World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:18, No:4, 2024

Open Science Index, Electronics and Communication Engineering Vol:18, No:4, 2024 publications.waset.org/10013591.pdf

Tape-Shaped Multiscale Fiducial Marker: A Design Prototype for Indoor Localization

Marcell S. A. Martins, Benedito S. R. Neto, Gerson L. Serejo, Carlos G. R. Santos

Abstract—Indoor positioning systems use sensors such as Bluetooth, ZigBee, and Wi-Fi, as well as cameras for image capture, which can be fixed or mobile. These computer vision-based positioning approaches are low-cost to implement, mainly when it uses a mobile camera. The present study aims to create a design of a fiducial marker for a low-cost indoor localization system. The marker is tape-shaped to perform a continuous reading employing two detection algorithms, one for greater distances and another for smaller distances. Therefore, the location service is always operational, even with variations in capture distance. A minimal localization and reading algorithm was implemented for the proposed marker design, aiming to validate it. The accuracy tests consider readings varying the capture distance between [0.5, 10] meters, comparing the proposed marker with others. The tests showed that the proposed marker has a broader capture range than the ArUco and QRCode, maintaining the same size. Therefore, reducing the visual pollution and maximizing the tracking since the ambient can be covered entirely.

Keywords—Multiscale recognition, indoor localization, tape-shaped marker, Fiducial Marker.

I. INTRODUCTION

THE indoor positioning system is a technology that continuously estimates the position of objects or people indoors. Some technologies applied in positioning systems use sensors such as Bluetooth, ZigBee, and Wi-Fi. Other technologies use devices with cameras for image capture. These cameras can be fixed or mobile. An example of a mobile camera is the smartphone. The positioning approach based on computer vision technologies can be considered low-cost implementation — especially when it is not necessary to buy cameras to implement the system — providing reliable services in real-time.

Detecting fiducial markers occurs continuously to find a unique character, signal, or color in a target image, known as finder pattern. These markers are like visual clues for easy identification in places with a wide range of colors, lighting, and geometric shapes [1].

The literature shows that the fiducial markers are more efficient than natural markers due to their distinction of shape and color in the most diverse environments. In this perspective, many commercial indoor positioning products have adopted these artificial markers due to their static labeling form [2].

In the literature, a range of fiducial markers are designed for different purposes, such as ARToolKit [3], ARTag [4], reacTIVision [5], CALTag [6], AprilTag [7], RUNETag [8], ChromaTag [9], and TopoTag [10].

In an indoor localization system, the fiducial markers are positioned far from each other, as shown in the work of Muñoz et al. [11], generating a gap that can cause the discontinuity of reading and freezing the location service. If the markers are positioned linearly with minimum distance, it will allow a continuous reading of the environment, so there is always a marker in the image capture scene.

Therefore, the present study proposes a fiducial marker design that allows continuous reading, robust to changes in distance, with the possibility of low-cost implementation, for location in indoor environments. The reading of this marker can use two algorithms, one for long and one for short distances.

Consequently, this study aims to assist researchers in choosing the technology for a low-cost indoor location system to be applied in the most varied environments such as shopping malls, airports, subway stations, and others. An advantage of this study is the economic feasibility of its implementation since the applications developed can run on mobile devices. This scenario is more evident considering that, currently, there is an increase in the popularity of smartphone access (which can be carried by the user wherever he goes) in addition to their growing computing power and ubiquitous Internet connection.

In addition to this introductory section, Section II addresses the related works. Section III presents the design of the tape-shaped marker. Section IV reports the process of locating and reading the Marker. Section V presents the tests performed. Finally, Section VI presents the conclusions as well as future work.

II. RELATED WORKS

State of the art on fiducial markers has different characteristics, such as colors, shape, and numerical coding. These works show the evolution of characteristics over time such as the robustness of indoor localization systems based on computer vision.

Circular-shaped markers are available in the literature; Concentric Contrasting Circle (CCC), designed by [12], is a monochromatic marker applied to an autonomous system that recognizes the positioning of objects by a robot. Reference [13] proposes a polychromatic circular-shaped marker with various detection levels called multi-ring color for augmented reality applications. Both markers have in common the coding form generated from the features extracted from the circles.

For [14], one of the advantages of using markers with a circular shape in localization systems is their robustness to occlusion at different detection angles.

Marcell S. A. Martins, Benedito S. R. Neto, Gerson L. Serejo, and Carlos G. R. Santos* are with the Programa de Pós-Graduação em Ciência da Computação (PPGCC), Univeridade Federal do Pará (UFPA), Belém, Pará, Brazil (*corresponding author, e-mail: carlosresque@ufpa.br).

CyberCode is a square and monochrome visual marking system based on bar code technology proposed by [15] aimed at indoor location systems and augmented reality. The ChromaTag marker proposed by DeGol et al. [9] also has a square shape of polychromatic colors which has fast detection compared to AprilTag markers by [7]. Cromatag's numerical encoding form is based on extracting features from their geometric shape, unlike AprilTag. Wang and Ye's Maxicode [16] has hexagonal data cells, each with a binary value of 1 (black) or 0 (white), identical to the QRCode according to ISO 18004 [17], which uses binary encoding to detect its labels.

