Enhancing Thermal Fiducial Marker Detection: Focus on Image Processing Techniques

Andre França[1,2], Guido Berger[1,2] André Mendes[1,2], José Lima[1,2]

[1] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
[2] Laboratório para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
{a54187@alunos., guido.berger@, a.chaves@, jllima}ipb.pt

Abstract. This article proposes methods for maximising the detection rates of thermal fiducial markers using thermography. By exploring the combination of image processing techniques with the use of an affordable thermographic camera, the aim is to mitigate the negative effects of thermography and improve accurate marker identification in a variety of mounting and distance conditions. The research identified a diversity of processing techniques capable of improving thermal marker recognition, offering the potential to surpass previous results. The results highlight the possibility of using low-cost thermographic cameras for this purpose, which could democratise and reduce the costs of recognition processes. This methodology validates the proposed approach, providing a robust basis for future improvements in thermal marker detection and promoting the feasibility of practical, low-cost applications in an assortment of fields.

Keywords: Thermal · Fiducial Markers · Image processing · Processing techniques

1 Introduction

Accurate recognition of fiducial markers is crucial in applications such as Augmented Reality [16] and Autonomous Navigation [1,6,7]. In this context, the combined use of thermographic cameras and thermal markers offers a promising approach to equipment positioning, as discussed by Claro, Silva, and Pinto [1] and also by Khattak, Papachristos, and Alexis [5] where a combination of techniques, including thermography, is used to improve marker detection.

The identification of fiducial markers using thermography has advantages over visual capture, as it allows the equipment to identify the markers even in occlusive situations, such as fog, smoke, light reflection or even the influence of lighting [1,5].

However, as observed in the work carried out by Claro, Silva, and Pinto [1], marker detection using thermography showed inferior detection rates when compared to the RGB camera, except for the distance range of 2 to 4 metres. Even in this range, detection rates were lower than 85%, and for distances greater than 6 metres, the diffusion of heat affects the sharpness of the thermography, significantly reducing the detection of the marker [1].

This paper proposes researching image processing techniques to improve the detection rate of these markers in each captured frame. The proposed strategy combines filters and techniques native to the library OpenCV 4.7 [11] with information obtained by the thermographic camera SeekThermal Compact [17], an accessible, low-cost tool available to the general public. The aim was to identify the combination of features applied to the image that increases the robustness of the thermal marker detection system at distances between 60 cm and 500 cm.

By using affordable resources, this research seeks not only to advance thermal marker recognition but also has the potential to democratise and reduce the costs of recognition processes using thermographic cameras, opening the door to a variety of practical and affordable applications.

This paper covers a brief bibliographical review of image processing for the detection of fiducial markers in Section 2, the methodology used to get information and process the images collected in Section 3, the results obtained and the analysis performed on the data in Section 4, followed by a conclusion Section 5.

2 Related work

Fiducial markers are special elements widely used in computer vision to track and position virtual objects in the real world [1,6,7,9]. Among the various types of markers available, the Augmented Reality University of Cordoba (ArUco) markers are 2D binary patterns that allow fast and accurate detection and identification by computer vision algorithms. They are made up of a wide black border and an internal binary matrix that determines their identifier (ID) [4,9].

With this technology, in 2019, Khattak, Papachristos and Alexis [5] proposed the use of ArUco markers extended beyond the visual spectrum of light, through markers that, by means of the difference in thermal conductivity between two materials, produce the characteristic patterns of ArUco markers. They demonstrated the effectiveness of this method for environments with marker occlusion in the visible light spectrum.

Then in 2023, Claro, Silva and Pinto [1] proposed the creation of a fiducial marker based on ArUco technology that would use 3 layers of sensing to detect the marker: visual, thermal and depth. They demonstrated the effectiveness of sensor fusion using a proprietary algorithm, which achieved good values in the marker detection rate. However, thermal detection showed significantly lower values than visual detection.

In this way, this study aims to identify methods that provide improvements in marker detection rates, especially in the infrared spectrum. The focus will be on the native techniques of the OpenCV library [12,13], seeking to transform the

frames in a dynamic way that is compatible with real-time applications, based on the results presented in [1] as a reference.

3 Methodology

The methodology adopted in this study is described in detail in the following sections, providing a solid basis for understanding and reproducing the results obtained. Section 3.1 covers the data collection process, where the experiment was conducted using a thermal marker set and a capture kit. The thermal marker set consisted of a heating plate, a power source and a 2 mm thick cork marker, cut in the shape of the ArUco ID 0, based on the approach adopted in [1]. On the other hand, the capture kit consisted of a thermographic camera connected to a microcomputer, positioned on a tripod to ensure stability during data collection. The collection points were determined for a range of distances between the camera and the thermal marker from 60 cm to 500 cm, and the data was recorded using the Robot Operating System (ROS).

Analysis of the captured images is described in Section 3.2, where a series of transformations are applied to the collected data to improve the effectiveness of the ArUco recognition algorithm. Initially, the images are extracted from the *rosbag* files and saved in *jpeg* format, along with other associated messages. Next, groups of data are randomly selected for each point of recording and configuration, totalling 8000 frames, on which the transformations are applied. These include Adaptive Contrast, Filter and Thresholding. Each of these steps is explained in detail and their implementation is discussed, providing a complete understanding of the proposed processing process.

