2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Multi-Spectral Visual Servoing

Enrico Fiasché[1,2], Ezio Malis[2] and Philippe Martinet[2]

Abstract— **This paper presents a novel approach for Visual Servoing (VS) using a multispectral camera, where the number of data are more than three times that of a standard color camera. To meet real-time feasibility, the multispectral data captured by the camera are processed using dimensionality reduction techniques. Instead of relying on traditional approaches that select a subset of bands, the proposed method unlocks the full potential of a multispectral camera by pinpointing individual pixels that hold the richest information across all bands. While sacrificing spectral resolution for enhanced spatial resolution - crucial for precise robotic control in forested environments - this fusion process offers a powerful tool for robust and real-time VS in natural settings. Validated through simulations and real-world experiments, the proposed approach demonstrates its efficacy by leveraging the full spectral information of the camera while preserving spatial details.**

I. Introduction

The application of Visual Servoing (VS) techniques has gained significant attention for various robotic tasks [1]. VS involves controlling robot motion based on visual feedback from a camera, enabling tasks such as object tracking, manipulation, and navigation. In complex environment, such as natural ones, accurate perception is crucial and the choice of an appropriate camera system plays a vital role. Multispectral cameras, a type of imaging system, have emerged as promising tools for perception in challenging environments. Unlike traditional RGB cameras that capture images in three color channels (red, green, and blue), multispectral cameras are capable of capturing images across a wider range of wavelengths, composed from 5 to 15 bands. This provides valuable spectral information beyond the visible spectrum, offering significant advantages in natural environments where comprehensive perception is essential. However, deploying VS techniques using multispectral cameras in real-time natural scenarios poses a key challenge: processing the large amount of multispectral data while maintaining high temporal performance. Addressing these processing and interpretation challenges is crucial to effectively utilizing the spectral information provided by multispectral cameras. Dimensionality reduction techniques play a crucial role in this regard, as they enable the extraction of key information while reducing the computational complexity of the analysis [2]. In hyperspectral imaging (HSI), a type of multispectral imaging with extremely high spectral resolution, various approach such as feature extraction [3] and band selection [4] have been explored to enable accurate and efficient classification of HSI images. Feature extraction combines multiple bands to create a compressed but informative subset, though it can be computationally demanding and may hinder interpretability due to the creation of new features. Conversely, band selection techniques choose a subset of bands from the multispectral data, preserving spectral meaning and reducing computational burden [4]. Various unsupervised and supervised methods have been proposed in the literature for band selection in hyperspectral data analysis. In [5], the Multigraph Determinantal Point Process (MDPP) approach efficiently finds the optimal band subset, while Principal Component Analysis (PCA) prioritizes the energy of variances in band images [6]. Another method, Sparse Nonnegative Matrix Factorization (SNFM), is used to solve the band selection problem [7]. Similarly, minimizing the correlation of selected bands has been used to identify the best subset [8]. In [9], an alternative approach involves clustering original bands and selecting representative bands from each cluster, proposed as an automatic band selection method. Furthermore, hybrid methods that combine the advantages of previous techniques are introduced [10]. In the context of Multi-Spectral Visual Servoing (MSVS), researchers have explored integrating VS techniques with RGB cameras to enhance robotic vision and control. One notable approach among these techniques is the colorimetry-based approach [11]. The authors investigated the utilization of visual features derived from the individual R, G, and B components of the image. Their objective was to evaluate the suitability of employing features derived from different linear combinations of two or all three components for visual servoing tasks. The results of their study illustrated the promising potential of this approach for enhancing VS applications. Another significant contribution comes from [12]. In this work, authors presented innovative parametric models and optimization methods for the robust and direct registration of color images. Their methodology involved representing a color image obtained by stacking n channels multiplied by corresponding surfaces to achieve precise alignment with a reference image. However, their methodology focused on the use of only the readily available RGB channels. In a different line of inquiry, researchers delved into frequency domain [13]. Rather than analyzing the image itself in the spatial domain, they explored its transformation into the frequency domain. While these approaches have yielded valuable insights, it's important to note that they are limited to RGB images and do not leverage the additional spectral information provided by multispectral cameras. The primary objective of this work is to bridge this gap and extend the applicability of visual servoing techniques to multispectral

[1]Université Côte d'Azur DS4H, Nice, France, enrico.fiasche@inria.fr
[2]Inria Centre at Université Côte d'Azur, Sophia Antipolis, France, {ezio.malis, philippe.martinet}@inria.fr

cameras with n bands, thereby harnessing the full potential of multispectral imaging in the context of robotic control. To the authors' knowledge, this is the first work investigating Visual Servoing for multispectral cameras' data. This work focuses on using the multispectral information to improve the performance of visual servoing tasks, rather than emphasizing the control aspect.