Liu [18] proposed a triangular-shaped marker. The marker has varied coloration and an inner circle that facilitates the object detection and tracking process, being considered by the authors with the best performance in detections through environmental lighting variations compared to the ARToolKit marker proposed by [3]. The way of encoding the marker occurs by the RGB values of the pixels and by extracting the triangular features.

Jurado-Rodriguez proposed a cube-shaped marker [19] as a new method for designing, detecting, and tracking customizable fiducial markers. This 3D marker has polychromatic coloring that allows its binary encoding. The marker can be used for internal guidance systems and augmented reality. This marker proved robust to different lighting intensities, with the ability to detect up to 80% of occlusion. The authors show that other proposals, such as DeepTag [20], have occlusion robustness of up to 30%, and the same occurs for TopoTag [10], ArUco [21], and AprilTag [7] when the occlusion is greater of 10%.

Costanza and Robinson [22] proposed markers with geometric format independence that Bencina et al. [5] applied in the reactTIVision system. These markers have a detection tolerant to severe distortions using adjacency trees of regions in their methodology. The marker design considers the number of nodes and adjacency levels of the object's centroids present in the marker.

Related works point to diverse fiducial markers for location applications in indoor environments. However, there is a need for a comprehensive review to verify the existence of a marker that can be continuously visible throughout the user's walk in an indoor environment, and that is robust to varying distances between the camera and the target using a multi-scale technique of detection.

From this perspective, a tape-shaped fiducial marker like the one proposed in this article has the advantage of having a sequential tape format, having a hierarchical pattern for multi-scalar detections, and can be fixed in places such as skirting boards or at the top of the wall, reducing visual pollution.

III. MARKER DESIGN

This section describes the proposed design of a tape-shaped marker that aims to map an internal environment. Therefore, it allows continuous reading along the path traveled by the user, even with the occlusion of some segment of the tape.

The proposed marker is composed of a sequence of Code Markers (CM) that can be read individually, especially when the camera is near the target. Furthermore, this CM sequence generates an encoding string that is called Tape Markers (TM). This sequence can be read, especially when the camera is far from the tape. The TM reading algorithm is similar to Standard Pattern [23] algorithm.

The marker structure were designed to cover the environment in areas such as the footboard or ceiling edges, allowing the user to move in all spaces without losing sight of the tape.

The monochromatic marker facilitates the detection and extraction of information, including conditions of varying lighting and low resolution. Fig. 1 shows that the QR Code inspired its Finder Pattern because its geometric shape is easy to detect. Alignment Patterns and Encoding Region are similar to CyberCode [15] for adjusting the ideal focal length in the recognition and extraction of target features.

A. Code Marker

Fig. 1 (a) shows that each CM has squares arranged in a regular rectangular matrix containing the Finder Pattern, Alignment Patterns, and the Encoding Region. Furthermore, white squares that surround this entire coding region are called Quiet Zone.

Finder Patterns are specific position detection marks located at the marker ends (left and right) of each CM. They consist of a 7×7 black square overlapped by a 5×5 white square and overlapped by a 3×3 black square in the center.

Alignment Patterns are the black squares that form an ‘L’ in CM, with a black square in the upper right region, providing the marker orientation in scenes.

The Encoding Region contains black and white squares encoding the code itself. The overall CM can be positive, as shown in Fig. 1 (a), when the background is white and the bits are black. Alternatively, CM can be negative, as shown in Fig. 1 (b), when the background is black and the bits are white.

Quiet Zone is a padding that does not contain data and ensures that the CM's neighborhood does not disorient the extraction of code.

B. Tape Marker

A TM is a segment containing a sequence of positive or negative CMs. The arrangement of these CMs forms the TM code, which can be read similarly to a bar code.

The TM contains four parallel vertical bars delimiting the beginning and end of the reading. The CMs (negative or positive) between these vertical bars become the TM code. The Finder Pattern of the CMs present along the TM must be ignored in its reading.

CM quantity varies within the tape according to the desired number of bits and the physical length readability of the tape. Because the more bits (CMs) present in a TM, the longer this tape segment becomes. Hence, the longer the segment is, the further away the user needs to be from the target, and the aperture angle of his camera can frame an entire segment.