3.1 Data collection setup

The setup of the experiment for data collection followed the procedure sketched out in Figure 1, consisting of two main elements: the thermal marker pack on the left and the capture pack on the right of the figure.

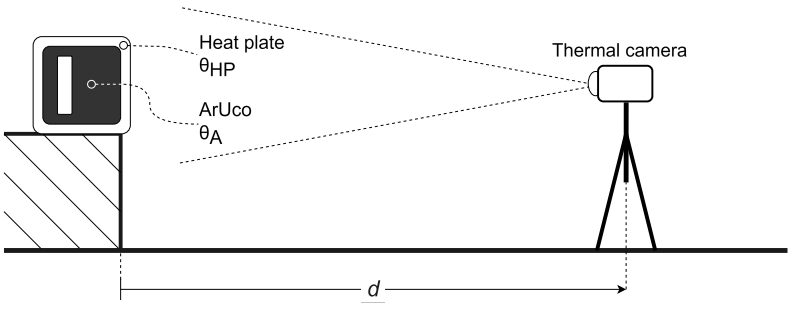
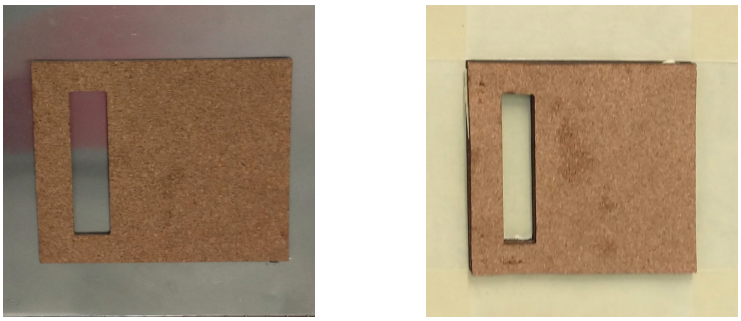


Fig. 1: Schematic of the data collection set-up

The thermal marker pack consisted of a polished aluminium heating plate, its power source and the 2 mm thick cork marker cut out in the shape of the tag ID 0 of ArUco [9]. The heating plate kept the metal part at a temperature of approximately 70 °C, which was measured using a culinary thermometer with a resolution of ± 0.5 °C. The exposed metal part of the hotplate was analysed in ways to analyse the influence of the reflectance of the material on the ArUco detection: without a masking tape coating, shown in Figure 2a, and with a masking tape coating, demonstrated in Figure 2b.



(a) Without masking tape (b) With masking tape

Fig. 2: Thermal fiducial marker

The capture kit consisted of a digital thermal imaging camera SeekThermal Compact [17] connected to an NVIDIA Jetson Nano [8] microcomputer, powered by a battery and positioned on a tripod to ensure stability for the time needed to collect the data. The data capture system for this experiment was based on ROS, due to its robust ecosystem and active community, maximising the scope of this research.

Afterwards, collection points were determined for the d value, illustrated in Figure 1, starting at 60 cm and with an interval of 20 cm until the value of 200 cm was reached. After this point, the interval increased to 25 cm until reaching a distance of 500 cm between the camera and the thermal marker, totalling 20 data collection points.

The data capture process consists of acquiring data from the camera via the messages published by the ROS package developed to interpret and publish data from SeekThermal cameras. These messages are captured using the recording and playback function of the ROS, the rosbag [2], for approximately 5 minutes, controlled by a timer, thus obtaining a large selection of frames to analyse. This results in around 1840 messages and frames per collection point, at a rate of approximately 6 Hz, containing information about thermography, camera data, captured frames and other information extracted from the camera.

This process is repeated for each previously demarcated position, applying changes to the experiment parameters as described in the schematic tree in

Figure 3. In this way, 4 different configurations are obtained for each camera position: with or without the masking tape; using the colour palette at the time of image capture White Hot or Black Hot.

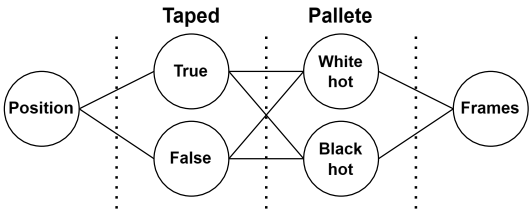


Fig. 3: Schematic tree of the data collection process

3.2 Image Analyses

Once the rosbags have been captured, the images are extracted from these files and saved in *jpeg* format, listed with a sequential identifier. Similarly, the other published messages are obtained by associating the content of these messages with a specific image using the identifier and saved in *csv* files. In this way, the aim is to facilitate the subsequent analysis processes that will be applied to this data and files, allowing simple and direct access and identification of data correspondence.

