ONLINE FIRST July 1, 2025

Engineering Transactions, **73**, https://doi.org/10.24423/engtrans.3499.2025
Institute of Fundamental Technological Research Polish Academy of Sciences (IPPT PAN)
Université de Lorraine • Poznan University of Technology

SiMRTRANS

A Concept of a Vision System for Position Determination Using Spatial-Time-Frequency Analysis

Piotr MIŚ*, Przemysław SZULIM, Jędrzej MĄCZAK

Warsaw University of Technology
Faculty of Automotive and Construction Machinery Engineering
Warsaw, Poland; e-mails: przemyslaw.szulim@pw.edu.pl, jedrzej.maczak@pw.edu.pl

*Corresponding Author e-mail: piotr.mis.dokt@pw.edu.pl

This paper presents the concept of a vision system designed to determine the position of a camera within a given coordinate system. The system focuses on identifying pulsating light markers in images recorded by the camera. These markers are characterized by specific colors, pulsation frequencies and known locations. Detection is achieved using a spatial-time-frequency processing method developed by the authors. The identified markers serve as input data for a mathematical model that determines the camera's position within the reference system. The article discusses the theoretical foundations of the proposed system design. The system was subjected to testing to verify its operational accuracy and precision in position determination. The results of these tests are presented. The article concludes with a summary of the work and a discussion of the system's further development and practical applications.

Keywords: image processing; spectrum analysis; light markers; localization.

1. Introduction

The topic of localization remains a significant area of research, driven not only by the need to improve the estimation quality of key parameters such as position and angular orientation, but also by the necessity to ensure redundancy or supplement existing systems that may malfunction under certain conditions. Vision-based localization systems have been developed for a long time [1–3]. Ongoing efforts are made to both improve positioning accuracy and ensure operational continuity by utilizing visual images for localization tasks. These systems are particularly valuable in environments where radio-based technologies (e.g., global navigation satellite systems (GNSS) and ultra-wideband (UWB)) may encounter problems such as disruptions.

The concept of the vision system discussed in this work involves determining the position of an object using pulsing light markers placed at known locations

and a camera capturing their images. The vision system can determine the position of the camera (or the object to which it is mounted) in space based on the recorded images of the markers.

Markers are objects with distinctive features, or a set of features, that make them recognizable. The use of markers in vision systems is not a novel concept and has been extensively discussed in numerous studies [4–8]. Markers can be classified as either passive (e.g., a simple QR code placed at a known coordinate) or active (e.g., light-emitting markers). Using light markers is advantageous because they can be not only objects with known positions but also transmit information using light signals. Over the past few years, there has been a growing interest in the use of light markers for data transmission, as existing infrastructure (e.g., traffic signals [1], streetlights, or vehicle headlights [9]) can be utilized to build a vision-based systems.

It is also worth noting that these solutions are also gaining attention due to their resilience against accidental or intentional interference, a common weakness of radio-based techniques.

Following the studies [3, 10–13], systems utilizing light markers can be divided into two groups:

- **VLC** (**visible light communication**): communication systems focused on detecting light marker signals, interpreting them, and extracting the embedded information,
- **VLP** (**visible light positioning**): positioning systems that use light markers to determine the object's position. These systems rely on detecting the presence of an object at a specific location to ascertain the position of the receiver (camera or photodiode).

Recently, there has been noticeable development in VLP systems employing light markers controlled by signals modulated using pulse position modulation (PPM) [14–18]. It is possible to capture streaks in recorded images by appropriately adjusting the camera's exposure time, which uniquely identify a specific light marker. An example of such a method is presented in [14], which describes an indoor positioning system based on images captured by a smartphone camera. The system can accurately determine the camera's position relative to the marker by analyzing the visible light stripes in the image, where each marker has a unique stripe pattern generated by using a signal modulated with PPM. In [15], the authors proposed an enhanced mechanism for reading PPM signals between an LED and a camera. This mechanism should increase data transmission efficiency while maintaining position detection. In this approach, LED markers are also identified using a unique stripe sequences. Additionally, the marker recognition algorithm can read information encoded using these stripe sequences. The studies presented in [16–18] focus on more advanced developments of PPM modulation. In [16], the behavior of the multiple pulse position modulation (MPPM)

signal is analyzed in terms of its spectral representation, allowing better adjustment of transmission parameters for varying system conditions. In [17], the authors investigate the effect of brightness control and the application of multi-heralded hybrid pulse position modulation (MH-HPPM) on the signal's spectral structure. This is important for designing systems that are resilient to interference. Meanwhile, [18] presents an analysis of spatial variant of PPM, which allows for better differentiation of markers within environments with limited operational space.

The authors of this paper present a concept for a system that uses pulsating light markers with predefined spatial positions and known pulsing frequencies. These markers are recorded by a camera. A distinguishing feature of the presented system is its unique approach to image processing. Specifically, a sequence of recorded two-dimensional frames (captured images) is transformed into a collection of signals representing pixel intensity level changes over time. This concept determines the operation of the spatial-time-frequency algorithm, developed in [19, 20], which will be discussed in detail in Sec. 2.

The presented system belongs to the VLP category and uses light markers, but focuses on different aspects compared to the previously cited works. In [14–18], the authors analyze two-dimensional signals (images) in order to detect stripe patterns that characterize specific light markers. In contrast, the approach proposed in this paper employs a spatial-time-frequency method that processes a sequence of recorded images into a set of signals representing pixel intensity level changes over time. This enables to focus on the analysis of one-dimensional signals and the detection of light markers based on their unique pulsation characteristics.