On the other hand, if a few bits are used, there will not be enough code to cover all the space, depending on the available

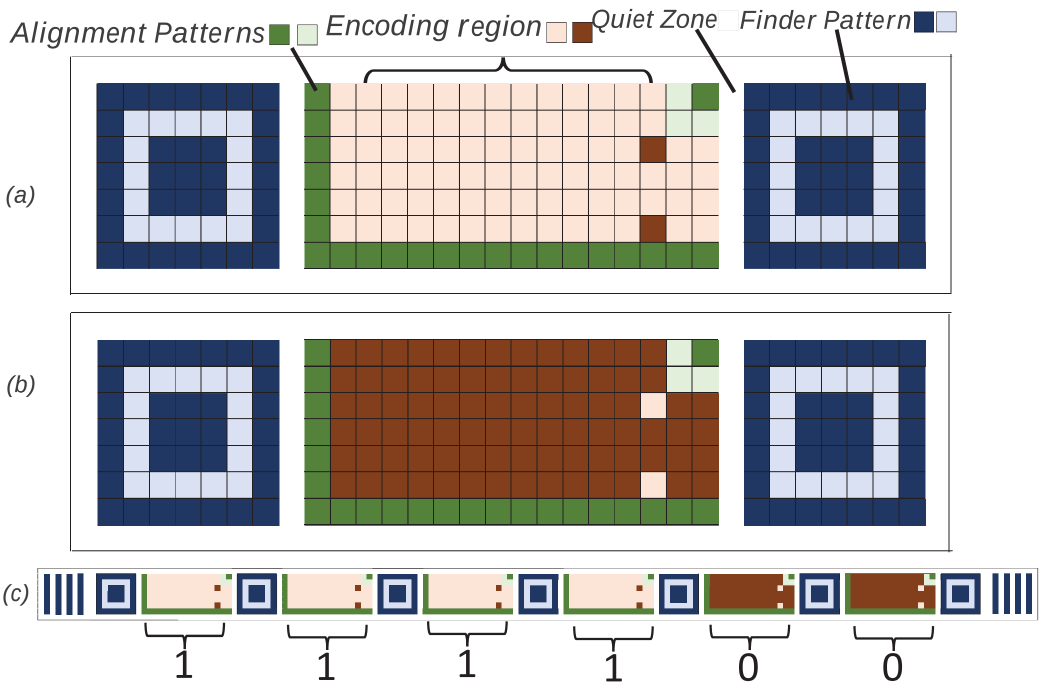

Fig. 1 The marker's design: (a) encoding for CM positive values; (b) encoding for CM negative values; (c) encoding for TM larger-scale detection

length. For example, using only 3 CMs per TM, there were only eight different codes to map the space.

C. Multiscale Functionality

This section covers the multiscale reading feature. The tape was designed to be robust to changes in distance between the camera and the target. Hence, the user can freely move in space without losing a target code.

Fig. 2 shows the flowchart showing multiscale tape detection. This flowchart starts with the frame capture function that will read an image. Then, the system starts the detection process of TMs, displaying the IDs. If TM detection fails, the system starts the detection of CMs, showing the IDs of detected ones.

D. CM Binary Encoding

The numeric codes of CMs should use as slight bit variation of tint (bits 1) as possible for near numbers so that when viewed from a long distance, it can be recognized as a single bit, with little ink for positive and much ink for negative. Therefore, the default binary encoding, the Weighted Binary Code (WBC), could not be used since the number of bits can vary significantly between nearby numbers. For instance, the binary version of 128 has one 1 and seven 0's, whereas the number 127 has seven 1's and one 0, considering a binary code with 8 bits of length.

Hence, a binary encoding that the first numbers are all combinations of a bit 1 and other bits 0, followed by all combinations of two 1's and the remaining 0's, and so on, was designed. Fig. 3 shows an example of a binary encoding of length 6 using both encodings. The figure shows that in the proposed encoding, the number of bits 1 grows in a stable and ordered way, and the number of bits grows or changes in an unordered way for the same sequence in WBC encoding.

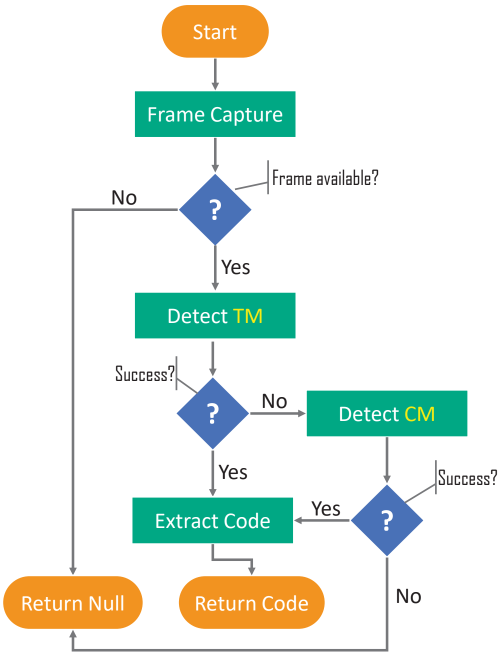

Fig. 2 Flowchart of the multiscale detection process

The green highlights show that the proposed code enables the usage of numbers from 0 to 21 for less ink and from 42 to 63 for more ink (those ranges grow along with the code's length). Additionally, the red arrow shows some problems in the same ranges when using the WBC encoding.

