does not induce any oscillatory behavior, typically $\kappa = 1$. We compare the constraint outcomes on the visual error, the angular velocity, the sharpness metric, and the image trajectory. For sake of visibility, the sharpness metric (9) is shown normalized on every figure, i.e., divided by the highest value S^* of S recorded during the servo to get $S/S^* \in [0, 1]$. The accompanying video also illustrates the obtained results.

A. Translation-less pan-tilt tracking a static target

In the first configuration, we consider a nominal case in simulation where the pan-tilt unit is fixed (i.e., no external motions) and has to center a static object. This classic case for a standard pan-tilt camera allows us to analyze clearly the impact of the non-predictive and the predictive sharpness constraints and to conclude on the best one to use.

1) Non-predictive constraint: In the non-predictive case, depictedin Fig. 5 with $\tau = 0.7$ $(*)$, the sharpness drops as soon as the centering task is activated (Fig. 5c). Indeed, since the image is initially sharp, the non-predictive constraint is satisfied so the centering task is initially not constrained and generates motion blur. Just after this drop, for $\tau = 0$ that corresponds to not considering the sharpness constraint in the QP, we can note that the sharpness naturally increases due to the decrease of the angular velocity. Whereas, for $\tau = 0.7$ $(*)$, we are in the case $S < \tau S_{max}$ and the constraint is involved for increasing as fast as possible the sharpness over τS_{max}. The velocity is slow downed (Fig. 5b) and the visual error converges more slowly (Fig. 5a). Finally, the trajectory of the target in the image is a pure straight line (Fig. 5d), and the centering task is successfully completed.

2) Predictive constraint: The case considering the predictive constraint is also reported in Fig. 5. Unlike the previous case, since this constraint is based on a one-step prediction, the centering task is constrained as soon as it is activated, which allows avoiding any initial large drop (apart of course when $\tau = 0$). We are now in the case $S \geq \tau S_{max}$ and the control decreases the sharpness up to τS_{max} (Fig. 5c), which preserves much more the quality of the image, especially for $\tau = 0.9$. The angular velocity is now saturated at a value that is lower with the increase of τ (Fig. 5b), which induces a slower convergence of the visual error (Fig. 5a) with a linear decrease. The image trajectory of the target is still a pure straight line (Fig. 5d), with a perfect centering of the target.

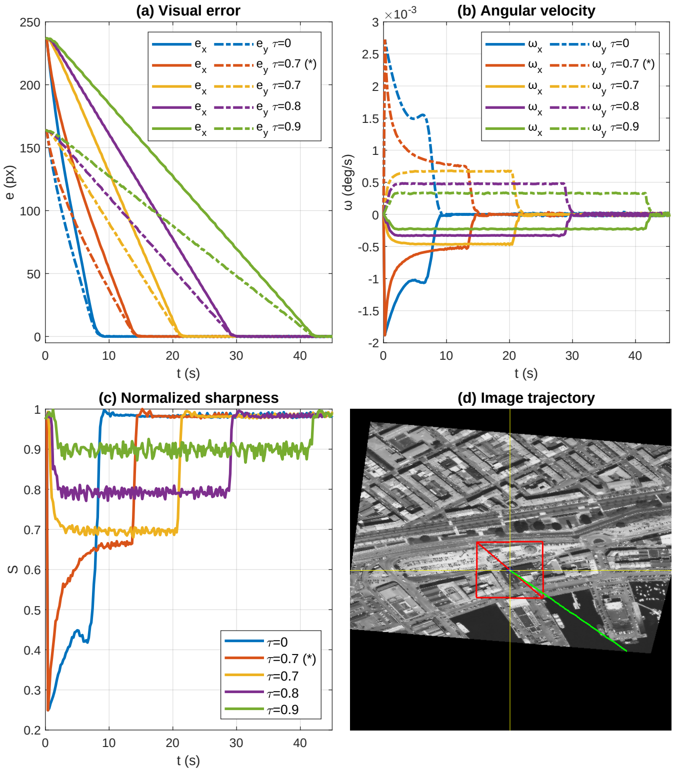

Fig 5. Simulation results of the non-predictive (35) (*) and predictive (37) constraints applied to the fixed pan-tilt configuration tracking a location in the port of Brest, wrt. τ. $S^* = 393.2$.