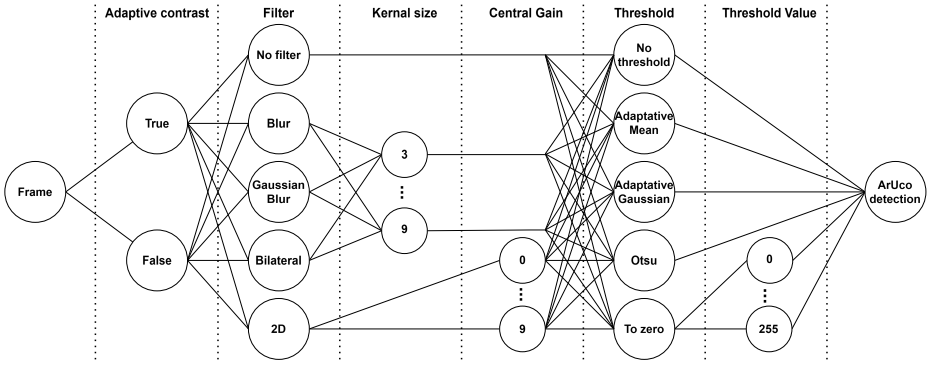


Fig. 4: Schematic tree of the frame analysis process

Subsequently, 100 frames and their metadata are randomly selected for each combination of recording point and configuration, thus obtaining a dataset of 8000 varied images, on which the sequence of analyses presented in Figure 4 will

be applied. In general, three transformations [3] are applied to the frame matrix to improve the effectiveness of the ArUco recognition algorithm: Adaptive Contrast; Filter; and Thresholding.

Adaptive Contrast adjusts the contrast between temperatures on the image by modifying the angular coefficient used by the camera to adjust the temperature values of the scene to the range of pixel values. Subsequently, the Filter is applied to reduce noise and distortions in the heat distribution that can negatively affect ArUco recognition. Various filtering techniques, such as Gaussian Blur and Bilateral Filter, are applied to the images to assess their impact on marker detection. Finally, Thresholding techniques are applied to enhance the characteristics of the marker in the frame, highlighting hot and cold areas and facilitating the detection process. More than 13500 combinations of transformations were applied to each frame, generating a large volume of comparative data.

Adaptive contrast: The adaptive contrast proposes adjusting the contrast between temperatures by modifying the value of the angular coefficient (α) used by the camera's interpreter to adjust the scene's temperature values to the pixel's value range (ρ), varying from 0, corresponding to the scene's minimum temperature (θ_{min}), to 255, corresponding to the scene's maximum temperature (θ_{max}), as illustrated in Figure 5.

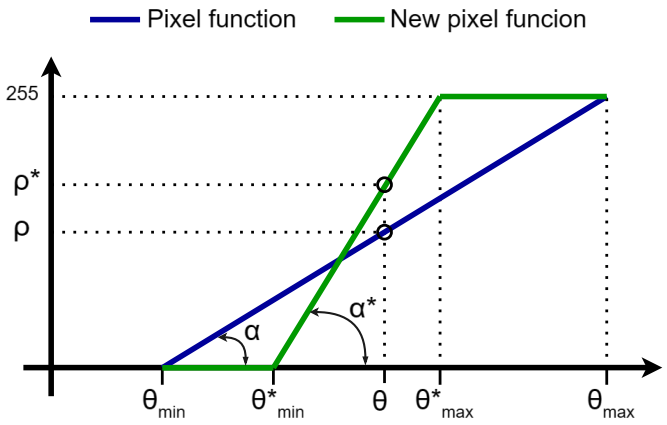


Fig. 5: Pixel values with and without adaptive contrast

[3] Results were obtained both with and without the application of the steps, to evaluate the effectiveness of the transformations

This is possible if the linear temperature adjustment is used by the camera's interpreter, Automatic Gain Control (AGC) linear [15], where the corresponding value for each pixel is given by Equation 1.

$$\rho = \alpha * (\theta - \theta_{min}) \tag{1}$$

In this way, it is possible to mathematically evaluate the value of α for a given frame using Equation 2. Since the values of θ_{max} and θ_{min} are provided through the metadata of each frame via the camera messages [3].

$$\alpha = \frac{255}{\theta_{max} - \theta_{min}} \tag{2}$$

To determine the value of the adjusted coefficient (α^*), a temperature range is determined which you want to emphasise, limited by θ_{min} to θ^*_{max}. Thus following Equation 3.

$$\alpha^* = \frac{255}{\theta^*_{max} - \theta^*_{min}} \tag{3}$$

The Equation 4 is then applied to each pixel in the image matrix, thus adjusting every pixel to its new value. The characteristic of this function is that for temperatures above θ_{max} and below θ^*_{min}, the pixel's saturation behaviour is observed, highlighting only the variations within the defined interval.

$$\rho^* = \begin{cases} \theta_{min} \leq \theta < \theta^*_{min}, & 0 \\ \theta^*_{min} \leq \theta \leq \theta^*_{max}, & \alpha^* * (\theta - \theta^*_{min}) \\ \theta^*_{max} < \theta \leq \theta_{max}, & 255 \end{cases} \tag{4}$$

Filter: The step following adaptive contrast is the application of filters, whose main objective is to reduce noise, and minimise details and distortions in heat distribution that could negatively impact the recognition of the ArUco. Blurring techniques are used, which convolve a kernel with an image, thus altering the perception of noise and edges [14].

The Averaging, Gaussian Blurring, Bilateral Filtering and 2D Convolution filters are applied in the image that came from Adaptative Contrast, native techniques of the OpenCV 4 [12] library, which require few parameters for their basic configuration. In these functions, the kernel parameters are varied, changing its size or gain.