In this way, it is possible to create a sequence of codes that

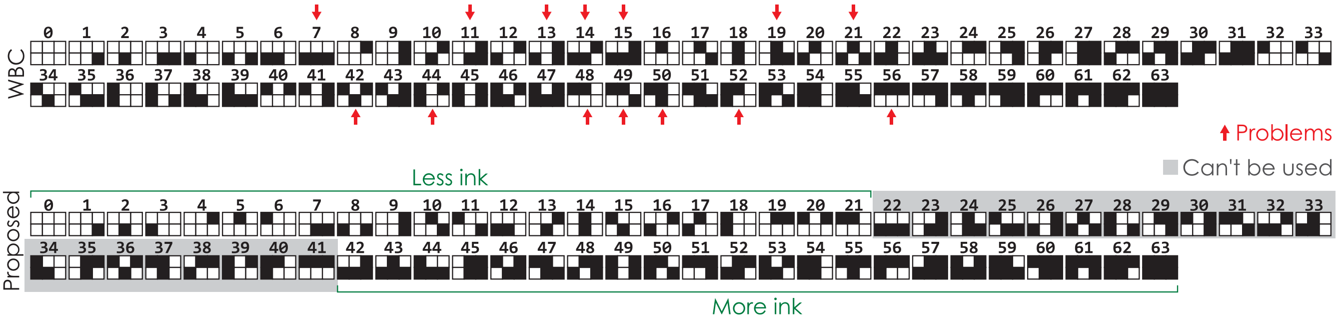

Fig. 3 Binary encoding proposed to CM's

uses the smallest number of bits 1 possible, thus efficiently controlling the amount of ink in the available area of the marker. This feature is relevant because a TM is composed of CMs that can be positive (white majority) or negative (black majority) that will be distinguished depending on the amount of ink used in these CMs.

The conversion from decimal to the proposed encoding starts by verifying how many 1s bits are needed to represent the decimal value using the combinatorial formula. Then, all needed 1s are on the right and the remaining 0 on the left, forming an initial code. The decimal value of the initial code is the sum of all combinatorics with fewer bits 1. According to the Algorithm 1, the initial code and the decimal value increment +1 until the value is the desired one.

Algorithm 1 Increment Code

```
Require: $bincodestr$
Ensure: $bincodestr + 1$                    ▷ The input incremented
$p \leftarrow$ position of the first bit 1 that has 0 in left
$aux \leftarrow bincodestr[p]$
$bincodestr[p] \leftarrow bincodestr[p-1]$
$bincodestr[p-1] \leftarrow aux$
if $bincodestr[p-2] = 1$ then               ▷ Verify out-of-bounds
    $left1 \leftarrow$ position of first bit 1 in the left of $p-2$
    Remove all bits 0 between $p-2$ and $left1$
    Prepend all 0 removed to $bincodestr$
end if
```

E. Environment Setup

The user can choose the tape height, and the program calculates the tape width based on the aspect ratio of a CM, which is 4.57 (or 32:7). To find the width of a CM, multiply the desired height by 4.57. For example, if the user wants a 6cm ribbon, the width will be 27.4cm. Considering that in the concatenation of the CMs to form the tape, a Finder Pattern is overlapped, the aspect ratio of 25:7 was used to calculate the length of the tape.

Then, having the width of a CM, the program can calculate the amount of CM that will fit in the available space, dividing the space by the width of each CM. The formula below shows how to do this calculation.

$$CM_{Qty} = \left\lfloor \frac{100i}{4.57t_h} \right\rfloor$$

Where CM_{Qty} is the quantity of CMs that fills the available space, i is the input available space in meters, and th is the desired tape height.

IV. MARKER LOCATION AND READING

Indoor location applications could use fiducial markers to track the user position [8]. Therefore, the tape-shaped marker proposed in this study tracks user position within an indoor environment. This marker can cover the environment as shown in Fig. 4, which shows an image of a floor plan consisting of three rooms and one corridor. The red line represents the strip with its respective CMs that will compose the TMs represented by blue brackets.

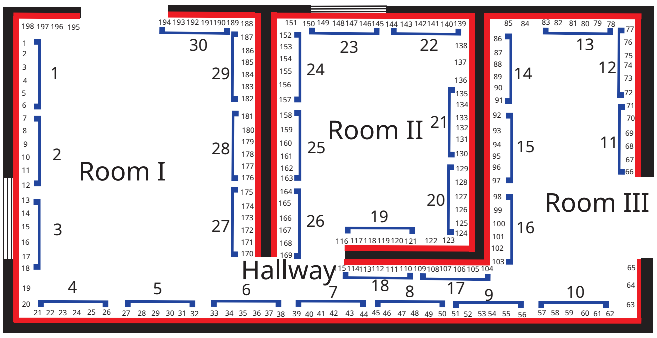

Fig. 4 Example of a floor plan containing the tape (in red) for each room with their respective CM codes. The blue square brackets show the TMs formed by a set of 6 CMs

Each TM (blue brackets in Fig. 4) is indexed to an environment position so that when detecting and reading a TM, its extracted code references a location in the environment. When approaching the camera to the TM region, the program will do a new reading, this time looking for a CM that again represents an environment location, making the location process occur at different distances from the tape.

A. Locating and Reading the TM

The TM reading relies on the number of black pixels displayed on the tape. For the reading to be possible, the user must be at a certain distance that frames an entire segment in the image. Four parallel bars indicate the beginning and end of the segment. The following steps describe a simple algorithm to locate and read the TM.