II. Background

To enhance the paper's readability, prior to delving into Multi-Spectral Visual Servoing, this section revisits the process of band selection as outlined in prior research. Additionally, it introduces the notation used throughout the paper and provides a brief overview of the foundational models and methodologies employed.

A. Band Selection

The band selection problem is a well-known problem in the field of remote sensing. It is defined as the process of selecting a subset of bands from a hyperspectral image that is most relevant to a given task. The goal is to reduce the dimensionality of the data while retaining the most relevant information. The band selection problem is a challenging task due to the elevated dimensionality of the data and the vast number of potential band combinations.

In a previous study [14], the authors employed the Rank Minimization-Based Selector to extract the most informative bands. This technique incorporates a spatial discrete gradient operator to reveal the spatial structural information in the band. Subsequently, it represents the spectral bands as a linear combination of a dictionary $\mathbf{D}$ and a coefficient matrix $\mathbf{Z}$. The low-rank representation [15] is applied in order to solve the problem. The authors applied a convex approximation method and introduced real-world noise considerations, leading to the followings:

$$\begin{aligned} \arg\min_{\mathbf{Z},\mathbf{E}} \quad & ||\mathbf{Z}||_* + \lambda||\mathbf{E}||_{2,1} \\ \text{s.t.} \quad & \mathbf{X} = \mathbf{D}\mathbf{Z} + \mathbf{E} \end{aligned} \tag{1}$$

where $||\cdot||_*$ is the trace norm and $||\mathbf{E}||_{2,1}$ is the $\ell_{2,1}$ norm. Then, the problem can be effectively addressed through the use of the augmented Lagrange multiplier method. Following the acquisition of the coefficient matrix $\mathbf{Z}$, the bands are ultimately categorized into K distinct groups employing spectral clustering techniques, as delineated in [16], using the affinity matrix $\mathbf{A} = (|\mathbf{Z}| + |\mathbf{Z}^\top|)$. Within a given cluster $\mathbf{C}$ featuring R bands, the weight of each band $\mathbf{x}_i \in \mathbf{C}$ is computed as follows:

$$\mathbf{W}_i = \frac{\sum_{\mathbf{x}_j \in \mathbf{C}} \mathbf{E}_i}{R\mathbf{E}_j} \sum_{\mathbf{x}_j \in \mathbf{C}, j \neq i} \mathbf{A}_{ij} \tag{2}$$

where the representation residual $\mathbf{E}$ in (1), and $\mathbf{E}_i = \sum \mathbf{E}_{ij}$. The bands ultimately chosen consist of those with the highest weights within their respective clusters. These bands are anticipated to encompass the majority of the information present in the original image cube.

B. Visual Servoing Notations

Let $||\mathbf{v}||$ and $\mathbf{v}'$ to represent, respectively, the Euclidean norm and a transformed version of the variable $\mathbf{v}$. Additionally, $\mathbf{v}^*$ is employed to indicate that $\boldsymbol{v}$ is defined with respect to the reference frame $\mathcal{F}^*$. The notation vex($\boldsymbol{\omega}$) is employed to denote $\text{vex}([\boldsymbol{w}]_\times) = [w_1; w_2; w_3]$.

C. Direct Visual Servoing

The Direct Visual Servoing (DVS) approach is a well-known technique that is widely employed in the field of robotics [13] [17]–[20]. It is a visual servoing technique that is based on the minimization of the image error. Indeed, instead of errors on features extracted from the image, the error is defined as the difference between the current image and the desired image. These methods typically aim to minimize a dissimilarity measure, such as the sum-of-squared differences (SSD), between the reference image and the current image using parametric models. While the frequency-based approach, as demonstrated in [13], has shown promise and potential for producing favorable results, a distinct approach is pursued in this paper. It is worth noting that the frequency-based approach remains an interesting avenue for further exploration, which could yield valuable insights in future research. One notable approach, as presented in prior research [17], involves the direct estimation of the projective transformation between a selected reference template and the corresponding area in the current image. Beginning with an initial prediction of the homography, the algorithm iteratively refines the estimation to find the optimal homography that minimizes the SSD between the reference pattern and the current pattern. In [18], the authors define the geometric parameters $\mathbf{g} = \{\mathbf{G}, \mathbf{e}, \rho_i^*\}_{i=1}^n$, where $\mathbf{G}$ is a projective homography relative to a dominant (virtual) plane, $\mathbf{e}$ denotes the epipole and ρ^* is the projective parallax of the 3-D point. The geometric parameters are estimated directly from pixel intensities, allowing the application of the algorithm to object of generic shape and texture. Within the framework of the DVS, the control errors for translation and rotation, referred to as $\boldsymbol{\varepsilon}_v$ and $\boldsymbol{\varepsilon}_\omega$ respectively, are as follows:

