A Versatile High-Performance Visual Fiducial Marker Detection System with Scalable Identity Encoding

Peter Lightbody
plightbody@lincoln.ac.uk

Tomáš Krajník
tkrajnik@lincoln.ac.uk

Marc Hanheide
mhanheide@lincoln.ac.uk

Lincoln Centre for Autonomous Systems Research,
School of Computer Science, University of Lincoln, United Kingdom

ABSTRACT

Fiducial markers have a wide field of applications in robotics, ranging from external localisation of single robots or robotic swarms, over self-localisation in marker-augmented environments, to simplifying perception by tagging objects in a robot's surrounding. We propose a new family of circular markers allowing for a computationally efficient detection, identification and full 3D position estimation. A key concept of our system is the separation of the detection and identification steps, where the first step is based on a computationally efficient circular marker detection, and the identification step is based on an open-ended 'Necklace code', which allows for a theoretically infinite number of individually identifiable markers. The experimental evaluation of the system on a real robot indicates that while the proposed algorithm achieves similar accuracy to other state-of-the-art methods, it is faster by two orders of magnitude and it can detect markers from longer distances.

CCS Concepts

•CCS → Computing methodologies → Artificial intelligence → Computer vision → Computer vision tasks → Vision for robotics

Keywords

Swarm Robotics; Computer Vision; Fiducial Markers

1. INTRODUCTION

Although originally designed for Augmented Reality (AR) applications, fiducial-based visual localisation systems are widely used in a number of areas throughout the field of robotics where robust and efficient full pose vision-based estimation is required. Thus, typical applications of such marker-based systems include swarm and bio-inspired robotics

(a) WHYCON (b) ARTAGS (c) APRILTAG (d) WHYCODE
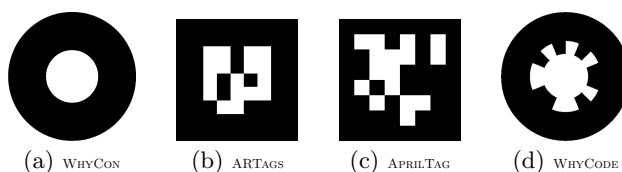
(e) Swarm of robots tagged with WHYCODE markers

Figure 1: Four types of fiducial markers: the state-of-the-art WhyCon, ARTags, AprilTag and the proposed WhyCode and a robotic swarm tagged with the WhyCode markers.

[1, 2, 10], which requires reliable localisation of large number of robots from an external camera(see also Figure 1(e)), visual-servoing that needs highly precise robot motion [19, 23], and semantic scene understanding [6, 15], where the scene objects are tagged with the fiducial markers to mitigate the drawbacks of general vision-based object recognition.

In any of these applications, and also more generally, visual fiducial marker detection and tracking systems must, ideally, fulfil all the following requirements to a high standard:

- **Robustness:** Markers must be robustly detectable in adverse conditions, e.g. while rapidly moving, from a considerable distance, under varying lighting conditions, etc.

- **Distinguishability:** In most applications, a single marker is not enough, either because several robots are to be tracked in parallel, or several objects or features in the environment need to be identified simultaneously. Thus, markers need to be both robustly identifiable and distinguishable to the vision system. The number of markers required, however, varies considerably from application to application. The marker tracking method must, therefore, be able to scale accordingly to the requirements imposed by the specific application or scenario.

- **Economic Feasibility:** Markers should ideally be affordable and easy to produce in large quantities, whilst also using cheap and readily available sensor(s). Hence, fiducial makers are often printed on paper and detected with standard RGB or grey-scale cameras. This, together with freely-available, open-source software, makes them customisable, minimising the burden for developers and researchers alike.

- **Precision:** The purpose of fiducial markers is to provide a precise position, and most often also an orientation, in 3D space. Accordingly, most markers often have properties that allow the estimation of their 6 degrees of freedom (DoF) pose, while still being planar.