Open Science Index, Electronics and Communication Engineering Vol:18, No:4, 2024 publications.waset.org/10013591.pdf

1) Detect the horizontal lines contained in the image, selecting only the most representative line;
2) Rotate the image to be parallel to the abscissa axis through the angular adjustments according to the line detected in the previous step, leaving the detected line horizontal and cutting the image longitudinally;
3) Apply grayscale, contrast, smoothing, and threshold filters;
4) For each image's horizontal line, check the presence of patterns that indicate the TM beginning and end. When detecting two patterns, convert the line to string;
5) Bit string simplification into unit values;
6) Store the simplified bits for voting in an array;
7) Apply polling of detected bit strings by selecting the string with the highest frequency. If the confidence is greater than or equal to 50%, this string is returned, otherwise the algorithm returns null;

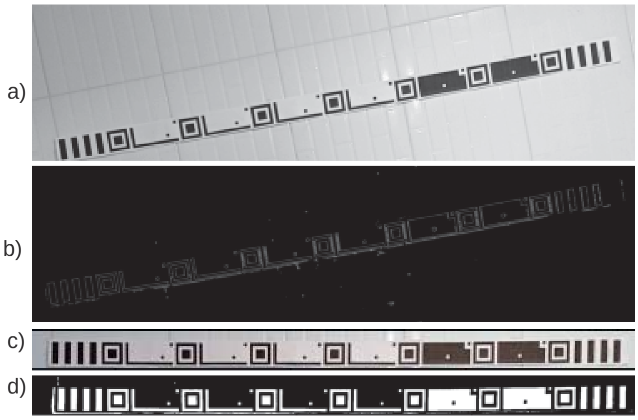

Fig. 5 (a) Grayscale image; (b) Canny filter; (c) Image cropped and rotated 180 degrees; (d) Binarized image

Line detection starts with transforming the color image to grayscale, as shown in Fig. 5 (a) and edge detection by applying the Canny filter with minimal value 230 and maximal value 255, as shown in Fig. 5 (b). Then, the Hough Transform method [24] detects lines with parameters rho=1, $theta = \pi/180$ and $threshold > 240$, as shown in Fig. 5 (c), which allows the rotation and clipping of its surrounding area. This area could contain a TM. Finally, the image is binarized to extract the bits with the horizontal scan, as shown in Fig. 5 (d).

Successful horizontal scans of the binary image participate in voting to ensure decoding reliability. For instance, if the result found "10101011" seven times and another three different results. The decoder provides the result "10101011" with reliability of 0.7, using a confidence of $C \geq 0.5$. When reliability is below 0.5, the algorithm returns null for this iteration.

B. Locating and Reading the CM

The CM reading occurs when the detection of TM fails. The process starts with detecting the Finder Patterns, as shown in Fig. 7 (a), transforming the coding region perspective, as shown in Fig. 7 (b), and extracting the bits from each CM on the tape, as shown in Fig. 7 (c). The following steps describe the process.

1) Applying pre-processing image filters: grayscale conversion, contrast enhancement, and threshold filters.
2) Extraction of the square contours that delimit the region of each marker present on the tape, also extracting points corresponding to the markers' region. Otherwise, the process flow goes back to the beginning.
3) Correction of the image's perspective to represent the actual marker aspect ratio.
4) Reading the grid pixels of the marker's coding region to divide the image into cells. Each grid cell will correspond to 1 for pixels values greater than 127 and 0 for smaller values.
5) Transforming the output of the resulting bits converting it into a bits vector, selecting only the bits on the Encoding Region. The process repeats until the algorithm extracts the last pair of Finder Patterns.
6) Return the detected bits if there is no contour in the processing queue.

The CMs reading is successful when the perspective correction allows the Alignment Patterns to fit correctly in the reading grid, enabling the consistent extraction of bits in the Encoding Region. Otherwise, the CM read fails, and the algorithm proceeds to read the next detected CM.

V. EXPERIMENTAL TEST

The test consists of detecting markers at different distances using the TM detection algorithm, shown in Fig. 6 (a), detection of CM, as shown in Fig. 6 (b), detection of ArUco, as shown in Fig. 6 (c) and detection of QRCode, as shown in Fig. 6 (d).

A. Test Apparatus

The smartphone devices Motorola Moto G8 Power Lite, Motorola G6 Play and Xiaomi redmi note 9 were used to capture the image from the tape. The Motorola Moto G8 Power Lite with Android 10 Operating System, has triple camera systems that involve multiple image sensors, each with its own lens, with an aperture of F2.0 and a 16-megapixel main camera that allows you to take pictures with a resolution of 4619x3464 pixels. The Motorola G6 Play with Android 8 operating system, has a main camera system of 13 megapixels 4128x3096 resolution with F2.0 camera aperture. The Xiomi Redmi Note 9 is a smartphone with Android 10 operating system, has a main camera of 48 megapixels with a resolution of 8000x6000 pixels with a camera aperture of F1.79.