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_v \\ \boldsymbol{\varepsilon}_\omega \end{bmatrix} = \begin{bmatrix} (\mathbf{H} - \mathbf{I})\mathbf{m}^{*'} + \rho^* \mathbf{e}' \\ \theta\mu \end{bmatrix} \tag{3}$$

where $\rho^* \in \mathbb{R}$ is the parallax of the chosen control point $\mathbf{p}^*$.

$$\mathbf{H} = \mathbf{K}^{-1}\mathbf{G}\mathbf{K} \tag{4}$$

$$\mathbf{m}^{*'} = \mathbf{K}^{-1}\mathbf{p}^* \tag{5}$$

$$\mathbf{e}' = \mathbf{K}^{-1}\mathbf{e} \tag{6}$$

the matrix $\mathbf{K} \in \mathbb{R}^{3\text{x}3}$ contains the camera intrinsic parameters

$$\theta = \begin{cases} \psi, & \text{if trace}(\mathbf{H} \geq 1) \\ \pi - \psi, & \text{otherwise} \end{cases} \tag{7}$$

$$\boldsymbol{\mu} = \frac{\boldsymbol{r}}{||\boldsymbol{r}||}, \quad \boldsymbol{r} = \frac{1}{2}\text{vex}(\mathbf{H} - \mathbf{H}^\top) \tag{8}$$

where $\psi = \text{real}(\arcsin(||\boldsymbol{r}||))$. Indeed, let the control inputs be translational and rotational camera velocities, gather in $\mathbf{v} = \left[\boldsymbol{v}; \boldsymbol{\omega}\right]$. The control law is defined by:

$$\mathbf{v} = \boldsymbol{\Gamma}\boldsymbol{\varepsilon} \tag{9}$$

with $\boldsymbol{\Gamma} = \text{diag}(\gamma_v \mathbf{I}, \gamma_\omega \mathbf{I})$. In the study conducted by [18], it was analyzed and demonstrated that the control law achieves local asymptotic stabilization of the equilibrium state $\boldsymbol{\varepsilon} = \mathbf{0}$.

III. Multi-Spectral Visual Servoing

This section introduces the proposed Multi-Spectral Visual Servoing technique, a novel approach designed to enhance robotic capabilities through the integration of multispectral imaging. Leveraging a multispectral camera system capturing a diverse range of spectral information across n fixed bands, the objective in this section is to leverage multispectral information to empower robots for advanced perception, tracking, and control in diverse real-world and natural scenarios. Multispectral imaging, with its rich spectral data, poses challenges due to high dimensionality. While adapting existing Direct Visual Servoing methods, the research work paves the way for a paradigm shift in multispectral-based robotic control. This novel integration unlocks the potential of multispectral information within VS, enabling informed robotic decision-making and ultimately leading to more versatile and robust robotic systems capable of handling diverse tasks in challenging environments.

A. Pixel Selection

The proposed approach to dimensionality reduction of multispectral data fundamentally differs from the conventional state of the art in a pivotal way: this approach do not exclude any bands from consideration. The key distinction here lies in the focus on spatial information. It is acknowledged that each band within a given image may contain distinctive spatial details, and a dedication to comprehensive evaluation ensures that no potentially relevant information is overlooked prematurely.

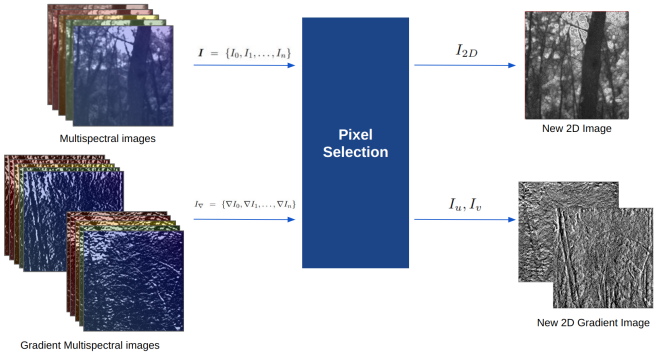


Fig. 1: Pixel Selection schema.