For the Averaging filter, the *blur(..., ksize)* function was used, and for the Gaussian Blurring filter, the *GaussianBlur(..., ksize, sigmaX)* function was used. Where *sigmaX* has been defined as 0, being calculated internally by the function [12,14], and *ksize* is a tuple in the format *(k, k)*, with values of k assuming 3, 5, 7, 9 and 11, therefore applying more variation to the analysis of this filter.

Meanwhile, Bilateral Filtering was applied using the function *bilateralFilter(..., d, sigmaColour, sigmaSpace)*. For the values of d, the integers 3, 5, 7, 9 and 11 were assumed, it being recommended that for real-time applications $d \leq 5$, since filter processing becomes slow for values greater than 5 [12]. The

sigma values were both set at 75, opting for an intermediate value to the 150 mentioned as a "very strong effect" [12,14]. Thus, the variability of this filter lies in the multiple values of *d*.

The 2D Convolution filter was based on the function *filter2D(..., ddepth=-1, kernel)*, where *kernel* is a floating point matrix that can be defined globally or for each channel of the image [12]. The purpose of applying this filter is to emphasise the edges in the image and check the effects of this approach on tag detection. This work considered the *kernel* matrix presented in the equation 5, where i is called the central gain and takes positive integer values from 0 to 9.

$$kernel = \begin{bmatrix} -1 & -1 & -1 \\ -1 & i & -1 \\ -1 & -1 & -1 \end{bmatrix} \tag{5}$$

In this step, the frame is subjected to 25 transformations, which will be reprocessed in the next Threshold step, together with the frame with no filter applied, thus offering a wide range of variations to be analysed and compared.

Threshold: This step aims to increase the contrast between the hot and cold parts of the image, thus highlighting the tag of the surrounding area. It should be noted that the ArUco detection algorithm itself uses thresholding techniques as the first stage of image processing [4,9]. Therefore, the effects of this processing will be amplified when analysing the algorithm itself.

Three types of threshold were applied to the frames received from the previous step: Simple Thresholding, Adaptive Thresholding (Mean and Gaussian) and Otsu's Binarisation [10,13]. These are native techniques for OpenCV 4 [13], usually applied to grey-scale images. Like filters, the thresholding application reduces noise and alters the perception of edges [13].

To apply Simple Thresholding, the function *threshold(..., thresh, maxval, type)* [13] has been applied. Where, the input *maxval* is the maximum value that the pixel can take on, in the context of this work, it is a fixed value of 255. The TOZERO *type* was chosen, which saturates pixels below the *thresh* value to 0, but preserves pixels equal to or greater than the *thresh* value. In this way, *thresh* values from 0 to 255 were applied to each frame, generating 256 frames for each input frame, which are then sent to the ArUco detection algorithm.

Adaptive Thresholding has advantages over Simple since it can deal with different levels of illumination in the scene, but its behaviour only highlights the edges, altering the filling of elements [10]. As the marker is filled in, it is not expected to recognise the marker well. These thresholds are applied using the function *adaptiveThreshold(..., maxValue=255, adaptiveMethod, thresholdType, blockSize, C)* [13]. Where the value for the substrate constant, *C*, and the size referring to the pixel's neighbourhood, *blockSize*, the values shown in [10] were used. The *thresholdType* used was THRESH_BINARY, as it does not invert the image values, while for the *adaptiveMethod*, the ADAPTIVE_THRESH_MEAN_C and ADAPTIVE_THRESH_GAUSSIAN_C types were applied to each frame coming from the filter stage.

For Otsu's Binarisation, the function *threshold(..., 0, maxval=255, THRESH_BINARY + THRESH_OTSU)* [13] was used. This method applies a global threshold to the image, just like Simple Thresholding, but the Otsu method determines the optimal threshold value for the image. It is hoped that this method, combined with the Gaussian filter, will significantly reduce noise while preserving the edges and filling of the image [10].

4 Results and Analysis

On completion of the implementation of the proposed methodology, more than one hundred million (100,000,000) combinations of frames were analysed. Therefore, it was obtained a data set made up of a wide variety of possible paths for more accurate tag detection. In the analyses carried out in this section, we consider the detection effectiveness value (η), defined by dividing the number of frames that were successful in correctly recognising the marker (i.e. which reported the recognition of ID 0 without any other false recognition, in addition to the reported distance having an error of up to 10% concerning the actual distance), by the total number of frames in the analysed configuration.

Due to the large number of possible configurations for manipulating the images, this section will only analyse those that obtained a η value greater than 75%. To segment these configurations, the following aspects are considered: choice of camera colour palette (White hot or Black Hot), application of masking tape, application of adaptive contrast, filter used and threshold applied.

The analysis of the results obtained has been segmented in this section as follows: interpretation of the colour palette and application of the masking tape, in Section 4.1; Evaluation of the effects of Adaptive Contrast, Filter and Threshold, in Section 4.2.