B. Dataset Setup

CMs with IDs 19, 20, 21, 22, -23, -24 were generated to start the tests. This sequence of CMs generates the TM with code '111100', which will be linked to the ID number 500 that will be shown when the TM is read (Fig. 6 (a)). After that, the tape was printed on glossy photographic paper with size 160.8×6 cm.

(a) TM detection (b) CM detection

(c) ArUco detection (d) QRCODE

Fig. 6 Tape marker detection tests

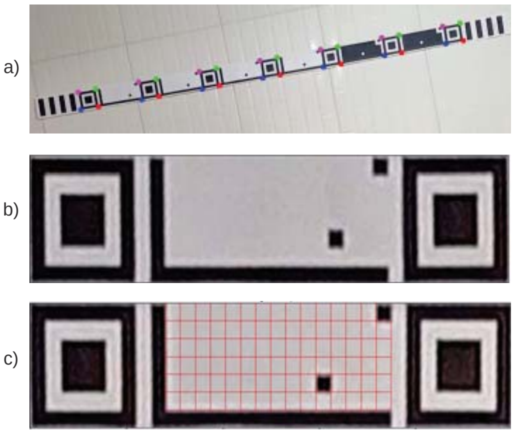

Fig. 7 (a) Detection of finder patterns; (b) image Warp; (c) Grid on encoding region

ArUco and QRCode were used to compare the accuracy of the proposed marker with two markers available in the market. ArUco features its black background and white bars and has occlusion and lighting variations robustness. The QRCode has a white background and black bars, and it can be encoded in its body with a large volume of bytes, in addition to being sensitive to lighting and occlusion. For testing, the proposed tapes have the same amounts of CMs present in the TMs with dimensions 6×6, 3×3, and 1.5×1.5 cm. After printing, the tapes were placed on a baseboard to capture images ranging from 0.5 m to 10 m with size 1600×1200 pixels. For each distance, six images were taken from various angles, thus creating a dataset (Fig. 8).

C. Accuracy Test

the accuracy is calculated by $S/(S+F)$, where the Success S occurs when the marker is detected correctly. The Failure F occurs when the algorithm does not correctly identify IDs or does not find the marker.

The test employed the two algorithms, variating the distance every 0.5 meters. Fig. 9 shows the TM (blue line) and CM (red line) reading accuracy. CM reading performs better than TM when the distance is short. Between 2 and 6 meters, the accuracy is evenly approximate, variating in an unordered manner. When the distance is greater than 6 meters, the TM reading performs better than CM.

The accuracy of QRcode shown in orange line (Fig. 9) stopped readings at the distance of 3.5m, the ArUco shown in Green line (Fig. 9) was limited at 5.5m and the CM's (Fig. 9) at 7.0m. The TM's cannot be detected at a distance of 0.5 to 1.5 m because the camera cannot completely frame the TM dimension. Then, the detections were complemented by the CM's detections for these distances. Even with variations in TM accuracy, it proved effective for a range of distances long than 1.5m, where it was possible to frame the camera and continue the detection service.

The accuracy tests with the proposed marker (Fig. 9) show its robustness for variations of distances in an indoor environment.

Fig. 10 shows area that markers occupy in each captured scene for each distance in range from 0.5 to 10 m. In this picture it is possible to understand the size of the occupied area of each marker in the scene, showing that the CM area is smaller than the area occupied by the QRCode and ArUco. In the use of the tape proposed in this study minimizes the use of a sequence of markers with dimensions 25x25 cm, the size that each marker needs to have to be detected at 10 m distance