In this paper, we propose a novel marker system that can generate suitable markers that can easily be printed on paper and an integrated software component that addresses the above requirements to a very high standard. In this work, we extend an open-source detection system for circular markers called WHYCON [11] by adding a novel encoding system based on the concept of *Binary Necklaces* [18], which we shall refer to as WHYCODE. Necklaces are a mathematical concept of combinatorics providing a *generator* for rotation invariant, uniquely identifiable patterns that can scale to a theoretically infinite number of individual markers, similar to the one shown in Fig. 1(d). The resulting markers are robustly and efficiently detectable in the environment and at the same time allow for discrimination between individual markers using the Necklace coding. With our extension of the original system we now present a 6-DoF fiducial marker system. The performance of the proposed system is demonstrated through a range of experiments which compare the new WHYCODE method against the frequently used ARTAGS and APRILTAG fiducial marker detection systems.

2. RELATED WORK

As discussed above, the need for vision based markers within robotics is evident. To meet this demand, several marker-based tracking and identification methods have been developed. Many of the markers commonly used within this area are designed to store large amounts of information, one such example being a QR code. This marker consists of a two-dimensional matrix barcode which encodes data in a pattern of black and white squares. In-built error correction codes allow the information to be correctly read, even if the marker is partly damaged, although these characteristics do restrict the range and angles from which the codes can be read. Consequently, although there is the potential to use such markers as part of a larger tracking system, their design makes them less suitable for tracking than both the methods discussed below and the proposed method.

Although this system utilises passive vision-based markers, a widely used approach within the field at the minute is the commercial motion capture system, ViCon [21]. By combining high-resolution and high-speed cameras with strong infra-red (IR) emitters, this system enables tracking with sub-millimetre precision by attaching IR reflective markers to mobile robots. Although these attributes give ViCon a solid ground truth, this approach remains a very costly system and is therefore not always an appropriate solution. This issue has, however, motivated a variety of works which propose a number of alternative low-cost tracking systems, many of which focus on passive vision-based tracking. Many of these newer methods use simple planar printable patterns which significantly lower not only the cost, but also the difficulty of use and set-up.

Another system often employed in this area is augmented-reality orientated markers, which, similarly to the proposed marker, allow additional information to be encoded such as a target ID. These systems often use the ARTag [7] and ARToolKit+ [22] software libraries.

The current ARTAGS developed from these software libraries utilise a square box fiducial marker which encodes information through the use of a large 2D black and white bar code. The real time performance of the system, coupled with its accuracy and robust nature, make it an ideal candidate for a comparison to the proposed system.

Further to this, the square marker of the APRILTAG [17] system also provides a robust and reliable system which can provide the position, orientation and identity of a tag. Similarly to the proposed system, the APRILTAG also stems from a lexicographic coding system [20] which, although conceptually similar to the QR code mentioned above, is designed to encode far smaller data payloads. Although this system is capable of providing robust detection at both short and long range, computational simplicity is sacrificed.

Despite the success of square markers within this field of research, the use of circular markers is quickly becoming a regular occurrence in many applications. This is largely due to the need to counter the shifting of the centroid of a square marker under perspective transformation. The less expensive centroid operation for circles has led to its use in a number of systems such as SyRoTek e-learning platform [13], which uses ring-shaped patterns with binary tags, and [23], a planar pattern which consists of the letter 'H' surrounded by a ring. In the latter, the pattern is initially detected using adaptive thresholding and later processed for connected component labelling. In order to determine whether the marker has been correctly tracked, its geometric properties are then tested after which false matches are discarded and a Canny edge detector and ellipse fitting method are applied to the positive matches.

The ARToolKit and ARTAGS markers mentioned above have also been developed into another system known as ArUco [9], which promotes a robust ID system which itself involves an error correction system which can handle up to 1024 individual codes. The detection process of the AR markers

used within the ArUco system combines the aforementioned adaptive thresholding step with contour extraction and code identification, thus determining the extrinsic parameters of the marker using the intrinsic camera parameters.

Finally, a system relatively similar to the proposed markers is the TRIP localisation system [5], where the marker comprises a number of concentric circles broken into several angular regions and coloured either black or white. This method can distinguish between 39 patterns, a performance comparable to the 30 options available when 8 beads are used within the proposed marker. Similarly to the ArUco system mentioned above, this system appropriates an adaptive thresholding method, with the system as a whole extracting the edges of the markers and processing the edges which correspond to the circular border of the ring patterns. These detected edges are then passed through an ellipse-fitting method which checks the concentricity of the ellipse. As the adaptive thresholding can be quite computationally expensive, this system can be costly in its performance, however, this disadvantage may be counteracted by the system's ability to achieve a precision of between 1% and 3% of relative error.