As shown in Fig.1, to further optimize the utility of multispectral imagery, this approach implements a discrete gradient operation on each band. This operation, applied individually to every band, serves the paramount purpose of highlighting regions where valuable information is concentrated, enhancing the contrast and making relevant data more discernible. The proposed method conducts a pixel-by-pixel analysis across all bands. This granular examination enables us to identify, pixel-by-pixel, the band containing the highest information content. To achieve this discrimination and pinpoint the most informative bands for each pixel, it is considered the multispectral images $\boldsymbol{I} = \{I_0, I_1, \ldots, I_n\}$ composed by n bands and their corresponding discrete gradient represented as $\boldsymbol{I}_\nabla = \{\nabla I_0, \nabla I_1, \ldots, \nabla I_n\}$. The approach employs the following equation, which assesses the spatial details:

$$k = \max_{i=1:n} \{\nabla I_i[u]^2 + \nabla I_i[v]^2\} \tag{10}$$

where k is the band index, and $\nabla I_i[u]$ and $\nabla I_i[v]$ are the horizontal and vertical gradients of the i^{th} band, respectively at the pixel $[u, v]$. This process is repeated for every pixel in the image, resulting in a pixel-wise selection of the most informative band. Consequently, it can directly construct the new 2D image by taking the selected band for each pixel from the original multispectral data. This new image effectively compresses the essential data into a singular image. This operation was performed this operation on the reference image to encapsulate the most valuable spatial information for further analysis and applied the same pixel selection process to the multispectral current image. The approach effectively recreates a similar image in the multispectral domain, preserving the desired spatial details. The proposed method prioritizes spatial details over spectral resolution, achieving significant benefits in visual servoing applications where precise spatial information is paramount. While spectral resolution hold value in various applications, it assumes a secondary role in visual servoing, making this trade-off advantageous.

B. Visual Servoing Formulation

The focus shifts towards the application of Direct Visual Servoing techniques. In this specific application, feature-free visual servoing techniques are adapted and tailored based on the principles outlined in the theoretical background to suit the use of multispectral camera data. These techniques provide valuable advantages in scenarios where objects exhibit generic shapes and textures. The implementation involves adapting the approach to the multispectral camera setup and optimizing it for precise positioning tasks in diverse real-world and natural scenarios. The resulting estimated homography serves as a crucial input for the MSVS approach. The control error, $\varepsilon \in \mathbb{R}^6$ for the translation and rotation, is calculated based on the estimated parameter $\mathbf{G}$ and is defined using the equation (3). However, it is important to note that the control error calculation has been customized to align with the specific application. As suggested in [19], this work uses a reduced version of the general control error as the rotational control error is equivalent to the original one via $\boldsymbol{\varepsilon}_\omega \approx 2\boldsymbol{r} \approx 2\theta^{-1}||\boldsymbol{r}||\boldsymbol{\varepsilon}$. In this setup, let the control inputs be the translational and rotational velocities of the camera, gather in the vector $\mathbf{v} = \left[\boldsymbol{v}; \boldsymbol{\omega}\right] \in \mathbb{R}^6$, defined in (9). The stability analysis was proven by [19].

IV. Simulation and Experiments Results

A. Simulation Setup

In the pursuit of evaluating the effectiveness of the propesed Multi-Spectral Visual Servoing (MSVS) approach, this section reports a series of simulations using a multispectral camera system. In all scenarios, the control objective is to accurately position a six-degree-of-freedom (6-DoF) controlled camera so that the current object image aligns precisely with the image captured at the reference pose, referred to as the reference image. The camera system, known as TOUCAN [21], is a multispectral camera capable of capturing data across ten distinct spectral bands. To replicate real-world scenarios, four different images in natural environments are used, all of which were captured with the multispectral camera, in proximity to Inria laboratory. In Fig.2 is shown an example of the multispectral images used in the simulations. The first ten images show the original multispectral data, and the result shows the corresponding 2D images, which is the result of the pixel selection presented in the previous section.