In previous works, authors studied the properties of the spatial-time-frequency algorithm and the selection of its optimal operating parameters. Article [19] focuses on the description and main assumptions underlying spatial-time-frequency processing, including studies that guided the selection of marker colors and pulsation frequencies. In a subsequent article [20], the research on the influence of the recording camera's exposure time on the obtained binary images was presented. However, the present article does not focus on analyzing the algorithm itself; instead it presents its specific application to a localization task. This work is experimental in nature and is intended to show the practical operation of the developed method.

Section 3 presents the mathematical assumptions that enable the determination of the camera's position using identified pulsating light markers. Section 4 presents the concept of the vision system and its operational scheme. The subsequent two sections describe the algorithm testing and the verification of positioning accuracy of the determined position in both stationary and moving cases. Section 7 summarizes the key findings of the conducted work, while Sec. 8 provides conclusions and directions for future development.

2. Spatial-time-frequency analysis

The spatial-time-frequency algorithm is designed to identify pixels in recorded images that capture the pulsating light emitted by markers. A detailed description of the algorithm and related studies were presented by the authors in [19, 20].

This spatial-time-frequency algorithm focuses on analyzing intensity level changes in pixel intensity over time. The recorded sequence of images must first be appropriately prepared to use this algorithm. For each pixel, signals representing the intensity level changes are generated across consecutive frames. These signals are then subjected to time-frequency processing to locate the pixels where the pulsation frequency of a given marker is recorded. Time-frequency methods help reduce averaging effects and allow estimation of the temporal instances at which frequency amplification occurs for specific pixels. The operational scheme of the algorithm is illustrated in Fig. 1.

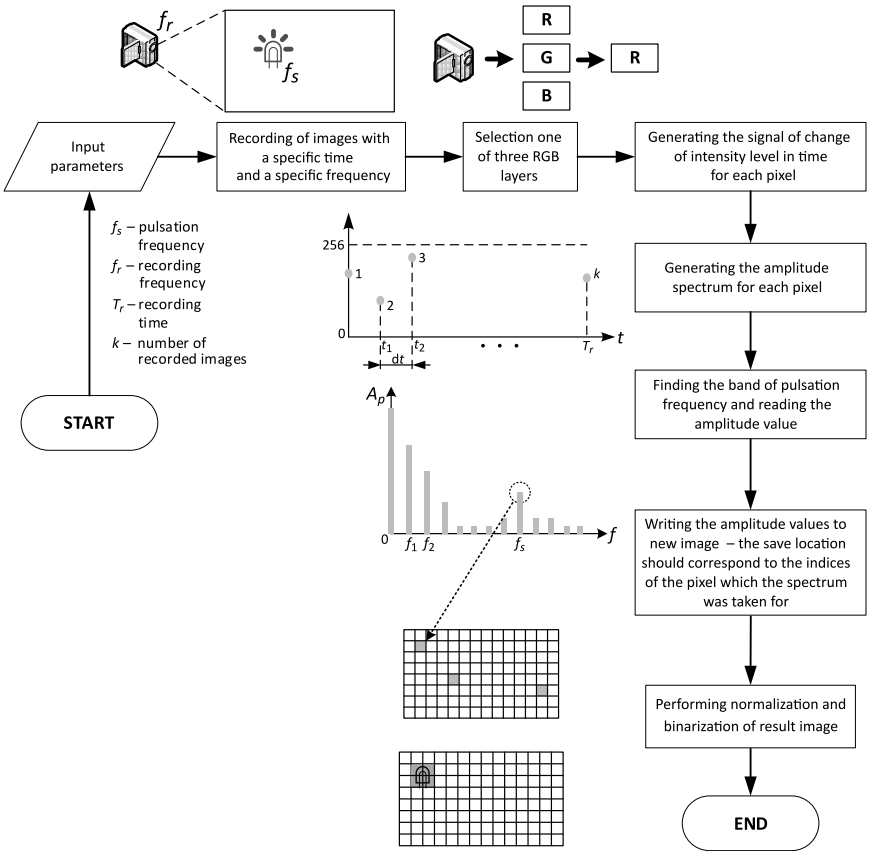

Fig. 1. The block diagram showing the individual steps of the operation of the spatial-time-frequency processing algorithm [20].

The output of the algorithm is a binary image, where pixels with a value of one indicate the presence of a detected marker. Image processing operations like erosion and dilation are applied to avoid errors and misinterpretations caused by phenomena such as light reflections from shiny surfaces. Additionally, the shape of the detected pixels is verified for its similarity to a circle. The closer the shape resembles a circle, the higher the probability that the marker has been correctly identified. Approximating the detected area with a circle also facilitates determining the center of the circle in the image.

It is worth noting that the algorithm can differentiate light markers based on their color and pulsation frequency. Marker recognition is achieved by analyzing the amplitude of the pulsation frequency of the markers used and by comparing the average brightness levels of individual pixels across the RGB layers. The practical implementation of this description will be demonstrated with examples in Sec. 5 focusing on experimental studies.

3. Calculation of position

The detection of markers in images enables the determination of the camera's position within a given reference frame. Mathematical relationships associated with the use of the camera lens and the projection of light rays through the lens onto the sensor are used. The idea of using such relationships is frequently encountered in the literature [7, 21, 22]. The mathematical model employed by the authors is based on the concept presented in [23].