Although the aforementioned methods are widely considered to be the state-of-the-art methods currently existing within this field, the real-world performance and low computational cost of the proposed method makes it superior to many of the systems indicated above. The ability to expand the recognisable patterns by incorporating a greater number of segments also makes the proposed method preferable to a number of the most commonly used systems.

3. MARKER LOCALISATION

The WHYCON algorithm was originally intended for computationally efficient localisation of a large number of patterns composed of concentric black and white circles of known diameter. The article [12] shows that the method achieves the same precision as state-of-the-art black-and-white pattern detectors while being faster by two orders of magnitude.

To find a circular pattern, the algorithm searches the image using a combination of flood-fill techniques and on-demand thresholding. The algorithm gathers statistical data of the patterns during their segmentation, which allows for rapid rejection of false candidates. The pattern search can be initiated from any position in the image, which, when combined with tracking, typically causes the algorithm to process only the pixels that are occupied by the patterns tracked. From a computational complexity perspective, this results in a significant performance boost.

In the initial phase of the pattern detection, the image pixels are searched for a continuous segment of black pixels, which are classified by an adaptive thresholding method that ensures good robustness to adverse lighting conditions. Once a continuous segment of black pixels is found by the flood-fill method, a simple circularity test is performed. A pattern consisting of s pixels, with bounding box dimensions b_u, b_v and inner and outer diameters d_i, d_o is considered circular if its 'roundness' ρ_{out} is smaller than a predefined value ρ_{max},

i.e.

$$\rho_{max} > |\rho_{out}| = \left| \frac{\pi}{4s} b_u b_v \frac{d_o^2 - d_i^2}{d_o^2} - 1 \right|. \qquad (1)$$

Once the black segment passes the circularity test, a new flood-fill search is initiated from its centre to locate the inner white segment. Since the inner segment is a circle, its circularity test is simpler than Equation 1:

$$\rho_{max} > |\rho_{in}| = \left| \frac{\pi}{4s} b_u b_v - 1 \right|. \qquad (2)$$

If the inner segment passes Equation 2, the algorithm compares positions of their centres to verify if the segments are concentric. Then the method calculates the ratio of their number of pixels and verifies if this ratio conforms to the known ratio of the black and white segments' areas.

After passing these tests, the positions of the segments' pixels u_i, v_i that were stored during the flood-fill search are used to calculate the pattern's centre u, v and covariance matrix $\mathbf{C}$ as follows:

$$\mathbf{C} = \frac{1}{s} \sum_{i=0}^{s-1} \begin{pmatrix} u_i u_i & u_i v_i \\ u_i v_i & v_i v_i \end{pmatrix} - \begin{pmatrix} uu & uv \\ uv & vv \end{pmatrix}. \qquad (3)$$

Note that u_i, v_i are integers, and the computationally most expensive part of Equation 3 is calculated using integer arithmetic. The u_i, v_i and $\mathbf{C}$ actually represent an elliptical projection of the pattern in the image.

Then, we calculate the eigenvalues λ_0, λ_1 and eigenvectors v_0, v_1 of the covariance matrix $\mathbf{C}$ and use them to determine ellipse semiaxes e_0, e_1 as follows:

$$\begin{aligned} e_0 &= 2\lambda_0^{\frac{1}{2}} v_0, \\ e_1 &= 2\lambda_1^{\frac{1}{2}} v_1. \end{aligned} \qquad (4)$$

Knowing the length of the ellipse semiaxes, we perform a final segment test, which verifies if the number of pixels s corresponds to the area of the ellipse:

$$\xi > |\pi e_0 e_1/s - 1|. \qquad (5)$$

The constant ξ represents a tolerance value much lower than ρ_{max}, because the ellipse dimensions e_0, e_1 are obtained from the covariance matrix with sub-pixel precision. If the detected segments satisfy Equation 4, they are assumed to represent the pattern. The obtained eigenvalues and eigenvectors are then used to calculate the spatial position of the pattern.