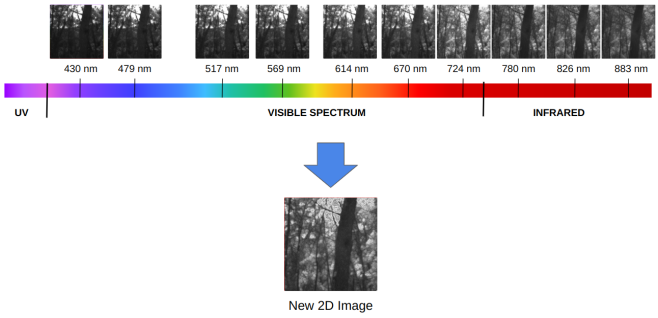
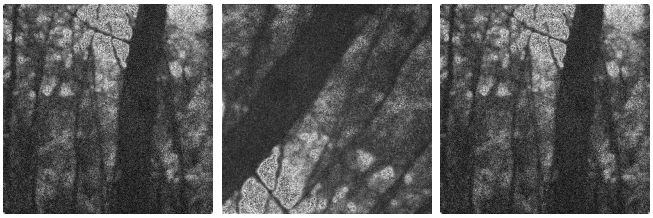


Fig. 2: From multispectral to 2D image.

In all simulations, a virtual camera system was simulated. The image processing and the control law computation are performed on a PC running Linux, equipped with a 14-cores 3.0 GHz Alder Lake. The code has been written in C using OPENROX open source library [22]. Each of these images presented unique challenges and characteristics, allowing for a comprehensive evaluation of the approach under varying conditions. Each image used in the system has dimensions of 520x520 pixels, providing ample detail for the intended assessments. To demonstrate robustness of the proposed work, a comparison is made with a traditional approach. The application of the same control law is compared in two contexts: Multi-Spectral Visual Servoing (MSVS), utilizing multispectral data, and the standard Direct Visual Servoing (DVS), which relies on RGB imagery. Both methods are evaluated based on their convergence and accuracy under the influence of introduced Gaussian noise. The detailed results of this comparison are presented in the next section.

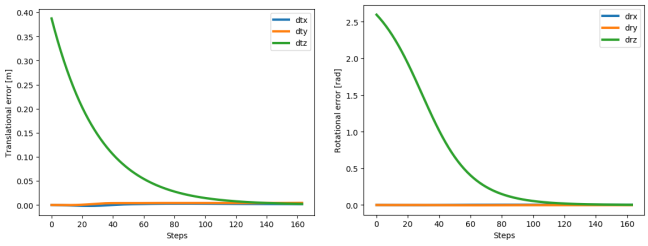
B. Simulation Results

During the simulation phase, rigorous assessments are conducted to evaluate the robustness and effectiveness of the proposed Multi-Spectral Visual Servoing approach across a range of different scenarios. As shown in [17], to simplify the simulation setup, it is assumed that the scene is planar, resulting in a value of ρ^* equal to 0 in the control error (3). In (9), a uniform value of 0.1 is chosen for the linear factor γ_v and uniform value of 0.05 is chosen for the angular factor γ_ω across all six λ_i. In addition, for the reconstruction of the current image in the virtual camera setup, a transformation matrix is applied to the reference image, which serves as the initial current image in all simulations. To show the robustness of the proposed algorithm, a 0.4 meter displacement is applied along the z-axis, coupled with a 150-degree rotation, across various images. It is then evaluated the performance of both MSVS and DVS in terms of convergence and accuracy. The performances of the algorithms are compared in four different scenarios, each influenced by Gaussian noise with a mean of 0.05 and a variance of 0.020. The Gaussian noise was added to the original multispectral image to simulate real-world environmental disturbances. This involved applying the noise directly to each band of the multispectral image. By incorporating Gaussian noise with defined parameters, the robustness of the algorithm is evaluated, providing insights into its performance in realistic conditions. In the first set of positioning simulation, shown in Fig.3, the snapshot was captured within a forest environment, with the objective to aligning the image to a prominent tree trunk. The reference image, in Fig.3a, shows the result of the pixel selection process, and the other images constitutes the outcomes of the MSVS approach.



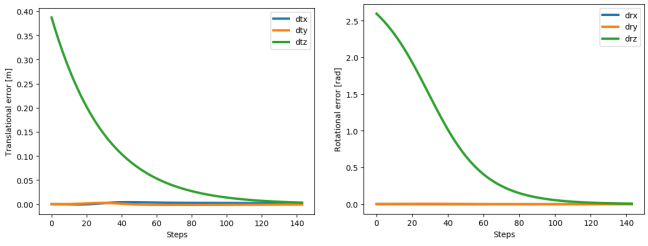
(a) Reference (b) Initial (c) Final

Fig. 3: MSVS simulation using big trunk image.