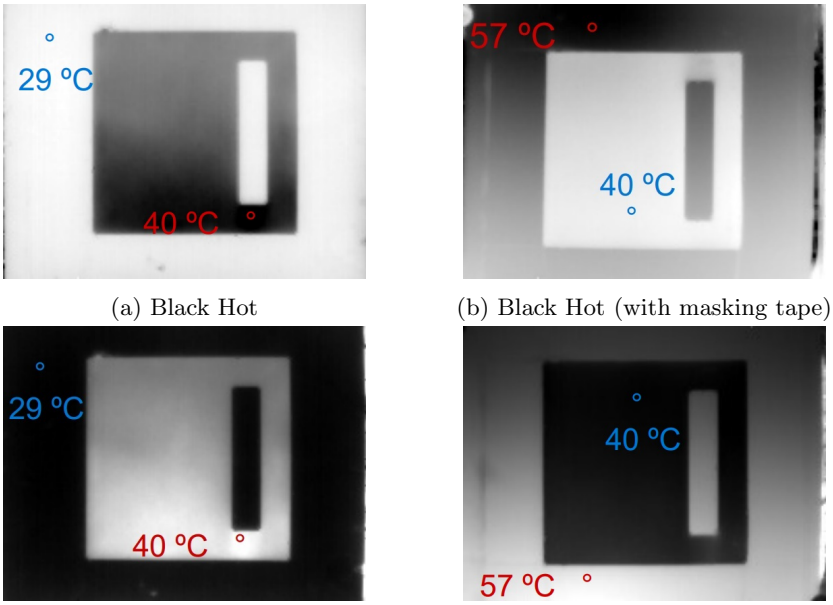
4.1 Interpreting the colour palette and applying masking tape

The first two variations proposed in the work methodology are changing the colour palette used by the camera's interpreter [3,15] and applying masking tape to the polished metal part of the thermal mark. Through these steps, it is possible to identify a phenomenon regarding the perception of temperature by the thermographic camera, where with the application of the tape a temperature of approximately 57 ºC is perceived by thermography for the metal part, on the other hand, in captures without the masking tape, the perceived temperature is 29 ºC. While the temperature of the cork marker is approximately 40 ºC for the 4 different configurations analysed, these values decrease with distance due to the diffusion effects of temperature [1], making them closer to the ambient temperature, which stayed close to 20 ºC in all the tests.

This interesting phenomenon occurs due to the emissivity of the material, as the polished metal section acts as a mirror, thus reflecting the infrared radiation from the environment [18], leading to the false impression that the metal section is low temperature. When the masking tape is applied to the polished part,

it blocks most of the reflective effect of the radiation from the surface. This procedure ends up "inverting" the perception of light and dark areas in the frame when comparing the Figures 6a with 6b or 6c with 6d, because it modifies the maximum and minimum temperature of the frame that is used to calculate the grey value of each pixel [15].

When comparing the images of Figure 6, the masking procedure worsens the contrast between the hot and cold areas of the thermal marker. However, this procedure has advantages when it comes to stabilising this temperature difference, as it minimises reflections in the infrared spectrum. Masking the metal surface makes the temperature difference between the hot and cold areas of the thermal marker more stable, predictable and consistent with the expected behaviour. The construction of the marker took into account the proposal by Claro, Silva, and Pinto [1] that the metal part (heat conductor) assumes higher temperatures than the cork part (heat insulator).



(a) Black Hot (b) Black Hot (with masking tape)

(c) White Hot (d) White Hot (with masking tape)

Fig. 6: Thermal image of tag at 80 cm (ambient temperature 20°C)

Analysing the combination of the colour palette together with the application of the masking tape, it is possible to observe the distribution of η by distance as presented in Figure 7. No combination achieved 75% of the η for all the collection points, i.e. at least one collection point did not achieve the desired detection efficacy. However, the configuration of Black Hot, combined with the non-application of masking tape, showed the best results, maintaining a η of over

85% for all collection points between 80 and 450 cm. Similarly, the configuration White Hot with masking tape showed values of over 80% at the vast majority of collection points, making it the best configuration with masking tape.

The application of the masking tape helped with detection at distances of 60 cm and 500 cm. For the other distances, the results were lower than when the masking tape was not applied, as well as presenting zones with η lower than 75%, as in the case of "With masking tape - Black Hot". Remember that this experiment was carried out in a laboratory environment, and the reflective effects of the polished metal ended up aiding detection by increasing the perceived temperature difference between the parts of the marker, also inverting the perception of the "hot part" and "cold part" of the marker by the thermographic camera. In outdoor environments, it is not expected to obtain such consistent values if there is thermal reflection.

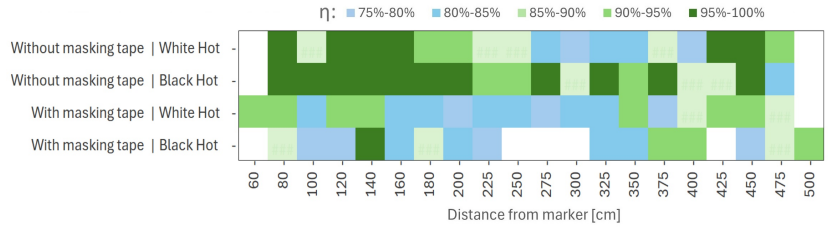


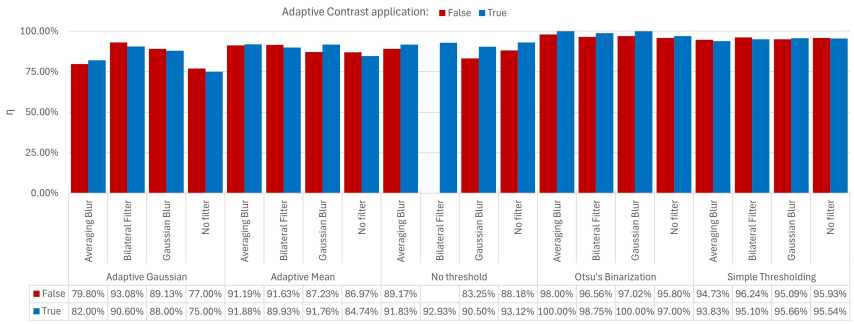
Fig. 7: Camera's colour palette and masking tape application analysis