To obtain the relative distance of the pattern, we calculate the pixel coordinates of the ellipse (co-)vertices, transform these into canonical camera coordinates using the intrinsic camera parameters that were obtained through standard camera calibration procedure. The transformed coordinates of the (co-)vertices are used to calculate the centre and axes of the ellipse in the canonical camera form. The vertices are used to calculate a conic $\mathbf{Q}$ such that all the ellipse points u', v' satisfy

$$\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix}^T \mathbf{Q} \begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0. \qquad (6)$$

Then, we calculate the eigenvalues $\lambda_0, \lambda_1, \lambda_2$ and eigenvectors q_0, q_1, q_2 of the conic $\mathbf{Q}$ and use them to obtain the spatial position of the pattern by the method presented in [23]:

$$\mathbf{x} = \frac{d_o}{\sqrt{-\lambda_0 \lambda_2}} \left(s_1 q_0 \lambda_2 \sqrt{\frac{\lambda_0 - \lambda_1}{\lambda_0 - \lambda_2}} + s_2 q_2 \lambda_0 \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_0 - \lambda_2}} \right),$$
(7)

where d_o is the circular pattern diameter.

In this work, we also implement a calculation of the pattern attitude. At first, we calculate the normal $\mathbf{t}$ by

$$\mathbf{t} = \left(s_1 q_0 \sqrt{\frac{\lambda_0 - \lambda_1}{\lambda_0 - \lambda_2}} + s_2 q_2 \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_0 - \lambda_2}} \right).$$
(8)

Note that the constants s_1 and s_2 are undetermined signs that have to be selected so that the $\mathbf{n}$ points towards the camera and x is in front of it. In other words, s_1 and s_2 are chosen so that the inequalities:

$$\begin{aligned} \mathbf{n}(0,0,1)^T &< 0 \\ \mathbf{x}(0,0,1)^T &> 0 \end{aligned}$$
(9)

are satisfied. While the roll and pitch of the pattern can be expressed from the normal $\mathbf{n}$, the yaw of the original circular marker can not be determined. However, the yaw can be calculated in the subsequent step, which uses the Necklace code for the pattern identification.

4. MARKER IDENTIFICATION

Considering our requirements outlined in the introduction, and building upon the good detection performance of the WHYCON system discussed in the previous section, our development of a new marker system focused on creating a marker which is compatible with the circular features of WHYCON, but also capable of providing a scalable encoding of IDs. The proposed encoding chosen for the WHY-CODE marker was originally identified within the combinatorics field of mathematics, and currently used widely in the fields of combinatorial chemistry [3] and computational biology [4]. These sequence patterns known as Necklaces are "lexicographically the smallest element in an equivalence class of strings under string rotation" [18].

Although not currently used in the field of robotics, this encoding is a highly suitable option for the system due to its rotational invariant nature. By bit-shifting the detected sequence of code until the lowest binary value is reached, the system is able to identify a starting point regardless of the position of the code that is being detected. This ability to alter the detected code without confusing the IDs also means that the inability to identify a yaw rotation, and thus a starting point, on the circular markers is circumvented.

In addition to this benefit, the Necklace Encoding system also allows the rotation of the marker to be calculated. By adjusting the detected code by the number of times the code was bit-shifted to achieve the lowest binary value, the marker's yaw rotation in 3D space can be calculated. As the ID is encoded by bit-shifting each number to their lowest binary value, both the ID calculation and subsequent yaw rotation can both be pre-calculated to minimise computational costs. However, for this to work reliably all codes which have

rotational symmetry, must also be removed from the encoding system, as they allow for the lowest binary value to be reached from multiple start locations, which would result in ambiguity when establishing the markers' yaw. To see an example of a marker with ambiguous yaw, see the leftmost quadcopter on Figure 1(e).
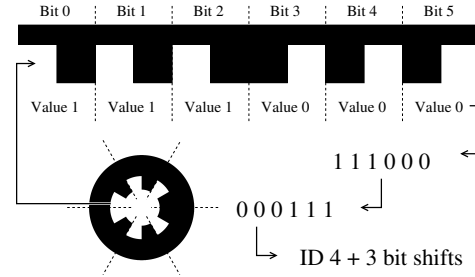

Figure 2: An example of how the Manchester Encoding is used with the Necklace System: The inner circle of the WhyCode marker encodes a binary string which is bit-shifted to match a Necklace code. Apart from identification, the number of bit-shifts allows us to identify the marker's rotation.