(a) Translational error MSVS

(b) Rotational error MSVS

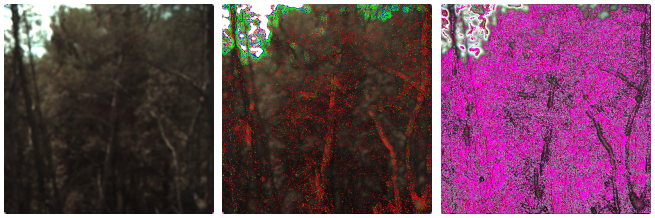


(c) Translational error DVS

(d) Rotational error DVS

Fig. 4: Convergence MSVS and DVS simulation with 0.4 m displacement and 150 degrees of rotation around the z-axis.

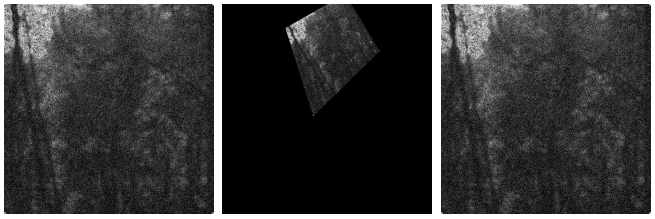
During the simulation the approaches provided similar results in terms of convergence and accuracy, as shown in the example in Fig.4. By subjecting the proposed approach to such challenges, the aim was to show MSVS ability to achieve precise object positioning under varying conditions. The control law is stable, both translational Fig.4a and rotational Fig.4b errors converge to zero. At the convergence the camera is positioned at the desired pose very accurately. Similar results were observed for other images, with both MSVS and DVS achieving consistent convergence and accuracy. However, a significant difference emerged in the fourth image. MSVS outperformed DVS, demonstrating robust convergence with clear results, while DVS diverged. This difference is attributed to the additional information provided by the multispectral images, expanding the data available for its control law to enhance navigation. Notably, both systems employed the same control law, highlighting the crucial role of image data in achieving accurate navigation. This finding underscores the potential advantages of multispectral cameras in situations where RGB cameras face limitations due to homogeneous scenes such as forests.



(a) RGB image (b) RGB pixels (c) Infrared pixels

Fig. 5: Pixel comparison between RGB and MS images.

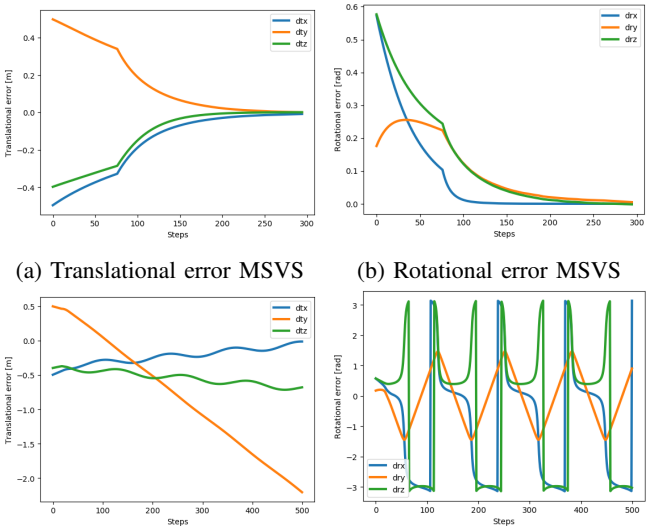
The Fig.5 shows the snapshot captured in a forest environment, capturing the vegetation. It is possible to see where the most valuable information are located in the multispectral image. The Fig.5a shows the original RGB image, while Fig.5b highlights the regions containing the most valuable Red, Green and Blue data in the multispectral image. In Fig.5c, the areas with the most valuable infrared pixels are represented. In this case, the multispectral image demonstrates more robustness as it leverages infrared bands, which contain crucial data about the vegetation. Both algorithms were evaluated with different displacements and rotations along all the axes, such as 0.5 m and 25 degrees, and the results were always the same. The reference image in Fig.6a shows the results of the pixel selection process, and the other images constitutes the outcomes of the MSVS approach.



(a) Reference (b) Initial (c) Final

Fig. 6: MSVS simulation using vegetation image.