4.2 Evaluation of the effects of Adaptive Contrast, Filter and Threshold

This section evaluates the effects of the Adaptive Contrast, Filter and Threshold steps on thermal marker recognition based on ArUco, following the sequential steps illustrated in Figure 4.

Adaptive Contrast: With the application of adaptive contrast, there were no significant effects on marker detection, as can be seen in Figure 8, which shows the η values for each type of Filter and Threshold. Only in the combination of a Bilateral filter and No threshold showed results below $\eta \leq 75\%$ when Adaptive Contrast was not applied, while when the technique was applied, detection accuracy exceeded 92%.

In some situations, there is an increase in η, reaching 100% frame detection. Generally, however, the percentage difference between the methods is low and does not give the expected result. In some combinations, there was even a deterioration with the application of adaptive contrast. This behaviour could be remedied in future work by refining this method and changing the temperature limits determined, among other aspects determined in Section 3.2.



	Adaptive Gaussian				Adaptive Mean				No threshold				Otsu's Binarization				Simple Thresholding			
False	79.80%	93.08%	89.13%	77.00%	91.19%	91.63%	87.23%	86.97%	89.17%		83.25%	88.18%	98.00%	96.56%	97.02%	95.80%	94.73%	96.24%	95.09%	95.93%
True	82.00%	90.60%	88.00%	75.00%	91.88%	89.93%	91.76%	84.74%	91.83%	92.93%	90.50%	93.12%	100.00%	98.75%	100.00%	97.00%	93.83%	95.10%	95.66%	95.54%

Fig. 8: Adaptive Contrast analysis

Filter: Through Figure 9, it is possible to observe a map representing the distribution of η as a function of distance, filter type and *kernel* value (in the case of Blur filters) or d parameter (in the case of Bilateral filters). There were no combinations with the 2D Convolution filter, as proposed in Section 3.2, capable of achieving significant detection values.

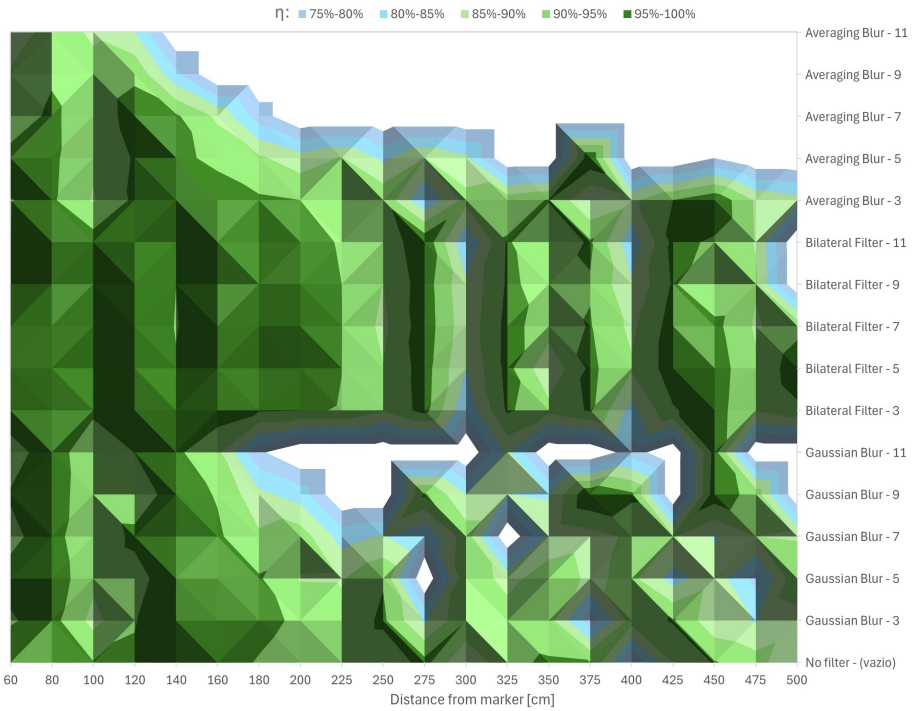


Fig. 9: Filter analysis

For the Averaging Blur Filter, it was observed that values of *kernel* greater than 5 resulted in a significant drop in detection for distances greater than 300 cm. However, this filter shows good results for smaller distances. Using a *kernel* value of 3, it was found that the η values were higher than those obtained without the filter for the majority of distances over 140 cm. On the other hand, the Gaussian Blur filters showed greater tolerance to increases in the *kernel* value but did not show appreciable gains in detection compared to not applying the filter.