To create a system which reliability identifies the markers and preserves backward compatibility with the WHYCON marker, we encoded the Necklace-based ID into the inner circle of the tag using Manchester Encoding [8]. Thus, each individual bit of the Necklace code is encoded by two consecutive segments of opposite colour, as demonstrated in Figure 2. Although the use of Manchester Encoding halves the number of segments available on the marker, it allows us to calculate an identification confidence rating based on the expected number of pixels in each segment of the Necklace code.

In theory, the Necklace Encoding supports higher-than binary bases, and it would be possible to encode the marker IDs in greyscale values along the inner circle rim. However, preliminary tests showed that the edge-based Manchester Encoding is more suitable due to its robustness. This has the benefit of making the system more robust, especially when subject to various lighting condition, but does have the negative effect of only allowing binary-code sequences when encoding IDs. As a result, this restricts the encoding system and limits the number of potential IDs to:

$$N = \frac{1}{n} \sum_{d=1}^{n} \varphi(d) 2^{n/d},$$
(10)

where $\varphi()$ is totient function [14] and n is the Necklace code length in bits. The Equation 10 is further illustrated in Table 1 which shows the number of combinations valid for the proposed marker, given that the Necklace code consists of a sequence of n bits:

5. EXPERIMENTS

To evaluate the performance of the proposed marker, we compared its localisation accuracy, detection range and identification reliability to state-of-the-art fiducial markers in a series of real experiments. Each of these tests used an RGB

Table 1: Necklace code length in bits and corresponding number of unique marker identities

Code length	[bit]	4	6	8	10	12	14	16
Unique IDs	[-]		3	9	30	99	335	1161	4080

camera of an ASUS Xtion RGB-D sensor, as it corresponds with the type of sensor that is widely used on robotic platforms, providing a standard 640×480 image at 25 frames per second. This sensor was fixed to a FLIR E46-17.5 Pan Tilt Unit (PTU) which provided a ground truth for the marker position, attitude and velocity. This PTU was also mounted atop a mobile platform with a SICK s300 laser scanner. As the detectable range of the markers exceeds the range of a ASUS depth camera, the laser scanner with a range of up to 30m provided a reliable distance measurement that was also used for the ground truth in some of the experiments. To allow for a fair comparison of the proposed marker against the ARTAGS and APRILTAG, each of these markers were resized to occupy the same area of 3.817cm^2. A default calibration was also used, rather than specifically calibrating the camera, to demonstrate the system's performance in standard circumstances.

5.1 Detection and identification range

The first test aimed to evaluate the effect that distance had on the performance of the system. The markers were affixed to the wall at a height equal to that of the camera. The mobile platform was then programmed to move backwards from a distance of 0.2 metres until the platform reached a distance of 7 metres from the wall. The movement occurred at a constant speed of 0.02 metres per second, which was selected in order to ensure that motion blur was not a factor.

Table 2: Maximum distances at which the markers were consistently detected and identified [m]

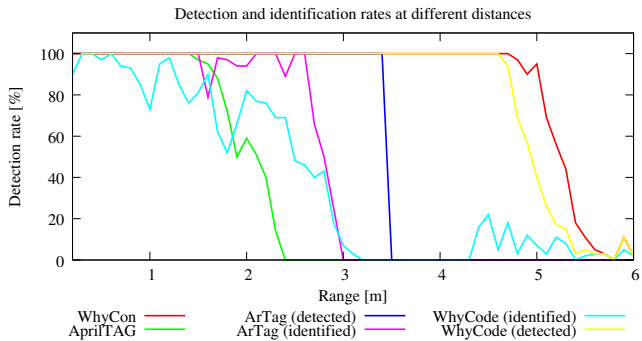|	WHYCON	APRILTAG	WHYCODE	ARTAGS
Detection	5.4	2.1	4.9	3.4
Identification	–	2.1	2.4	2.7