Three distinct light markers with known coordinates in the given reference frame are used: $(x_1,\ y_1,\ z_1)$, $(x_2,\ y_2,\ z_2)$, $(x_3,\ y_3,\ z_3)$ in the model proposed in [23]. Assuming that each marker represents the center of a sphere, and the camera position $(x,\ y,\ z)$ is assumed to be located on the surface of each sphere, the following equations can be written:

$$\begin{aligned} (x-x_1)^2+(y-y_1)^2+(z-z_1)^2 &= d_1^2,\\ (x-x_2)^2+(y-y_2)^2+(z-z_2)^2 &= d_2^2,\\ (x-x_3)^2+(y-y_3)^2+(z-z_3)^2 &= d_3^2. \end{aligned} \tag{3.1}$$

It is assumed that one coordinate is identical across all markers. This can be realized, for instance, by mounting the markers on the ceiling of the room where localization takes place. Therefore, the equations can be simplified as follows:

$$z_1 = z_2 = z_3.$$

The height difference between the marker and the camera can be expressed as:

$$z_1 - z = H.$$

Substituting it into Eqs. (3.1) yields the following form:

$$\begin{aligned}(x-x_1)^2+(y-y_1)^2+H^2&=d_1^2,\\ (x-x_2)^2+(y-y_2)^2+H^2&=d_2^2,\\ (x-x_3)^2+(y-y_3)^2+H^2&=d_3^2,\end{aligned}\tag{3.2}$$

where $(x,\ y,\ H)$ are the coordinates of the camera to be determined. Further considerations rely on the mathematical relationships related to image formation on the sensor, as illustrated in Fig. 2.

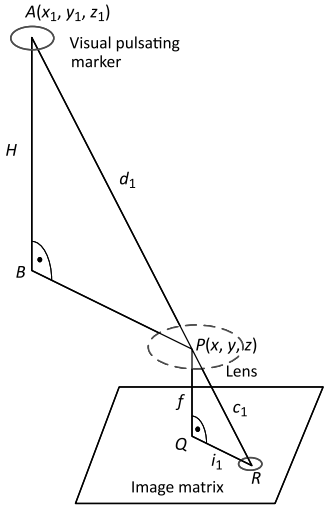

Fig. 2. Illustration showing the distribution of light rays incident from the marker onto the sensor matrix, along with mathematical relationships [23].

Using the similarity of triangles, the following relationship can be derived:

$$d_1=\frac{c_1}{f}*H,\tag{3.3}$$

where c_1 refers to the length depicted in Fig. 2, and f is the focal length of the lens.

Since,

$$c_1^2=f^2+i_1^2,$$

Eq. (3.3) can be rewritten as:

$$d_1^2=\frac{f^2+i_1^2}{f^2}*H^2.\tag{3.4}$$

Here, f is the camera's focal length, and i_1 represents the distance from the center of the detected marker's region in the image to the center of the image

sensor. This distance can be determined using the actual pixel size on the sensor. Using Eq. (3.4), the system of equations in (3.2) can be expressed as:

$$\begin{aligned}(x-x_1)^2+(y-y_1)^2+H^2&=\frac{f^2+i_1^2}{f^2}*H^2,\\(x-x_2)^2+(y-y_2)^2+H^2&=\frac{f^2+i_2^2}{f^2}*H^2,\\(x-x_3)^2+(y-y_3)^2+H^2&=\frac{f^2+i_3^2}{f^2}*H^2.\end{aligned}\tag{3.5}$$

After simplifications, the system of equations for determining the camera's position becomes:

$$\begin{aligned}(x-x_1)^2+(y-y_1)^2&=\frac{i_1^2}{f^2}*H^2,\\(x-x_2)^2+(y-y_2)^2&=\frac{i_2^2}{f^2}*H^2,\\(x-x_3)^2+(y-y_3)^2&=\frac{i_3^2}{f^2}*H^2,\\z_1-z&=H.\end{aligned}\tag{3.6}$$

The above equations enable the determination of the camera's coordinates (x, y, z) under the following assumptions:

- At least three light markers are required, mounted so that one coordinate is identical for all markers.
- The markers must be within the camera's field of view to appear in the captured images.
- Camera parameters (e.g., the focal length and actual pixel dimensions) must be known for calculations.

4. The concept of vision system

Figure 3 illustrates the conceptual workflow of the described vision system. The system requires the following parameters to be known for proper operation:

- the frequency and color of each marker,
- the image capture frequency of the camera (it should be at least twice the highest pulsation frequency of the marker and a multiple of it to minimize aliasing effects [19]),
- the resolution of the captured images.