To conclude on this nominal case, both sharpness constraints improve the quality of the image compared to the unconstrained case while the target is perfectly centered. Still, the comparison between the non-predictive and the predictive constraints shows that the predictive constraint brings a far better behavior of the system. That is why only the predictive constraint will be considered in the following.

B. Pan-tilt embedded on a satellite tracking a fixed target

In the second configuration, the pan-tilt unit is embedded on a satellite aiming to center a static selected terrestrial target. The pan-tilt is initially compensating for all the orbital motions thanks to the satellite inner loop control, i.e., the pan-tilt focuses on a particular terrestrial location such that the target is visible in the image before the centering task is activated. This use case is validated in real experiments using our Cartesian robot,and presented in Fig. 6 and Fig. 7.

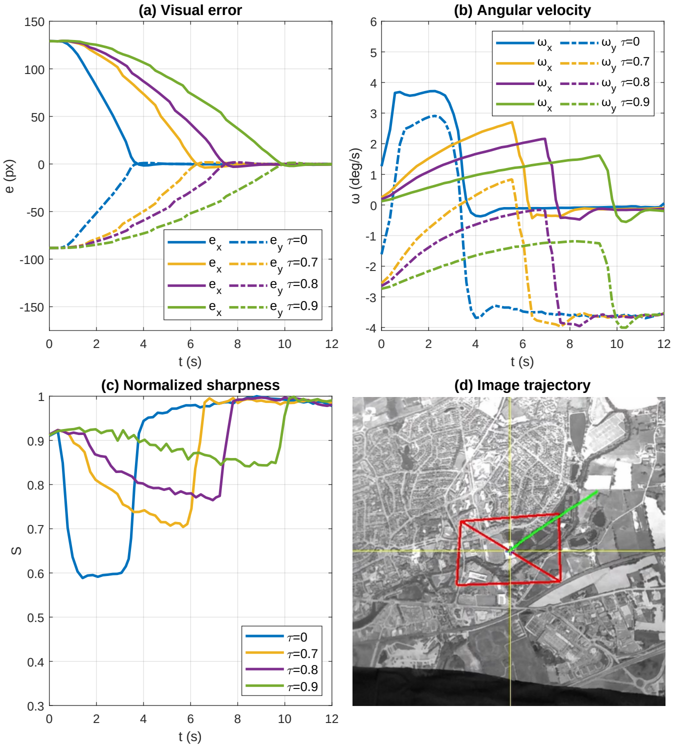

Fig 6. Experimental results of the predictive constraint (37) applied to the embedded pan-tilt configuration tracking a static terrestrial object wrt. τ. $S^* = 393.05$.

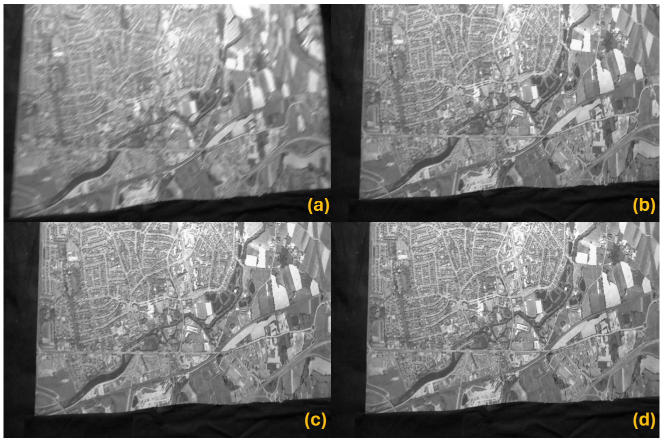

Fig 7. Comparisons between images acquired during experiments of Fig. 6 at t=2s in (a) the non-constrained case and (b) with the predictive constraint $\tau = 0.7$, (c) $\tau = 0.8$ and (d) $\tau = 0.9$.