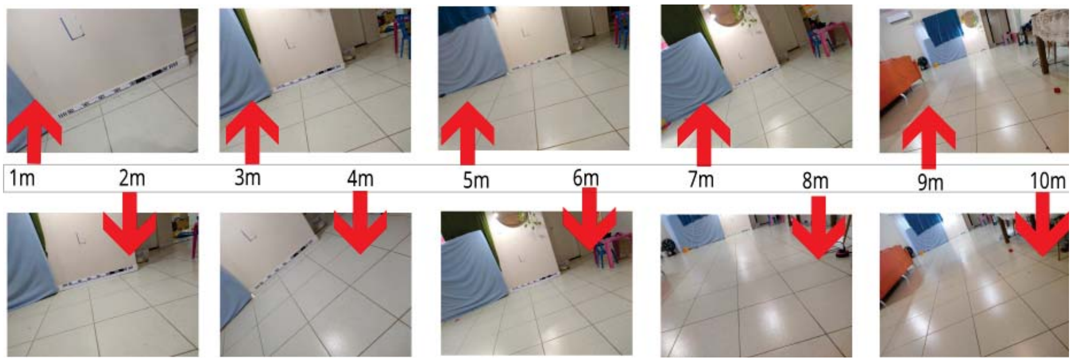

Fig. 8 A sample of the dataset at each distance

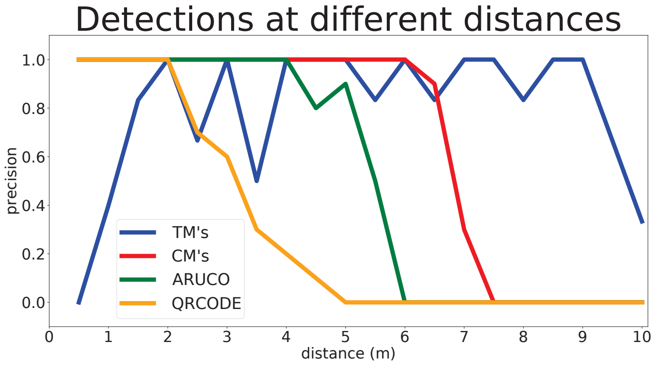

Fig. 9 Reading accuracy results in distances ranging from 0.5 to 10 m

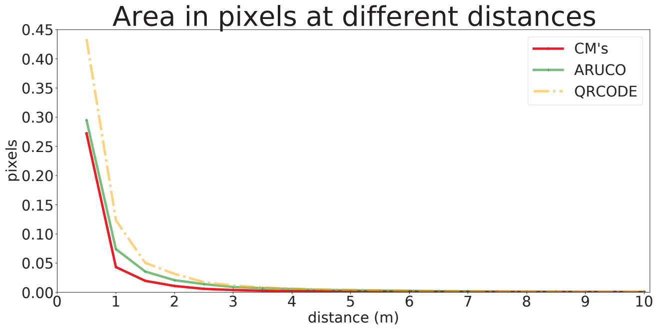

Fig. 10 The marker size and image size ratio ranging from 0.5 to 10 m

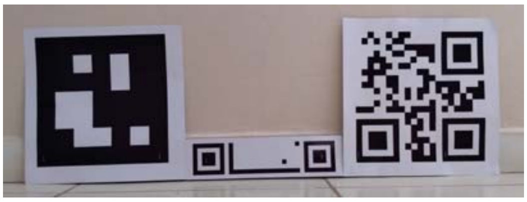

Fig. 11 Markers with sizes to be detected at 10 m distance

(Fig. 11). Therefore, the sequence of markers it would have to have an approximate area of 3750 cm² that will occupy very large areas to maintain the efficient detection service for a distance at 10 m, different from the TM that has an area 946.8 cm² to be detected at the both distances.

VI. CONCLUSION

Indoor position systems based on computer vision that use artificial markers tend to be more efficient than systems that use real markers. In addition, its implementation cost tends to be lower when compared to other systems that use radiofrequency, ultrasound, or other sensors. However, a fiducial marker's disadvantage is the relationship between aesthetics and the size of the area occupied by them because the smaller the distance between the markers (checkpoints) in the environment, the more significant the visual pollution. Also, reducing the distances between them will allow a continuous reading because the same camera frame can have more than one marker.

The tests showed that the location of CM's tape segments and TM full tape worked successfully for distances up to 10 meters. For short distances, like 0.5 meters, the TM's reading performed worse because the camera could not frame it entirely. On the other hand, when the TM is completely present in the camera frame, locating and reading the TM tends to perform better for longer distances. Therefore, the system switches between two reading algorithms for short and long distances, ensuring tracking and read reliability through the feasibility of the multiple scale feature.

The marker can be read using any camera, whether fixed or mobile. This equipment portability guarantees that the system is flexible to the hardware, which makes it possible to use it for different localization activities.

This work presented a fiducial marker with the following characteristics: 1. Continuous reading: The user will be able to always read the marker in the environment since the marker will be present in the scene, even with some occlusions. 2. Scalable: The marker can be read at various distances, allowing traceability without loss of reading in the environment. 3. Robust to partial occlusion: Even if one of the CMs is occluded, it will be possible to read the marker because the algorithm reads the next CM present in the scene, maintaining traceability in the environment. 4. Linear morphology: the marker has a tape format, allowing reading around the perimeter of the scenario.

For future work, improvements in the performance of the TM and CM reading algorithms were suggested, as well as the possibility that the marker contains a validity test similar to the QRCode. This study also highlights the low cost for applicability of this technology to the most varied indoor environments such as shopping malls, airports, and subway stations. In addition to the possibility of positioning it on a baseboard or on the edge of the ceiling, ensuring aesthetic standards, mitigating visual pollution in environments, and allowing traceability of the marker.

ACKNOWLEDGMENT

The authors would like to thank to the Programa de Pós-Graduação em Ciência da Computação (PPGCC) of the Universidade Federal do Pará (UFPA) for providing the facilities for the conduction of the experiments and data analysis. PPGCC is financially supported by the Coordenacão de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERENCES

[1] M. Košt'ák and A. Slabỳ, "Designing a simple fiducial marker for localization in spatial scenes using neural networks," *Sensors*, vol. 21, no. 16, p. 5407, 2021.

[2] H. Lim and Y. S. Lee, "Real-time single camera slam using fiducial markers," in *2009 ICCAS-SICE*. IEEE, 2009, pp. 177–182.

[3] H. Kato and M. Billinghurst, "Marker tracking and hmd calibration for a video-based augmented reality conferencing system," in *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)*. IEEE, 1999, pp. 85–94.

[4] M. Fiala, "Artag, a fiducial marker system using digital techniques," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2. IEEE, 2005, pp. 590–596.