Figure 3: Maximum distances at which the markers were consistently detected and identified

As can be seen in Table 2 and Figure 3 the original WHYCON

marker has proven to achieve the longest detection range of 5.4 m. Although the WHYCODE marker was almost able to achieve a similar range, the new marker started to provide incorrect IDs once the distance had surpassed 2.4 metres. Similarly to that, the ARTAGS were undetectable at a range of 3.5 metres or more, and their correct identification was not reliable when the distance of the marker exceeded 2.7 metres. As for the APRILTAG, no incorrect IDs were reported. However, the distance at which the marker was reliably detectable was the lowest of the markers tested at only 2.1 metres.

5.2 Identification range vs. code length

A similar test was also conducted on the WHYCODE marker to identify how changing the number of encoding bits affects the range at which the encoding can be correctly identified. As can be seen in Figure 4 using less than 8 bits for the code does not affect the range, while increasing it has a negative impact on the identification range. This corresponds with the expectation that the limiting factor of identification range is the size of the individual elements that make up the encoding pattern.
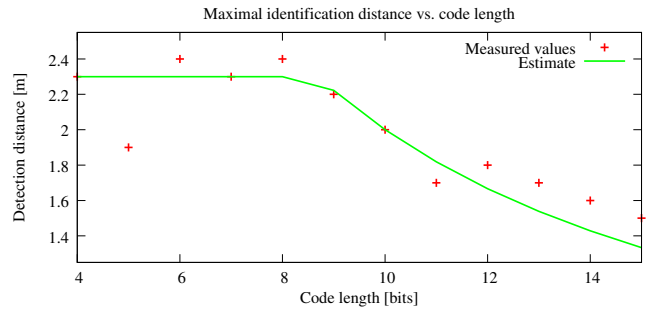

Figure 4: Dependence of maximal identification range on the Necklace code length n. The estimate is based on a formula $min(2.4, 200/n)$

5.3 Robustness to motion blur

This test, which was intended to analyse the effect of motion blur on the markers, involved keeping the markers stationary whilst rotating the PTU. This setup not only ensured the equal movement of all the markers, but also created a stable, continuous and repeatable experiment which represented one of the system's intended applications: mobile robotic platforms with a moving on-board camera. With the markers affixed to the wall, the camera was placed exactly 1 metre from the wall and the PTU rotated from -90 degrees to +90 degrees at a constant speed. Figure 5 shows the speeds that were tested during this experiment with the resulting detection and identification ratios.

These results indicate that while both WHYCODE and WHYCON systems are less susceptible to motion blur, the APRILTAG identification scheme is more robust to motion blur compared to WHYCODE.

When attempting to decode the ID, the WHYCODE marker reported a number of incorrect results at the faster motions, which is caused by the fact that the code does not employ any error detection or self-correction scheme. In contrast,
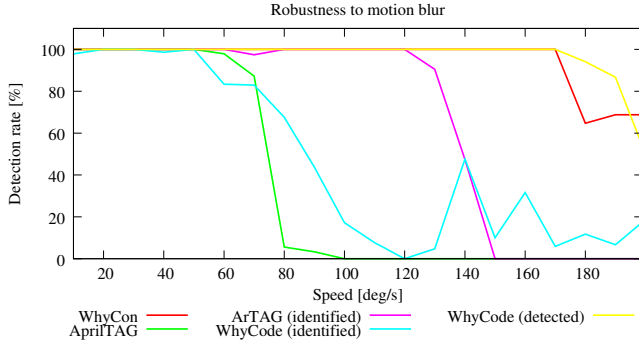
Figure 5: The results of the Motion Blur experiment - dependence of the detection rate on the marker velocity.

the lexicographic error correcting [20] used by the APRILTAG meant that no incorrect IDs were detected during our tests.