The Bilateral filter showed the best results among the filters, especially for d values of 3, 5 and 7, showing gains in η values for almost all collection points. The range of d values that gave the best results is at the limit established for real-time applications ($d \leq 5$) [12], making this filter the most suitable for future applications.

Threshold: To evaluate the Threshold, the Adaptive Gaussian and Mean methods will be analysed together with Otsu's Binarisation, as these methods do not require the insertion of values for their application, while Simple Thresholding will be evaluated by the distribution of the values applied in *thresh* [13].

This results in the map shown in Figure 10, which indicates the η distribution for the thresholding techniques. It can be seen that Otsu's Binarisation and Adaptive Mean techniques stand out from Adaptive Gaussian, obtaining better detection results. In addition, they have an unexpected feature of complementarity, since Otsu's is effective up to 300 cm and Adaptive Mean becomes effective after this value. Also of note is that not applying any Threshold technique showed η values higher than 80% for distances of less than 250 cm, making it an interesting solution for shorter distances as it would save processing time.

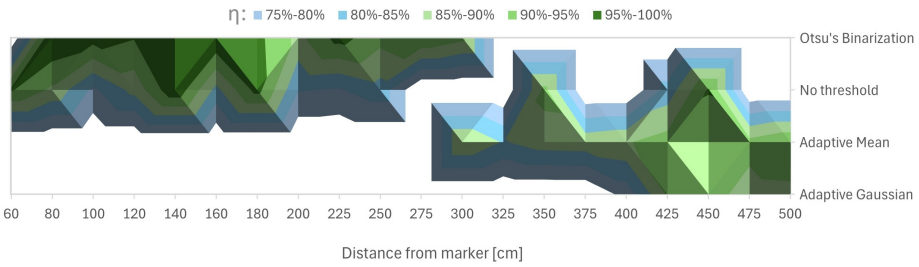
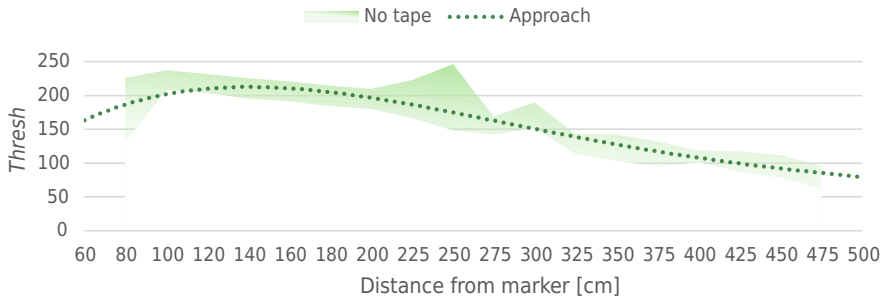


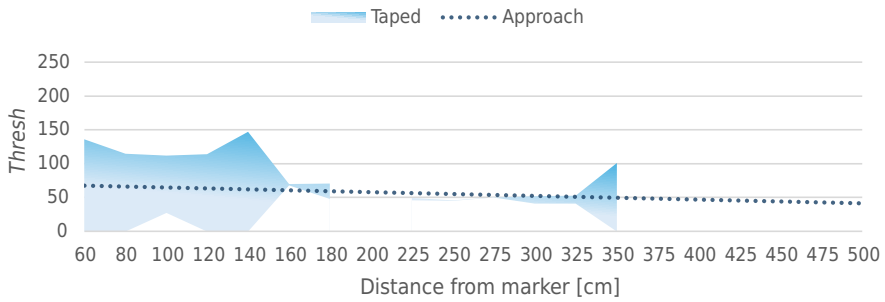
Fig. 10: Threshold analysis

The Simple Threshold is mainly influenced by the use or not of masking tape and by distance. The other factors such as palette, filter and application of adaptive contrast did not have a significant influence on determining a range of values for *thresh* that presents a relevant η. Thus, Figures 11a and 11b represent

the range of *thresh* values and a polynomial approximation of the average of the values presented.



(a) Without masking tape



(b) With masking tape

Fig. 11: Range of *thresh* by distance

It can be seen from Figure 11a that when the masking tape is not applied, the region representing *tresh* values has a narrow behaviour with large variations along the distance. For the polynomial approximation, a polynomial of degree 4 was used, with the equation 6, which resulted in an R^2 of 0.9615.

In Figure 11b, when the masking tape was applied, the *thresh* behaviour showed a higher tolerance for the first few points. However, for distance values greater than 160 cm, there was a drop in detection, with gaps, and distance intervals where no $\eta \geq 75\%$ occurred. The polynomial approximation was carried out using a polynomial of degree 1, i.e. a linear approximation, which resulted in the equation 7, with an R^2 of 0.3132, although it shows good tracking in the range presented.

$$thresh = -2.0277 * d^4 + 29.053 * d^3 - 146.66 * d^2 + 264.54 * d + 56.219 \quad (6)$$

$$thresh = -6.1392 * d + 70.594 \tag{7}$$

Where: d is the distance in m between the marker and the camera.

5 Conclusion

In summary, the analyses carried out in this study identified a variety of processing techniques capable of improving the recognition of thermal fiducial markers. Each sampling point has the potential for a specific combination of techniques that can surpass the thermographic marker detection results presented in previous work. It is important to emphasise that these results were obtained in a sampling environment and with a thermographic camera, among other factors, which differ considerably from the context used by the aforementioned authors. However, this study shows that there are ways to improve the identification of these markers.