[5] R. Bencina, M. Kaltenbrunner, and S. Jorda, "Improved topological fiducial tracking in the reactivision system," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*. IEEE, 2005, pp. 99–99.

[6] B. Atcheson, F. Heide, and W. Heidrich, "Caltag: High precision fiducial markers for camera calibration." in *VMV*, vol. 10, 2010, pp. 41–48.

[7] E. Olson, "Apriltag: A robust and flexible visual fiducial system," in *2011 IEEE international conference on robotics and automation*. IEEE, 2011, pp. 3400–3407.

[8] F. Bergamasco, A. Albarelli, E. Rodola, and A. Torsello, "Rune-tag: A high accuracy fiducial marker with strong occlusion resilience," in *CVPR 2011*. IEEE, 2011, pp. 113–120.

[9] J. DeGol, T. Bretl, and D. Hoiem, "Chromatag: A colored marker and fast detection algorithm," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 1472–1481.

[10] G. Yu, Y. Hu, and J. Dai, "Topotag: A robust and scalable topological fiducial marker system," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 9, pp. 3769–3780, 2020.

[11] R. Muñoz-Salinas, M. J. Marín-Jimenez, E. Yeguas-Bolivar, and R. Medina-Carnicer, "Mapping and localization from planar markers," *Pattern Recognition*, vol. 73, pp. 158–171, 2018.

[12] L. B. Gatrell, W. A. Hoff, and C. W. Sklair, "Robust image features: Concentric contrasting circles and their image extraction," in *Cooperative Intelligent Robotics in Space II*, vol. 1612. SPIE, 1992, pp. 235–244.

[13] Y. Cho, J. Lee, and U. Neumann, "A multi-ring color fiducial system and an intensity-invariant detection method for scalable fiducial-tracking augmented reality," in *In IWAR*. Citeseer, 1998.

[14] C.-C. Lin and R. L. Tummala, "Mobile robot navigation using artificial landmarks," *Journal of Robotic Systems*, vol. 14, no. 2, pp. 93–106, 1997.

[15] J. Rekimoto and Y. Ayatsuka, "Cybercode: designing augmented reality environments with visual tags," in *Proceedings of DARE 2000 on Designing augmented reality environments*, 2000, pp. 1–10.

[16] Y. P. Wang and A. Ye, "Maxicode data extraction using spatial domain features," Jun. 10 1997, uS Patent 5,637,849.

[17] B. ISO, "Iec 16022: information technology-automatic identification and data capture techniques-data matrix bar code symbology specification," *BS ISO/IEC*, vol. 16022, 2006.

[18] J. Liu, S. Chen, H. Sun, Y. Qin, and X. Wang, "Real time tracking method by using color markers," in *2013 International Conference on Virtual Reality and Visualization*. IEEE, 2013, pp. 106–111.

[19] D. Jurado-Rodríguez, R. Muñoz-Salinas, S. Garrido-Jurado, and R. Medina-Carnicer, "Design, detection, and tracking of customized fiducial markers," *IEEE Access*, vol. 9, pp. 140 066–140 078, 2021.

[20] Z. Zhang, Y. Hu, G. Yu, and J. Dai, "Deeptag: A general framework for fiducial marker design and detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[21] 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*, vol. 47, no. 6, pp. 2280–2292, 2014.

[22] E. Costanza and J. Robinson, "A region adjacency tree approach to the detection and design of fiducials." 2003.

[23] M. Kabuka and A. Arenas, "Position verification of a mobile robot using standard pattern," *IEEE Journal on Robotics and Automation*, vol. 3, no. 6, pp. 505–516, 1987.

[24] V. F. Leavers, *Shape detection in computer vision using the Hough transform*. Springer, 1992, vol. 1.

Marcell Serra de Almeida Martins Graduated in the Computer Science course at the University of Amazonia. Research interest related to the areas of information and visualization, virtual and augmented reality, and human-computer interaction.

Benedito S.R. Neto received a master's degree. from the Federal University of Pará, in 2014. He is currently an EBTT professor at the Federal Institute of Education, Science and Technology of Pará-Campus Cametá. His research interests are Neural Networks, Fuzzy Systems, Geospatial Modeling, Landscape Ecology, Remote Sensing, Image Processing, Geoprocessing, Human-Computer Interaction and Gamification in Teaching.

Gerson L. Serejo Master in Computer Science (2018) in the Graduate Program in Applied Computing (PPCA) of the Nucleus for Amazonian Development in Engineering (NDAE), Federal University of Pará (UFPA), University Campus of Tucuruí; degree in Control and Automation Engineering from faculty Estácio of Pará and graduation in Data Processing from Center University of Pará. I am currently a PhD student at the Graduate Program in Computer Science (PPGCC) at the Federal University of Pará (UFPA).

Carlos G. R. Santos received the master's and Ph.D. degrees from the Federal University of Pará, in 2015 and 2017, respectively. He is currently a Tenured Professor with Federal University of Pará. His research interests are in information and scientific visualization, virtual and augmented reality, and human computer interaction.