5.4 Accuracy of angle estimation

Since the x, y, z position estimation is identical to the original WHYCON method [12], which reports that its localisation accuracy is comparable to ARTAGS based markers, we tested only the accuracy of angle estimation. In contrast to the earlier experiments, the markers were this time placed on the robot's PTU which, whilst facing the free-standing stationary camera, used the pan and tilt functions to vary the angle of the markers. The recorded positions and rotations of the markers were then compared to the angle taken from the PTU. This comparison was then used to calculate an error rate for the system, see Table 3.

Table 3: Average error of angle estimates [radians]

	WHYCON	APRILTAG	WHYCODE	ARTAGS
Pitch/roll	0.024	0.023	0.020	0.038
Yaw	——	0.034	0.042	0.044

As can be seen from the above table, all markers exhibited average errors lower than 0.05 radians demonstrating that the system's ability to establish the marker's orientation was successful across all four systems. It should be noted that while the original WHYCON marker is unable to provide the yaw rotation, WHYCODE can estimate the yaw rotation with a high level of accuracy using the Necklace Encoding.

5.5 Robustness to illumination changes

The last test aimed to verify the performance of the system when subjected to various lighting conditions. To achieve this, the markers were positioned next to a large window in order to utilise natural, ambient light and avoid the flickering sometimes caused by artificial light. By taking a photo every 10 seconds during the 25 minutes before and during sunrise, the markers were able to go from complete darkness to normal daytime lighting conditions. While the ARTAGS were detected in 64% of these images, APRILTAG, WHYCON and WHYCODE were detected in 71%, 72%, 74% of images

respectively. Since the slight differences in performance may be attributable to slight variations in light, we can state that all the markers demonstrated a similar robustness to variable illumination.

5.6 Computational complexity

In addition to the above tests, a number of computational performance tests were conducted on each of the systems. The first of these were conducted using procedurally generated images of size 5000×5000 pixels containing over 550 randomly placed markers. This test helped to evaluate each of the systems ability to handle, not only large images, but also images which contain high number of markers and varying levels of clutter. Although WHYCON and WHYCODE took more than a second to process the first frame, each subsequent frame was then processed significantly faster. The average time to process a single frame when comparing the APRILTAG and the WHYCODE systems can be seen in Table 4, which shows the main advantage of the WHYCODE method – its computational efficiency. Table 4 also shows that identification and yaw estimation step do not slow down the original WHYCON method, which is two orders of magnitude faster than the ARTAGS and APRILTAG.

The performance boost WHYCON and WHYCODE results from the on-the-fly calculation of the detected segment statistics, which is naturally achieved by the flood-fill segmentation technique and which allows tracking without any computational overhead. Although the computational efficiency of both ARTAGS and APRILTAG could be improved by employing some tracking scheme, it is unlikely to achieve a two-orders of magnitude speed-up.

Table 4: Average processing time of an image with 550 markers [seconds]

Clutter	WHYCON	APRILTAG	WHYCODE	ARTAGS
none	0.06	3	0.06	3
little	0.07	16	0.07	14
large	0.07	15	0.07	15

6. CONCLUSION

In this paper, we present an extension to the marker used by the WhyCon tracking system. The proposed marker not only utilises a new encoding method which allows identification of each marker, but also extends the algorithm to allow the full localisation of a marker with 6 DOF. By keeping the simple roundel design, the proposed marker is not only backwards compatible with the previous system, but also maintains its sub-pixel (2D) and millimetre (3D) precision, and high computational efficiency.

The results of our study show that the WHYCODE system, despite the additional overhead of having to decode marker IDs, performed similarly to the original WHYCON system and outperformed the comparative systems in both accuracy and speed. By exceeding the high level of performance demonstrated by the APRILTAG and ARTAGS, and at two orders of magnitude faster, the proposed system achieves

a strong level of accuracy without the high computational requirements. These achievements therefore make the proposed system particularly applicable to resource-constrained systems and scenarios, where the reliable and swift tracking of multiple robots is a necessity. Moreover, the WhyCon system can reliably detect smaller markers at longer ranges, which is also makes it a popular alternative to APRILTAG or ARTAGS. The entire system is available as an open-source package at https://github.com/LCAS/whycon.

In the future, we will explicitly model uncertainty of the marker locations, which should not only improve our system's accuracy [16], but also its coverage by allowing to fuse input from multiple cameras.