The initial ROI is sharp as in the previous case. Thus, when $\tau \neq 0$, the sharpness is preserved over τS_{max} (Fig. 6c) with a better image quality (Fig. 7) while the target is successfully centered with a perfect straight-line trajectory (Fig. 6d) and a convergence time increasing with τ (Fig. 6a). Velocity reductions are observed compared to the unconstrained case and they are stronger wrt. τ but they still allow achieving the centering task while compensating for the satellite/robot motion. Additionally, compared to the previous simulation results, maximum sharpness is slightly varying (Fig. 6c) due to small luminosity changes. They do not impact the control scheme.

C. *Target with an unknown motion*

In this specific scenario validated in simulation, the target now exhibits significant own motion (800 km/h), causing initial blurring of the ROI, as illustrated in Fig. 8d. This blur in the early iterations may potentially lead to template tracking failures. In this case, one may consider a more robust visual tracking algorithm. The estimation of the corresponding motion $\widehat{\frac{\partial \mathbf{x}_t}{\partial t}}$ in the image is handled by a Kalman filter with a constant velocity model. As a sharp image is initially not available, S_{max} is updated during the control according to (28). As a result, the sharpness threshold τS_{max} has initially a low value that will be improved by the natural increase of the sharpness due to the compensation of the target's own motion and the reduction of the visual error. In this case, we set $\tau = 0.9$, the best level of sharpness obtained so far. The results are depicted in Fig. 8. First, the predictive constraint allows tracking and centering the airplane, while the non-consideration of the sharpness constraint leads to an immediate failure of the template tracking due to blur accumulation. Indeed, the predictive constraint favors an increase in sharpness (Fig. 8c) induced by a larger influence of the compensation of the airplane's motion (Fig. 8b) compared to the centering task at the beginning (Fig. 8b inner window). This effect is also observed in Fig. 8d, the centering task starts to bend the image trajectory only from $2s$ and its impact becomes preponderant at $4s$ when the aircraft's motion is fully compensated. Thus, as

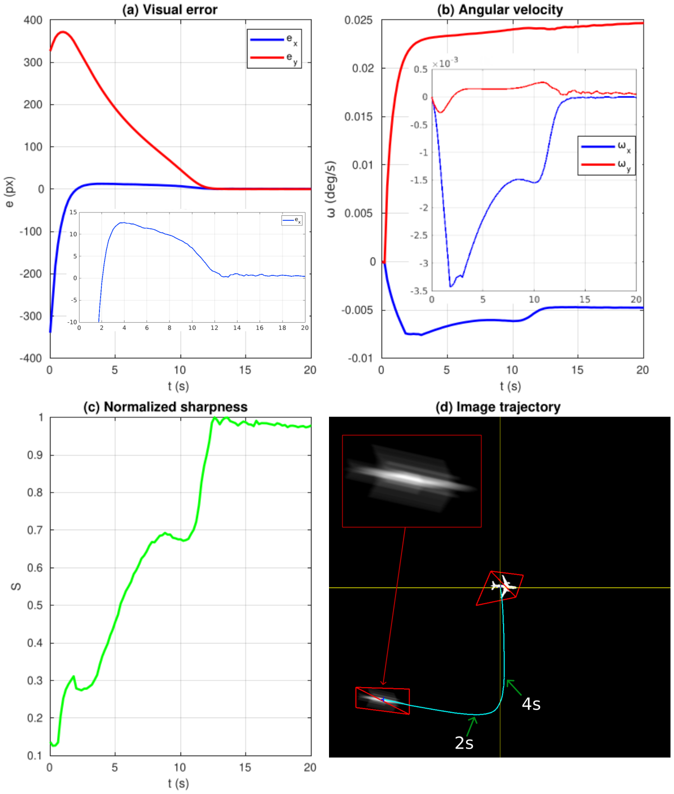

Fig 8. Simulation results of the predictive constraint (37) applied to the embedded pan-tilt configuration tracking an airplane flying at 800 km/h above water. $S^* = 523.2$. Velocity in (b) are subtracted from orbital motion compensations, the inner figure also subtracts the compensation for the aiplane motion.

already explained in Section IV-C, any target motion gradually delays the completion of the centering task (a part, of course, if the target moves naturally to the center of the image frame). Lately, the centering is perfectly accomplished (Fig. 8a) and the maximum sharpness is reached accordingly (Fig. 8c).