In Fig.7, it is shown the comparison results of the MSVS and DVS approach. The MSVS control law is capable of estimating the correct homography, leading to stability, where both translational and rotational errors converge to zero. However, the DVS control law is unable to estimate the correct homography due to the noisy information coming from the RGB image, which causes the control law to diverge and become unstable. At the convergence the multispectral camera is positioned at the desired pose very accurately.



(a) Translational error MSVS (b) Rotational error MSVS

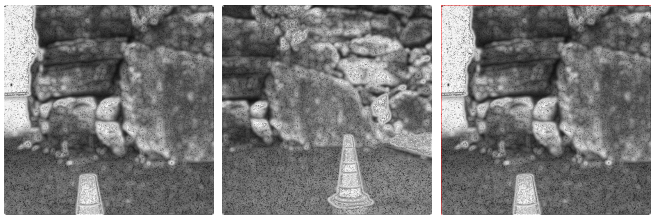
(c) Translational error DVS (d) Rotational error DVS

Fig. 7: MSVS and DVS comparison using the vegetation image under influence of Gaussian noise.

C. Real-World Experiments

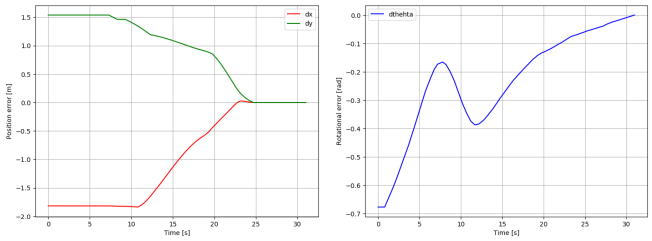
Real world experiments were conducted outdoor in the proximity of the laboratory at Inria. During the experiments, the SCOUT MINI AgileX robot, controlled as a holonomic robot, used ROS to interface with the multispectral camera system. The robot was equipped with a TOUCAN [21] multispectral camera. The robot task was to control the camera to align the current image with the reference image. The reference image was captured offline at a predefined position. The robot and the camera were controlled using the same PC system used in the simulations. An adaptive gain strategy was employed in the experiments. Initially, the robot used the gain proposed in the simulation. As the robot approached the target, the gain was gradually increased to speed up the alignment process without prolonging the waiting time to reach the final position. This adaptive gain strategy ensured efficient and timely convergence to the desired alignment, improving the overall performance of the proposed approach. In all experiments, the robot successfully aligned the current image with its corresponding location in the reference image, demonstrating the robustness of the MSVS approach, as shown in the two real-world scenarios in Fig.8 and in Fig.10. This was achieved even when starting from different positions. By employing the pixel selection technique, the multispectral data were significantly reduced, allowing for real-time processing and control. In Fig.9 and Fig.11 are presented the results of the experiments, showing

the evolution of position and orientation errors over time. Both figures demonstrate a clear convergence of errors, indicating the stability and effectiveness of the proposed approach.



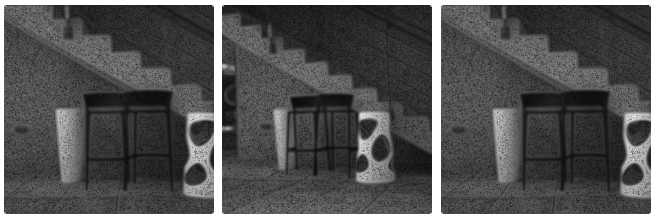
(a) Reference (b) Initial (c) Final

Fig. 8: First real-world scenario experiments.



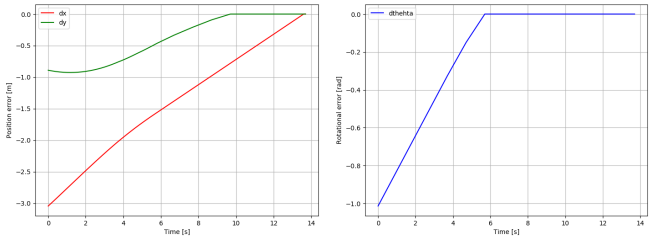
(a) Position error (b) Rotational error

Fig. 9: First real-world scenario MSVS results.



(a) Reference (b) Initial (c) Final

Fig. 10: Second real-world scenario experiments.



(a) Position error (b) Rotational error

Fig. 11: Second real-world scenario MSVS results.

