APPLICATIONS OF
MODELLING AND SIMULATION
http://arqiipubl.com/ams eISSN 2600-8084 VOL 9, 2025, 213-226

A Review: Obstacle Perception Based on Panoramic Vision in Low-altitude Rotorcraft UAVs

Jiang Xiaoyan[1,2], Khairul Hamimah Abas[1*] and Abdul Rashid Husain[1]

[1]Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[2]Hebei Institute of Mechanical Electrical Technology, Xingtai, Hebei, China

[*]Corresponding author: khairulhamimah@utm.my

Submitted 07 March 2025; Revised 04 May 2025; Accepted 16 May 2025; Available online 28 May 2025.

Abstract: In recent years, Low-altitude Rotorcraft Unmanned Aerial Vehicles (UAVs) have shown great potential in many tasks, such as aerial mapping, search and rescue, plant protection, and more. In unknown environments, UAVs face the risk of collisions with various unforeseen obstacles. For UAVs, the ability to effectively perceive and recognize obstacles from all directions is crucial for ensuring safe flight, and a variety of sensors have been employed for this purpose. Among these, panoramic vision-based obstacle perception systems offer significant advantages. This paper focuses on key technologies for obstacle detection based on panoramic vision in UAVs. It reviews the current state of research and applications both domestically and internationally, covering aspects such as hardware platforms, key algorithms (including model correction, distortion correction, feature matching, image stitching), panoramic imaging, and omnidirectional obstacle detection. The paper also identifies major bottlenecks and outlines future development directions. Based on this work, we will continue our research on autonomous omnidirectional obstacle detection and avoidance.

Keywords: Fisheye camera; Obstacle detection; Panoramic vision; Review; Low-altitude UAV.

1. INTRODUCTION

In recent years, Low-altitude Rotorcraft UAVs have gradually become the preferred platform for performing low-airspace tasks, such as aerial surveying, search and rescue, and aerial inspection, due to their advantages in portability and flexibility. In practice, the uncertainty and complexity of low-airspace environments pose significant challenges, as various obstacles (including trees, buildings, power lines, birds, and other flying objects) threaten the safe flight of UAVs. Therefore, it is urgent to further enhance their capabilities in autonomous perception, decision-making, and path-planning. Among these, autonomous perception of various obstacles is the most fundamental and critical capability.

Generally, obstacle perception methods used in UAVs rely on infrared sensors, ultrasonic sensors, laser sensors, radar sensors or visual sensors. With the advancement of hardware technology, the visual sensors have become increasingly smaller in size, more energy-efficient, and higher in resolution. In recent years, visual sensor-based obstacle avoidance has become an essential component in various UAV systems. In practice, mainstream visual sensor-based obstacle avoidance systems include monocular vision [1], binocular vision [2], trinocular vision [3], panoramic vision [4]. Among these, panoramic vision refers to an omnidirectional imaging system that typically employs multiple cameras or fewer wide-angle cameras to construct a surround-view system.

UAVs equipped with an omnidirectional field of view (FOV) can observe large visible regions, capturing scene information from all directions. With the ability to acquire comprehensive environmental data in a short time, panoramic vision is highly suitable for obstacle avoidance in unknown flight environments. For the obstacle avoidance tasks [5], it is essential to accurately identify and locate relevant obstacles from a large amount of environmental information. This demands not only the high-quality implementation of relevant algorithms but also the support of high-performance hardware. For instance, high-resolution visual sensors integrated into platforms such as the EVO II and Mavic 3 enable panoramic vision systems that not only capture real-time images and videos but also assist UAVs in achieving autonomous obstacle avoidance and navigation. These applications offer valuable insights and directions for research and practical implementation of real-time omnidirectional obstacle avoidance in UAVs.

A typical panoramic vision system consists of a hardware platform (including visual sensors and on-board computing units) and a software system (including image processing, panoramic imaging, and control signal output modules). Based on structural differences, the implementation of panoramic vision can be categorized into 6 methods as summarized in Table 1. The table presents the performance level of each method across various issues, where five stars represent the best possible rating.

Table 1. Comparison of different panoramic imaging systems.