This last scenario illustrates that the predictive constraint enables the tracking and centering of any high-speed target in Earth troposphere (validated up to 800 km/h) while maintaining an expected degree of sharpness as long as the image processing part is able to track it.

VI. CONCLUSION

In this paper, a QP-based IBVS controller has been expressed under a motion blur-free constraint to operate simultaneously a centering and target tracking task in various configurations of a pan-tilt camera. This constraint was designed through a control barrier function using the norm of the image gradient and proposed in two variations, either with current image information or after a one-step prediction. After demonstrating the validity of our approach in a nominal case, we considered the pan-tilt embedded on a satellite, where the tracking of a target, static or not, was addressed. For each case, the control successfully completed the centering task while ensuring the sharpness constraint, leading to preserve the quality of the image during the whole process. Specifically, the predictive constraint made possible the smooth target tracking of a high-speed object, which has a lot of interests either in the robotics or in the astronautical fields. Finally, we believe that this approach can be easily transferred to any system using a camera, provided that the visual control law can be formulated as a quadratic cost function. As future work, it would be interesting to integrate an optimization of the energy consumption in the QP. An MPC could also be designed for applications benefiting from high computational resources.

REFERENCES

[1] D.J. Agravante, G. Claudio, F. Spindler, and F. Chaumette. Visual servoing in an optimization framework for the whole-body control of humanoid robots. *IEEE Robotics and Autom. Lett.*, 2(2):608–615, 2016.

[2] G. Allibert, E. Courtial, and F. Chaumette. Predictive control for constrained image-based visual servoing. *IEEE Trans. on Robotics*, 26(5):933–939, 2010.

[3] A.D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada. Control barrier functions: Theory and applications. In *European control conference*, pages 3420–3431, 2019.

[4] S. Baker and I. Matthews. Lucas-Kanade 20 years on: A unifying framework. *Int. Journal of Computer vision*, 56:221–255, 2004.

[5] A. Bambade, S. El-Kazdadi, A. Taylor, and J. Carpentier. Prox-QP: Yet another quadratic programming solver for robotics and beyond. In *Robotics: Science and Systems*, 2022.

[6] F. Chaumette and S. Hutchinson. Visual servoing and visual tracking. In *Handbook of Robotics*, chapter 24, pages 563–583. Springer, 2008.

[7] S. Cho and S. Lee. Fast motion deblurring. In *SIGGRAPH Asia*, 2009.

[8] S. Cho, Y. Matsushita, and S Lee. Removing non-uniform motion blur from images. In *IEEE Int. Conf. on Computer Vision*, 2007.

[9] L. Cui, E. Marchand, S. Haliyo, and S. Régnier. Hybrid automatic visual servoing scheme using defocus information for 6-dof micropositioning. In *IEEE Int. Conf. on Robotics and Automation*, 2015.

[10] S. Dai and Y. Wu. Motion from blur. In *IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.

[11] A. Dame and E. Marchand. Mutual information-based visual servoing. *IEEE Trans. on Robotics*, 27(5):958–969, Oct. 2011.

[12] J. Dauner, L. Elsner, O. Ruf, D. Borrmann, J. Scharnagl, and K. Schilling. Visual servoing for coordinated precise attitude control in the tom small satellite formation. *Acta Astronautica*, 2023.

[13] D. Falanga, P. Foehn, P. Lu, and D. Scaramuzza. PAMPC: Perception-aware model predictive control for quadrotors. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2018.