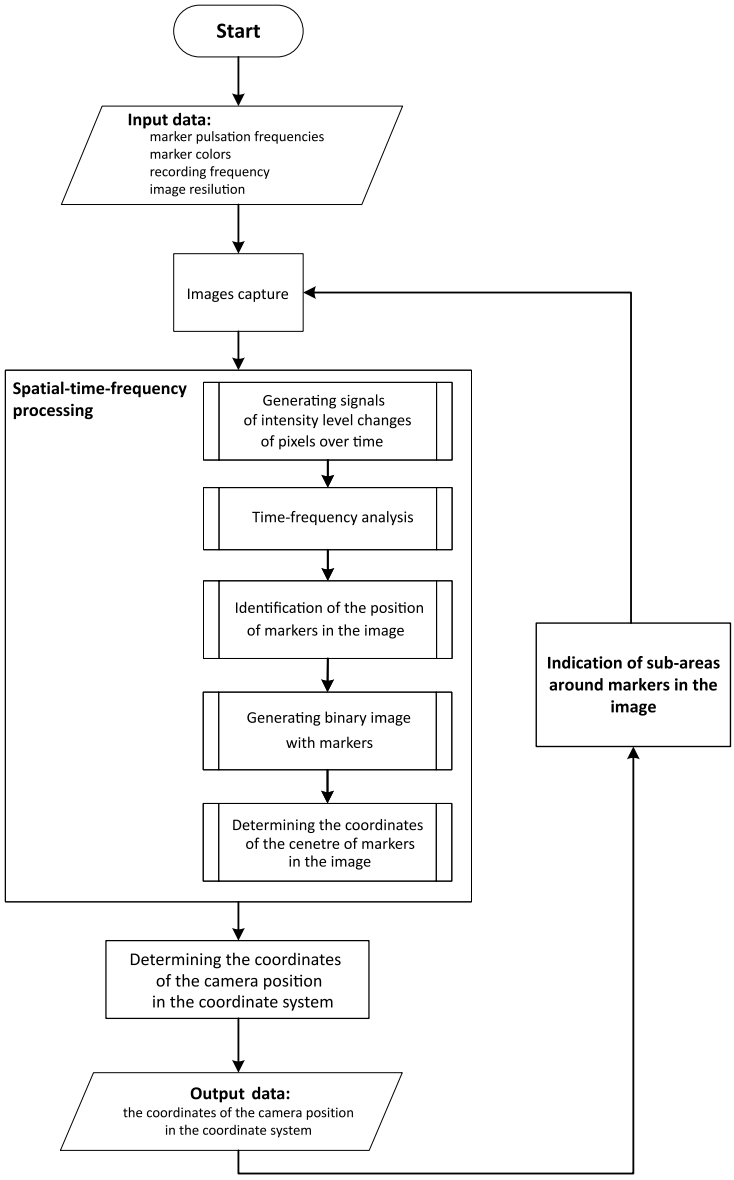

FIG. 3. Operation diagram of a vision system based on spatial-time-frequency image transformation.

The vision system records a sequence of images and proceeds to spatial-time-frequency processing. The system must analyze all pixels during the first iteration, because it initially does not know the camera's position or which pixels correspond to each marker. If the recorded images have dimensions $n \times m$, time-frequency analysis must be performed on $n \cdot m$ pixels. This first iteration of the

program is a calibration stage. The object with the recording camera should remain stationary during this stage, as this simplifies the processing to basic fast Fourier transform (FFT) analysis, and thereby speeding up the entire calibration process. Once the areas containing markers are located and their centers determined, the system can calculate the camera's position using the mathematical model and Eqs. (3.6). The output data consists of the camera's position coordinates in the reference coordinate system.

After identifying the marker positions in the images, the system determines groups of the nearest pixels to be processed in the next iteration. This approach significantly reduces the number of pixels requiring analysis, resulting in shorter processing times. It is worth noting that the size of the pixel neighborhood window depends on the speed of the object equipped with the camera – the faster the movement, the larger the window should be. The presented workflow of the vision system was implemented in MATLAB.

5. Stationary research results

5.1. Measurement preparation

The initial measurements were conducted in a stationary scenario, with the camera remaining fixed relative to the markers. This setup reflects the calibration stage of the system – localizing the markers in the images and determining the camera's position within the coordinate system. LED diodes covered with circular plastic diffusers were used for measurements. These covers significantly increased the probability of detecting the markers as circular shapes in the images. The markers had the following parameters:

- marker 1: green, frequency 20 Hz, coordinates [960, 107, 0] mm,
- marker 2: red, frequency 15 Hz, coordinates [712, 666, 0] mm,
- marker 3: red, frequency 20 Hz. coordinates [415, 317, 0] mm.

A Basler acA1600-60gc camera [24] equipped with a lens of 81.6 mm focal length [25] was used for the measurements. The following parameters were used for both measurements and calculations:

- resolution of recorded images: 1070×800 pixels,
- recording frequency: 60 Hz (60 frames per second),
- number of recorded images per measurement: 180,
- number of frames used for FFT analysis: 60,
- exposure time: 1 ms,
- focal length: 81.6 mm,
- pixel size: $4.5\ \mu\text{m} \times 4.5\ \mu\text{m}$.

The measurement setup is shown in Fig. 4. Symbol A denotes the board with the markers, and the symbol B represents the camera. By placing the markers on the board, the assumption of a single common coordinate was maintained. The camera was positioned perpendicular to the board.

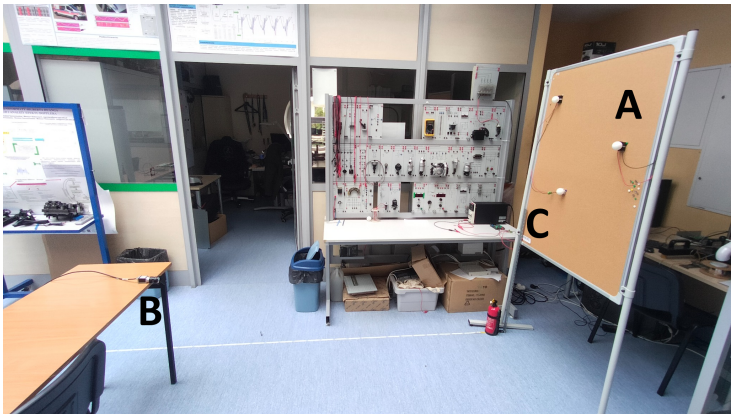

FIG. 4. The measuring station.

Measurements were performed for five different distances between the camera and the board with the markers. The camera's coordinates were known beforehand for testing purposes in order to facilitate comparison between actual values and those estimated by the vision system. The origin of the coordinate system was located at the bottom-left corner of the marker board and is marked by symbol C.

5.2. Measurements – marker recognition

Figure 5 shows the camera's view of the board with the markers. The markers in the image have been numbered according to the earlier description.