Throughout the analysis, it was observed that various factors influence the detection of markers, from construction aspects to the post-processing of the captured images. A crucial factor is the distance between the camera and the marker, where the diffusion effect makes it difficult to perceive the marker, a behaviour reported in related works. However, using the techniques proposed in this study, it is possible to mitigate this effect, even with low-cost cameras.

As a suggestion for future work, it is suggested that mechanisms be developed to automatically select the methods analysed. Using the data provided by the cameras and an estimate of the camera-marker position, it would be possible to return the post-processing configuration that maximises the η value.

To conclude, this study contributes significantly to advances in thermal fiducial marker recognition, offering valuable insights and proposals for future research. The appropriate combination of processing techniques can substantially improve marker detection, opening doors to more accurate and efficient applications in various fields, such as Augmented Reality and Autonomous Navigation.

Acknowledgments

The authors are grateful to the Foundation for Science and Technology (FCT, Portugal) for financial support through national funds FCT/MCTES (PIDDAC) to CeDRI, UIDB/05757/2020 (DOI: 10.54499/UIDB/05757/2020) and UIDP/05757/2020 (DOI: 10.54499/UIDP/05757/2020) and SusTEC, LA/P/0007/2020 (DOI: 10.54499/LA/P/0007/2020).

References

1. Claro, R.M., Silva, D.B., Pinto, A.M.: Artuga: A novel multimodal fiducial marker for aerial robotics. Robotics and Autonomous Systems **163**, 104398 (5 2023). https://doi.org/10.1016/J.ROBOT.2023.104398

2. Field, T., Leibs, J., Bowman, J., Thomas, D., Perron, J., Carroll, M.: rosbag - ros wiki (2020), http://wiki.ros.org/rosbag
3. França, A.: Seek thermal ros (2023), https://github.com/alf767443/SeekThermal_ROS
4. Garrido, S., Panov, A.: Opencv: Detection of aruco markers, https://docs.opencv.org/4.x/d5/dae/tutorial_aruco_detection.html
5. Khattak, S., Papachristos, C., Alexis, K.: Marker based thermal-inertial localization for aerial robots in obscurant filled environments. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **11241 LNCS**, 565–575 (3 2019). https://doi.org/10.1007/978-3-030-03801-4_49, http://arxiv.org/abs/1903.00782http://dx.doi.org/10.1007/978-3-030-03801-4_49
6. Lebedev, I., Erashov, A., Shabanova, A.: Accurate autonomous uav landing using vision-based detection of aruco-marker. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **12336 LNAI**, 179–188 (2020). https://doi.org/10.1007/978-3-030-60337-3_18/COVER, https://link.springer.com/chapter/10.1007/978-3-030-60337-3_18
7. Miranda, V.R., Rezende, A.M., Rocha, T.L., Azpúrua, H., Pimenta, L.C., Freitas, G.M.: Autonomous navigation system for a delivery drone. Journal of Control, Automation and Electrical Systems **33**, 141–155 (2 2022). https://doi.org/10.1007/S40313-021-00828-4/TABLES/2, https://link.springer.com/article/10.1007/s40313-021-00828-4
8. NVIDIA: Kit de desenvolvedor jetson nano para ai e robótica — nvidia, https://www.nvidia.com/pt-br/autonomous-machines/embedded-systems/jetson-nano/
9. OpenCV: Opencv: Aruco markers and boards detection for robust camera pose estimation, https://docs.opencv.org/4.x/de/d67/group__objdetect__aruco.html
10. OpenCV: Opencv: Image thresholding, https://docs.opencv.org/3.4/d7/d4d/tutorial_py_thresholding.html
11. OpenCV: Opencv 4.7.0 is now available! - opencv (2022), https://opencv.org/blog/opencv-4-7-0/
12. OpenCV: Opencv: Image filtering (2024), https://docs.opencv.org/4.x/d4/d86/group__imgproc__filter.html
13. OpenCV: Opencv: Miscellaneous image transformations (2024), https://docs.opencv.org/3.4/d7/d1b/group__imgproc__misc.html#gae8a4a146d1ca78c626a53577199e9c57
14. OpenCV: Opencv: Smoothing images (2024), https://docs.opencv.org/4.x/d4/d13/tutorial_py_filtering.html
15. SeekThermal: seekthermal/seekcamera-python: Python language bindings for the seek thermal sdk (2024), https://github.com/seekthermal/seekcamera-python
16. Simões, E.: Targets para realidade aumentada. o target, ou marcador, é uma imagem no... — by erick simões — nave recife — medium (2020), https://medium.com/nave-recife/targets-para-realidade-aumentada-60235dda01a0
17. Thermal, S.: Thermal cameras for your smartphone - seek thermal — affordable infrared thermal imaging cameras (2024), https://www.thermal.com/compact-series.html
18. Urzedowski, A., Migasiuk, D.W., Buraczyńska, B.: Visual effects of surface emissivity in thermal imaging. Advances in Science and Technology Research Journal **14**, 215–222 (6 2020). https://doi.org/10.12913/22998624/118103