[14] A. Hajiloo, M. Keshmiri, W.F. Xie, and T.T. Wang. Robust online model predictive control for a constrained image-based visual servoing. *IEEE Trans. on Industrial Electronics*, 63(4):2242–2250, 2015.

[15] H. Ji and C. Liu. Motion blur identification from image gradients. In *IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.

[16] O. Kanoun, F. Lamiraux, and P.B. Wieber. Kinematic control of redundant manipulators: Generalizing the task-priority framework to inequality task. *IEEE Trans. on Robotics*, 27(4):785–792, 2011.

[17] O. Kermorgant and F. Chaumette. Dealing with constraints in sensor-based robot control. *IEEE Trans. on Robotics*, 30(1):244–257, 2013.

[18] G. Klančar, S. Blažič, D. Matko, and G. Mušič. Image-based attitude control of a remote sensing satellite. *Journal of intelligent & robotic systems*, 66:343–357, 2012.

[19] E. Krotkov. Focusing. *Int. J. of Computer Vision*, 1(3):223–237, 1988.

[20] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 8183–8192, 2018.

[21] E. Marchand. Control camera and light source positions using image gradient information. In *IEEE Int. Conf. on Robotics and Automation*, pages 417–422, Roma, Italia, April 2007.

[22] E. Marchand and C. Collewet. Using image gradient as a visual feature for visual servoing. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Taipei, Taiwan, Oct. 2010.

[23] E. Marchand, F. Spindler, and F. Chaumette. ViSP for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robotics and Automation Mag.*, 12(4):40–52, Dec. 2005.

[24] M. McGuire, P. Hennessy, M. Bukowski, and B. Osman. A reconstruction filter for plausible motion blur. In *ACM Symp. on Interactive 3D Graphics and Games*, pages 135–142, 2012.

[25] S. Nah, T. Hyun Kim, and K. Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *IEEE Conf. on Comp. Vision and Pattern Recognition*, 2017.

[26] F. J. Navarro, F. Serón and D. Gutierrez. Motion blur rendering: State of the art. *Comp. Graphics Forum*, 30(1):3–26, 2011.

[27] S.K. Nayar and M. Ben-Ezra. Motion-based motion deblurring. *IEEE Trans. on PAMI*, 26(6):689–698, 2004.

[28] PhaseOne. iXM-SP150. https://www.phaseone.com/solutions/geospatial-solutions/components/ixm-sp150/.

[29] M. Potmesil and I. Chakravarty. Modeling motion blur in computer-generated images. *ACM SIGGRAPH*, 17(3):389–399, 1983.

[30] M. Robic, R. Fraisse, É. Marchand, and F. Chaumette. Vision-based rotational control of an agile observation satellite. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2022.

[31] P. Roque, E. Bin, Pedro M., and D. Dimarogonas. Fast model predictive image-based visual servoing for quadrotors. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 7566–7572, 2020.

[32] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Trans. on Graphics*, 27(3):1–10, 2008.

[33] M. Subbarao and J.-K. Tyan. Selecting the optimal focus measure for autofocusing and depth-from-focus. *IEEE Trans. on PAMI*, 20(8):864–870, 1998.

[34] Y. Sun and B.J. Duthaler, S.and Nelson. Autofocusing algorithm selection in computer microscopy. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2005.

[35] S. Tiwari, V.P. Shukla, A.K. Singh, and S.R. Biradar. Review of motion blur estimation techniques. *Journal of Image and Graphics*, 1(4):176–184, 2013.

[36] J. Tordesillas and J.P. How. Panther: Perception-aware trajectory planner in dynamic environments. *IEEE Access*, 10:22662–22677, 2022.

[37] C.Y. Wee and R. Paramesran. Measure of image sharpness using eigenvalues. *Information Sciences*, 177(12):2533–2552, 2007.

[38] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *European Conf. on Computer Vision*, 2010.

[39] Y. Yitzhaky and N.S. Kopeika. Identification of blur parameters from motion blurred images. *Graphical Models and Image Processing*, 59(5):310–320, 1997.