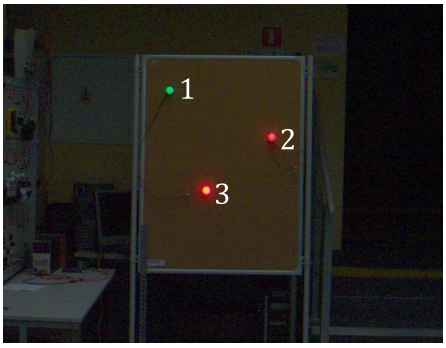

FIG. 5. The image from the camera's perspective – a board with numbered markers.

The spatial-time-frequency algorithm correctly identified the markers regardless of the camera's distance from the board. Sample binary images with the detected markers are shown on Figs. 6 and 7. It is worth noting here that the size of the detected area is not important – the main goal is the ability to determine the center of the pixel area containing the marker.

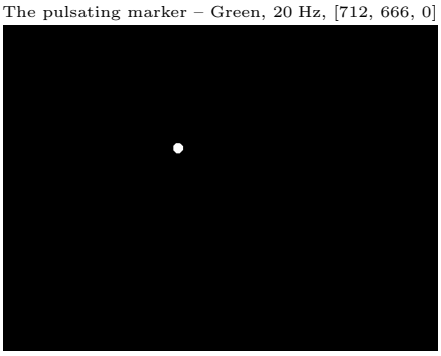

FIG. 6. The binary image with detected marker 1.

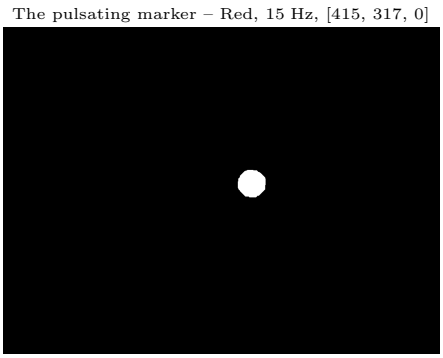

FIG. 7. The binary image with detected marker 2.

The vision system recognizes the light markers by analyzing the amplitude spectra of individual pixels. The first step involves checking the amplitude values at the frequency bands corresponding to the markers' pulsation frequencies for each pixel in the image. If these values exceed a certain threshold, the system then checks whether the amplitude is several times higher than that of the neighboring frequency bands. This condition aims to avoid false detection caused by high noise levels as an actual frequency enhancement. If this condition is satisfied, the next step is to compare the mean spectral values for the R and G color layers to identify the marker's color. To provide a better understanding of how these conditions operate in practice, Fig. 8 presents a collection of amplitude spectra from pixels where markers were identified.

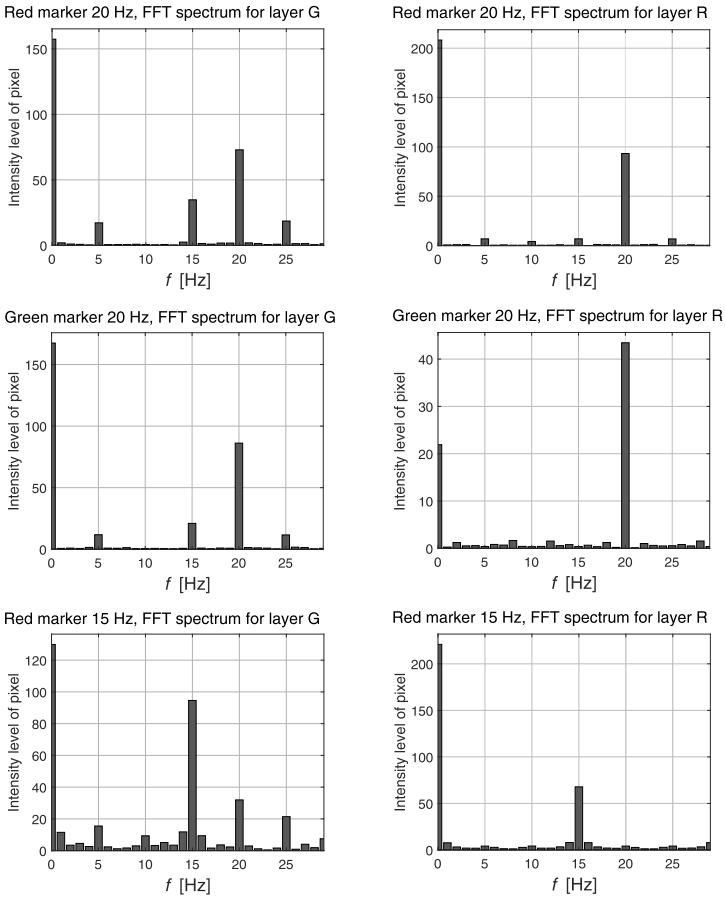

FIG. 8. A set of amplitude spectra for example pixels where pulsating light markers were recorded.

It should be noted that the spatial-time-frequency algorithm detects the area in the image where a pulsating marker has been recorded (Figs. 6 and 7). However, this area may not necessarily have a shape that is easy to analyze further. Therefore, the identified region in the image undergoes dilation and erosion operations once the pixels containing the marker are detected. The purpose of these operations is to approximate the detected marker area to a circular shape. This helps to indirectly reduce the impact of geometry distortions in the image and facilitates the estimation of the area's center.

5.3. Measurements – verification of position estimation

Measurements were performed for five different distances between the camera and the marker board. Table 1 presents the coordinates of the camera calculated

Table 1. Camera coordinates and relative error.