Acknowledgements

7. REFERENCES

[1] F. Arvin, T. Krajník, A. E. Turgut, and S. Yue. COS-ϕ: artificial pheromone system for robotic swarms research. In *IROS*. IEEE, 2015.

[2] F. Arvin, A. E. Turgut, T. Krajník, and S. Yue. Investigation of cue-based aggregation in static and dynamic environments with a mobile robot swarm. *Adaptive Behavior*, 2016.

[3] L. Batten et al. Classification of Chemical Compound Pharmacophore Structures. 1999.

[4] W. Y. Chen and J. D. Louck. Necklaces, mss sequences, and dna sequences. *Advances in applied mathematics*, 18(1):18–32, 1997.

[5] D. L. De Ipiña, P. R. S. Mendonça, and A. Hopper. TRIP: A low-cost vision-based location system for ubiquitous computing. *Personal and Ubiquitous Computing*, 2002.

[6] C. Feng. *Camera Marker Networks for Pose Estimation and Scene Understanding in Construction Automation and Robotics*. PhD thesis, University of Michigan, 2015.

[7] M. Fiala. Artag, an improved marker system based on artoolkit. *National Research Council Canada, Publication Number: NRC*, 2004.

[8] R. Forster. Manchester encoding: opposing definitions resolved. *Engineering Science & Education Journal*, 9(6):278–280, 2000.

[9] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292, 2014.

[10] C. Hu, F. Arvin, C. Xiong, and S. Yue. A bio-inspired embedded vision system for autonomous micro-robots: the lgmd case. *IEEE Transactions on Cognitive and Developmental Systems*, PP(99):1–1, 2016.

[11] T. Krajník, M. Nitsche, J. Faigl, T. Duckett, M. Mejail, and L. Přeučil. External localization system for mobile robotics. *Proceedings of the 16th International Conference on Advanced Robotics (ICAR 2013)*, pages 1–6, 2013.

[12] T. Krajník, M. Nitsche, J. Faigl, P. Vaněk, M. Saska, L. Přeučil, T. Duckett, and M. Mejail. A practical multirobot localization system. *Journal of Intelligent & Robotic Systems*, 76(3-4):539–562, 2014.

[13] M. Kulich, J. Chudoba, K. Košnar, T. Krajník, J. Faigl, and L. Přeučil. Syrotek distance teaching of mobile robotics. *IEEE Transactions on Education*, 2013.

[14] D. Lehmer. On Euler's totient function. *Bulletin of the American Mathematical Society*, 38(10):745–751, 1932.

[15] P. Lightbody, C. Dondrup, and M. Hanheide. Make me a sandwich! intrinsic human identification from their course of action. 2015.

[16] R.-G. Mihalyi, K. Pathak, N. Vaskevicius, T. Fromm, and A. Birk. Robust 3d object modeling with a low-cost rgbd-sensor and ar-markers for applications with untrained end-users. *Robotics and Autonomous Systems*, 66:1–17, 2015.

[17] E. Olson. AprilTag: A robust and flexible visual fiducial system. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 3400–3407. IEEE, May 2011.

[18] J. Sawada. Generating Bracelets in Constant Amortized Time. *SIAM Journal on Computing*, 31(1):259, 2001.

[19] M. C. Silverman, D. Nies, B. Jung, and G. S. Sukhatme. Staying alive: a docking station for autonomous robot recharging. In *Robotics and Automation, 2002. Proceedings. ICRA '02. IEEE International Conference on*, volume 1, pages 1050–1055 vol.1, 2002.

[20] A. Trachtenbert. Computational methods in coding theory. Master's thesis, University of Illinois at Urbana-Champaign, 1996.

[21] Vicon. Vicon MX Systems.

[22] D. Wagner and D. Schmalstieg. Artoolkitplus for pose tracking on mobile devices. In *Proceedings of the 12th Computer Vision Winter Workshop (CVWW'07)*, February 2007.

[23] S. Yang, S. A. Scherer, and A. Zell. An onboard monocular vision system for autonomous takeoff, hovering and landing of a micro aerial vehicle. *Journal of Intelligent & Robotic Systems*, 69(1-4):499–515, 2013.