V. Conclusions

In conclusion, the proposed study has demonstrated the feasibility of employing a multispectral camera system for visual servoing. The paper introduced a novel technique for transforming multispectral data into 2D images, which facilitated the application of Direct Visual Servoing. Through extensive simulations and experiments, system convergence was successfully achieved, highlighting the potential of the approach in enhancing robotic perception, tracking, and control. The integration of multispectral information, as evidenced by the methodology, opens up new possibilities for robotics in diverse real-world scenarios. While this research represents a significant step forward, it also suggests avenues for further exploration and refinement.

Acknowledgment

This work has been supported by the French government, through the UCA DS4H Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-17-EURE-0004. We would like to thank the engineer Louis Verduci for his code development for efficient multispectral camera data capture in the framework of RobForRisk.

References

[1] F. Chaumette and S. Hutchinson, "Visual servo control. i. basic approaches," *IEEE RAM*, vol. 13, no. 4, 2006.

[2] D. Ruiz Hidalgo, B. Bacca Cortés, and E. Caicedo Bravo, "Dimensionality reduction of hyperspectral images of vegetation and crops based on self-organized maps," *Information Processing in Agriculture*, vol. 8, no. 2, 2021.

[3] M. P. Uddin, M. A. Mamun, and M. A. Hossain, "Feature extraction for hyperspectral image classification," in *IEEE R10 HTC*, 2017.

[4] W. Sun and Q. Du, "Hyperspectral Band Selection: A Review," *IEEE GRS Magazine*, vol. 7, no. 2, 2019.

[5] Y. Yuan, X. Zheng, and X. Lu, "Discovering diverse subset for unsupervised hyperspectral band selection," *IEEE TIP*, vol. 26, no. 1, 2017.

[6] C.-I. Chang, Q. Du, T.-L. Sun, and M. Althouse, "A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification," *IEEE GRS*, vol. 37, 12 1999.

[7] W. Sun, W. Li, J. Li, and Y. Lai, "Band selection using sparse nonnegative matrix factorization with the thresholded earth's mover distance for hyperspectral imagery classification," *Earth Science Informatics*, vol. 8, 01 2015.

[8] Y. He, D. Liu, and S. Yi, "Recursive spectral similarity measure-based band selection for anomaly detection in hyperspectral imagery," *Journal of Optics*, vol. 13, no. 1, 2010.

[9] X. Cao, B. Wu, D. Tao, and L. Jiao, "Automatic band selection using spatial-structure information and classifier-based clustering," *IEEE JSTARS*, vol. 9, no. 9, 2016.

[10] Y. Sun and J. Pei, "Band selection based on hyperspectral piling Fisher graphs (HSPFiGs) analysis," *Infrared Physics & Technology*, vol. 133, Sept. 2023.

[11] C. Collewet and E. Marchand, "Colorimetry-based visual servoing," in *IEEE/RSJ IROS*, 2009.

[12] G. Silveira and E. Malis, "Visual servoing from robust direct color image registration," in *IEEE/RSJ IROS*, 2009, pp. 5450–5455.

[13] E. Marchand, "Direct visual servoing in the frequency domain," *IEEE RA-L*, vol. 5, no. 2, pp. 620–627, 2020.

[14] G. Zhu, Y. Huang, S. Li, J. Tang, and D. Liang, "Hyperspectral Band Selection via Rank Minimization," *IEEE GRS Letters*, vol. 14, no. 12, 2017.

[15] J. Chen and J. Yang, "Robust subspace segmentation via low-rank representation," *IEEE TCYB*, vol. 44, no. 8, 2014.

[16] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Squaring weighted low-rank subspace clustering for hyperspectral image band selection," in *IEEE IGARSS*, 2016.

[17] E. Malis and S. Benhimane, "A unified approach to visual tracking and servoing," *Robotics and Autonomous Systems*, vol. 52, no. 1, 2005.

[18] G. Silveira and E. Malis, "Direct visual servoing: Vision-based estimation and control using only nonmetric information," *IEEE T-RO*, vol. 28, no. 4, 2012.

[19] G. Silveira, L. Mirisola, and P. Morin, "Decoupled direct visual servoing," in *IEEE/RSJ IROS*, 2013.

[20] B. Tamadazte, N. L.-F. Piat, and E. Marchand, "A direct visual servoing scheme for automatic nanopositioning," *IEEE/ASME MECH*, vol. 17, no. 4, 2012.

[21] S. Tisserand, "Vis-nir hyperspectral cameras," *Photoniques*, pp. 58–64, 2021.

[22] Openrox open-source c library providing real-time algorithms for robotics. [Online]. Available: https://github.com/ACENTAURI-INRIA/openrox