	Coordinate	Real value [mm]	Estimated value [mm]	Error [%]
1.	x	340	344	1.18
	y	730	736	0.82
	z	2800	2786	0.50
2.	x	340	343	0.88
	y	730	738	1.10
	z	4300	4188	2.60
3.	x	340	344	1.18
	y	730	732	0.27
	z	5800	5637	2.81
4.	x	340	344	1.18
	y	730	738	1.10
	z	7300	7153	2.01
5.	x	340	340	0
	y	730	738	1.10
	z	8800	8958	1.80

by the system in the specified reference system compared to the actual values, along with the relative errors.

The average relative error values for the individual coordinates are as follows:

- x-coordinate: 0.88%,
- y-coordinate: 0.88%,
- z-coordinate: 1.95%.

The average relative error does not exceed 2%, which can be considered a good result. For instance, an error of 2% corresponds to a discrepancy of approximately 2 cm when determining distances over 1 meter – acceptable for many applications.

At this point, it is worth mentioning that the authors assumed an ideal image geometry. This means that the camera's lens plane is always parallel to the plane in which the light markers are located. In reality, the lens may be at a slight angle, which affects the calculation results. Therefore, there is potential to improve the accuracy of the camera position estimation by introducing calibration and geometry correction.

6. Research results obtained during camera motion

In the next phase of the research, the vision system was tested for its ability to determine the camera's position during motion. Due to hardware limitations, the

pulsating light markers were mounted on a device that moved along a single axis at a constant velocity. The camera remained stationary. From the perspective of an observer near the markers, the camera appeared to move uniformly in the same direction but with the opposite orientation. This approach does not introduce significant errors. The setup with the installed markers is shown in Fig. 9.

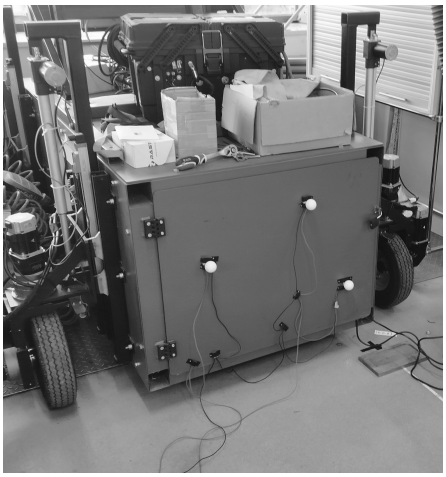

FIG. 9. The measuring station – light markers installed on a vehicle capable of uniform up-and-down movement.

Compared to the stationary tests, the resolution of the recorded images was adjusted to 800×600 pixels. The markers had the following parameters:

- marker 1: red, frequency 15 Hz, coordinates: [−763, 600, 0] mm,
- marker 2: green, frequency 20 Hz, coordinates: [−358, 760, 0] mm,
- marker 3: red, frequency 20 Hz, coordinates: [−165, 363, 0] mm.

The vision system detected the markers in the images, determined the initial position of the camera, and selected regions of neighboring pixels for further analysis. During subsequent image captures, the markers moved upward at a constant speed of 4 cm/s. The device was raised by 8 cm and thus increasing its height above the floor. Figure 10 shows the camera's view including the numbered markers and their neighboring pixel regions at the beginning of the motion.

Since the device moved in only one direction, only one coordinate (y-direction) changed. A model of the expected change in the camera's coordinates was created and compared to the vision system's estimates. The displacement of the light markers is shown in Fig. 11 and Table 2 presents the calculation results of the camera's coordinates.

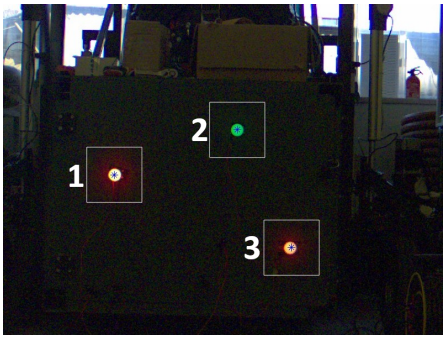

FIG. 10. View from the camera lens – detected light markers along with their centers and neighboring pixel areas selected for further analysis.

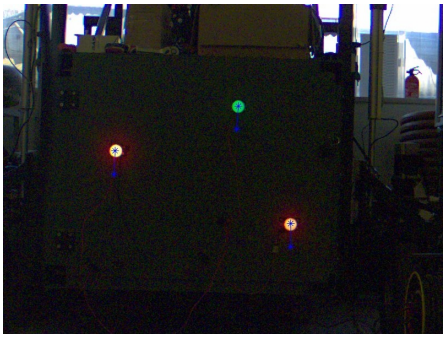

FIG. 11. Displacement of the light markers as observed from the camera lens.

TABLE 2. Calculated camera coordinates and corresponding relative errors.

	Coordinate	Real values [mm]	Estimated value [mm]	Error [%]
1.	x	−470	−473	0.64
	y	460	458	0.43
	z	3450	3371	2.30
2.	x	−470	−474	0.85
	y	380	383	0.79
	z	3450	3371	2.29

The vision system determined the camera's position with an error of approximately 2% similar to the stationary tests' results. The camera's displacement was calculated as 75 mm while the actual displacement value was 80 mm, resulting in a relative error of 6.25%. This error is higher than that observed in coordinate determination. The distance is calculated based on the coordinates so the distance error accumulates the errors in the camera's position estimation. However, the relative error in the camera's coordinates is small and may be

acceptable for many applications. Additionally, if the vision system is combined with another localization method, the error in determining the camera's (or object's) position could be reduced.

The examined concept of the localization system has not yet been implemented as a standalone real-time application. All data processing and computations were performed in MATLAB, after recording several sequences of images. The number of calculations is significantly influenced by the resolution of the recorded images (as the number of pixels determines the number of signals to be processed). For this reason, the system includes a calibration stage. This is the initial phase of operation when the system does not yet know the initial position of the camera. The algorithm processes all pixels to detect areas in the image containing markers. Once the markers are found, the system processes only the regions in the immediate vicinity of the markers (Fig. 10). However, even with this approach, the computational load remains considerable. To enable real-time operation, the system should be further optimized and can be implemented on FPGA platforms, which are known for high-speed processing capabilities.

A key aspect of this study involved evaluating marker detection during motion. Figures 12 and 13 present spectrograms showing the time-frequency spectrum of an example pixel where marker pulsations were detected. To improve

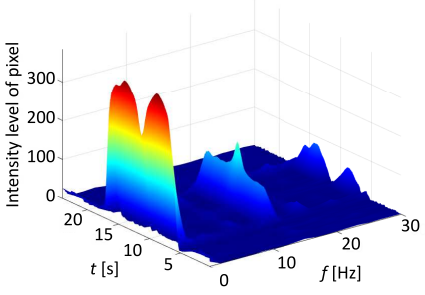

FIG. 12. Time-frequency spectrogram of a pixel where the marker appeared.

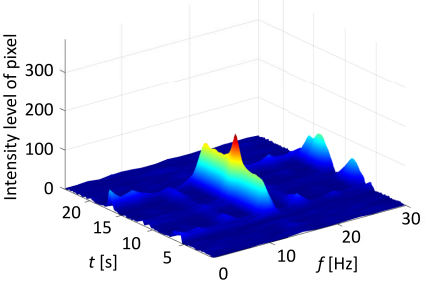

FIG. 13. Time-frequency spectrogram of the same pixel as in Fig. 11, shown without the average intensity level to enhance visibility.

readability, Fig. 13 shows the spectrogram with the average intensity level of the pixel. The marker with a pulsation frequency of 15 Hz was detected.

The spectrograms indicate the following points:

- No marker was present on this specific pixel at the start of the recording.
- The marker appeared at 5 seconds, as shown by an increase in both the average intensity level and the pulsation frequency amplitude.
- The marker's center reached this pixel in around 11 seconds, marked by the maximum pulsation frequency amplitude and a corresponding reduced average intensity level (Fig. 13).
- The marker left the pixel at 17 seconds, causing a significant drop in amplitude values.

By generating such spectrograms for various pixels in the image, the time interval during which the marker occupied a specific pixel can be estimated. Additionally, combining these spectrograms enables the creation of a trajectory that reflects the movement path of the detected marker.

7. Summary

The conducted studies allowed for evaluating the performance of the presented system. The vision system determines the camera's location within a reference frame with a relative error of 2% by applying spatial-time-frequency processing and the proposed equations. The system is capable of determining the camera's position under both stationary and dynamic conditions. The proposed vision system requires a calibration phase during which the camera must remain stationary at the initial stage of operation for proper functioning. Subsequently, the system identifies the region of neighboring pixels to be processed in subsequent iterations.

It is possible to generate spectrograms that enable precise motion analysis of the camera and identify the time intervals during which a marker occupied a specific pixel. Additionally, it is feasible to trace the movement paths of markers on recorded images. Using these trajectories for each marker, the camera's motion trajectory can also be estimated. However, the trajectory error may exceed the error in determining individual coordinates.

8. Conclusion

The presented vision system determines camera coordinates using spatial-time-frequency processing. It can effectively operate and support the localization process of objects in space (provided the conditions described in this article are met).

The authors aim to further develop the system and the spatial-time-frequency algorithm. Future work will focus on the region of neighboring pixels striving to more precisely define the structure and size of this analysis area. This should have an impact on processing time. Consequently, the authors aim to reduce the overall processing time. While the system's concept has been introduced, several interesting aspects remain to be explored and refined, such as employing alternative time-frequency methods to generate more accurate information about marker positions and their time of occurrence.

Authors' statement

The authors declare that there are no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

References

1. Liu H.S., Pang G., Positioning beacon system using digital camera and LEDs, *IEEE Transactions on Vehicular Technology*, **52**(2): 406–419, 2003, https://doi.org/10.1109/TVT.2002.808800.
2. Yoshino M., Haruyama S., Nakagawa M., High-accuracy positioning system using visible LED lights and image sensor, *2008 IEEE Radio and Wireless Symposium*, RWS, Orlando, FL, pp. 439–442, 2008, https://doi.org/10.1109/RWS.2008.4463523.
3. Hassan N.U., Naeem A., Pasha M.A., Jadoon T., Yuen C., Indoor positioning using visible LED lights: A survey, *ACM Computing Surveys*, **48**(2): Article 20, 2015, https://doi.org/10.1145/2835376.
4. Li Y., He L., Zhang X., Zhu L., Zhang H., Guan Y., Multi-sensor fusion localization of indoor mobile robot, *IEEE International Conference on Real-time Computing and Robotics (RCAR)*, Irkutsk, Russia, 2019, pp. 481–486, 2019, https://doi.org/10.1109/RCAR47638.2019.9044006.
5. Tanaka H., Ultra-high-accuracy visual marker for indoor precise positioning, *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, pp. 2338–2343, 2020, https://doi.org/10.1109/ICRA40945.2020.9196535.
6. Lu Z., Wu Y., Yang S., Zhang K., Quan Q., Fast and omnidirectional relative position estimation with circular markers for UAV swarm, *IEEE Transactions on Instrumentation and Measurement*, **73**: 5034911, 2024, https://doi.org/10.1109/TIM.2024.3472903.
7. Guan W., Wen S., Zhang H., Liu L., A novel three-dimensional indoor localization algorithm based on visual visible light communication using single LED, *2018 IEEE International Conference on Automation, Electronics and Electrical Engineering (AUTEEE)*, Shenyang, China, pp. 202–208, 2018, https://doi.org/10.1109/AUTEEE.2018.8720798.
8. Li Y., Zhu S., Yu Y., Wang Z., An improved graph-based visual localization system for indoor mobile robot using newly designed markers, *International Journal of Advanced Robotic Systems*, **15**(2): 1729881418769191, 2018, https://doi.org/10.1177/1729881418769191.

9. Novák M., Dobesch A., Wilfert O., Janík L., Visible light communication transmitter position detection for use in ITS, *Optical Switching and Networking*, **33**: 161–168, 2019, https://doi.org/10.1016/j.osn.2018.04.002.

10. Zhuang Y., Hua L., Qi L., Yang J., Cao P., Cao Y., Wu Y., Thompson J., Haas H., A survey of positioning systems using visible LED lights, *IEEE Communications Surveys and Tutorials*, **20**(3): 1963–1988, 2018, https://doi.org/10.1109/COMST.2018.2806558.

11. Soner B., Karakas M., Noyan U., Sahbaz F., Coleri S., Vehicular visible light positioning for collision avoidance and platooning: A survey, *IEEE Transactions on Intelligent Transportation Systems*, **25**(7): 6328–6344, 2024, https://doi.org/10.1109/TITS.2023.3349160.

12. Zekavat S.R., Buehrer R.M., Durgin G.D., Lovisolo L., Wang Z., Goh S.T., Ghasemi A., An overview on position location: past, present, future, *International Journal of Wireless Information Networks*, 2**8**(1): 45–76, 2021, https://doi.org/10.1007/s10776-021-00504-z.

13. Khan M.N., Jamil M., Gilani S.O., Ahmad I., Uzair M., Omer H., Photo detector-based indoor positioning systems variants: A new look, *Computers & Electrical Engineering*, **83**: 106607, 2020, https://doi.org/10.1016/j.compeleceng.2020.106607.

14. Rego M., Fonseca P., OCC based indoor positioning system using a smartphone camera, [in:] Santos V., Lau N., Neto P., Lopes A.C. [Eds.], *Proceedings of IEEE International Conference on Autonomous Robot Systems and Competitions*, *ICARSC*, pp. 31–36, 2021, https://doi.org/10.1109/ICARSC52212.2021.9429782.

15. Nguyen N.H., Dinh Nguyen Q., Enhanced bit-rate performance for visible light communication systems between led and mobile camera, [in:] Jeong S.H., Loc H.D., Fdida S., Le-Ngoc T. [Eds.], *Proceedings of Tenth International Conference on Communications and Electronics (ICCE),* pp. 544–549, 2024, https://doi.org/10.1109/ICCE62051.2024.10634616.

16. Zhou Z., Liang B., Cao Y., Zhang M., MPPM spectrum analysis based on PPM, [in:] *Proceedings of 14th International Conference on Computer Research and Development (ICCRD)*, pp. 356–362, 2022, https://doi.org/10.1109/ICCRD54409.2022.9730597.

17. Das S., Mandal S.K., Spectral analysis of dimming controlled MH-HPPM for visible light communication, *Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 291–296, 2021, https://doi.org/10.1109/WiSPNET51692.2021.9419413.

18. Bui T.-C., Biagi M., Kiravittaya S., Theoretical analysis of optical spatial multiple pulse position modulation, *IEEE Global Communications Conference (GLOBECOM)*, 2018, https://doi.org/10.1109/GLOCOM.2018.8647138.

19. Miś P., Szulim P., Analysis of the possibility of using markers emitting pulsating light in the task of localization, *Applied Computer Science*, 1**7**(1): 26–39, 2021, https://doi.org/10.23743/acs-2021-03.

20. Miś P., Szulim P., The effect of camera exposure on the results of spatial-frequency processing and the quality of the obtained amplitude images, [in:] Szewczyk R., Zieliński C., Kaliczyńska M. [Eds.], *Automation 2022: New Solutions and Technologies for Automation, Robotics and Measurement Techniques*, Springer International Publishing, Cham, pp. 306–316, 2022.

21. Zhu B., Cheng J., Wang Y., Yan J., Wang J., Three-dimensional VLC positioning based on angle difference of arrival with arbitrary tilting angle of receiver, *IEEE Journal on Selected Areas in Communications*, **36**(1): 8–22, 2018, https://doi.org/10.1109/JSAC.2017.2774435.

22. Rahman M.H., Sejan M.A.S., Kim J.-J., Chung W.-Y., Reduced tilting effect of smartphone CMOS image sensor in visible light indoor positioning, *Electronics* (*Switzerland*), **9**(10): 1–19, 2020, https://doi.org/10.3390/electronics9101635.

23. Hossen M.S., Park Y., Kim K.-D., Performance improvement of indoor positioning using light-emitting diodes and an image sensor for light-emitting diode communication, *Optical Engineering*, **54**(4): 045101, 2015, https://doi.org/10.1117/1.OE.54.4.045101. https://doi.org/10.1117/1.OE.54.4.045101

24. *Basler ace Classic acA1600-60gc*, n.d., https://www.baslerweb.com/en-us/shop/aca1600-60gc/ (accessed on 2024.12.18).

25. *8.5mm, 200mm-∞ Primary WD, HR Series Fixed Focal Length Lens*, n.d., https://www.edmundoptics.com/p/85mm-200mm-infin-primary-wd-hp-series-fixed-focal-length-lens/22903/ (accessed on 2024.12.18).

Received December 31, 2024; accepted version April 25, 2025.
Online first July 1, 2025.