Category	Ref	FOV	Imaging Qualification	Real-time	Structural Complexity	Cost	Applicability
Monocular-based	[6]	★★☆☆☆	★★☆☆☆	★★★☆☆	★☆☆☆☆	★☆☆☆☆	Small UAV
Multiple images stitching	[7]	★★★★★	★★★★☆	★★☆☆☆	★★★★☆	★★★☆☆	Middle-large UAV
Fisheye panoramic system	[8]	★★★★☆	★★☆☆☆	★★★★☆	★★☆☆☆	★★☆☆☆	Small UAV
Panomorph imaging system	[9]	★★★★☆	★★★★☆	★★★☆☆	★★★★☆	★★★★☆	Specific UAV
Catadioptric panoramic system	[10]	★★★★★	★★★★☆	★★☆☆☆	★★★★★	★★★★☆	Static Observation
Monocentric panoramic system	[11]	★★★★★	★★★★★	★★★☆☆	★★★★★	★★★★★	Advanced Experiment

By comparison: The Monocular Panoramic System is characterized by its simple structure and low cost, which facilitates integration into small UAVs. However, it is generally incapable of achieving true omnidirectional perception without relying on mechanical rotation or the motion of the UAV itself. As a result, the field of view is discontinuous, blind spots are present, and depth information is unavailable. The multiple image stitching system is capable of producing high-resolution panoramic images. The number and arrangement of cameras can be configured flexibly; however, accurate calibration and synchronization are required. The image stitching algorithms are complex, and during high-speed flight, stitching delays can negatively impact real-time performance. The fisheye panoramic system can achieve a viewing angle of 180° or greater, reducing the number of required cameras and making it suitable for small UAVs. However, it suffers from severe image distortion, blurred imaging in the edge regions, and difficulties in depth estimation. The panomorph imaging system enhances image resolution in specific directions (such as the front or bottom) and optimizes the use of computing resources. However, its system design is complex and challenging to mass-produce. In addition, it requires extensive image preprocessing, including specialized decoding and correction algorithms. The catadioptric panoramic system achieves a full panoramic view by combining mirrors and cameras. It produces distortion-free imaging at the centre, making it suitable for high-quality imaging in static scenes. However, the system has a complex structure and large size, and the mirror support components are prone to obstructing the field of view. The monocentric panoramic system adopts a unified optical axis design, features a compact structure, and is well-suited for high-end micro-UAVs. However, the precise arrangement of lenses and sensors makes the system extremely costly and difficult to calibrate and maintain.

In this paper, we present a detailed review and analysis of the hardware platform and key algorithms for panoramic vision-based obstacle avoidance in Low-altitude Rotorcraft UAVs. We compared the advantages and disadvantages of several representative solutions, and conducted preliminary experiments on key strategies, including hardware platform construction, panoramic data acquisition, panoramic image stitching, and omnidirectional obstacle detection.

2. HARDWARE PLATFORM OF PANORAMIC VISION

In this paper, we focus on two categories of panoramic vision-based hardware platforms: single camera-based and multiple cameras-based.

2.1 Single Camera-based Panoramic Vision Platform

There are two main methods for acquiring panoramic information using a single camera. One approach involves utilizing the principles of light refraction to directly obtain panoramic images. The work presented in Ref. [12] introduced several omnidirectional cameras sharing common components, including a panoramic imaging system composed of a paraboloidal mirror, an orthographic lens system, and a CCD video camera. Similarly, Ref. [13] designed a compact and cost-effective omnidirectional stereo vision system using a concave lens and a convex mirror, achieved with only a single camera. Ref. [14] proposed a rotating optics system consisting of prism sheets, circular or linear polarizing films, and a hyperboloidal mirror, offering two operational modes corresponding to the left and right eye images. All these systems rely on mirrors or reflective components, which inevitably increase the weight and structural complexity of the omnistereo imaging system.

The other approach involves capturing a series of images by rotating or moving a monocular camera [15]. By seamlessly stitching these images, a panoramic view can be generated. The key challenges lie in the accurate matching and stitching [16] of overlapping regions between adjacent images. The work in Ref. [17] presented an online-generated panoramic view of the user's environment, suitable for execution on current-generation mobile devices such as smartphones. Generally, these methods are applicable only in static scenes and are not suitable for real-time applications. Ref. [18] proposed two possibilities for capturing omnistereo panoramas using optics without any moving parts. They introduced a special mirror designed to replicate the same rays produced by rotating cameras, as well as a lens for omnistereo panoramas based on curves whose caustic forms a circle. In practice, several factors, such as visible seams, vertical disparities in moving objects, and uneven camera motion, can degrade the quality of stitching, potentially leading to significant visual artifacts or distortions in the final

panoramas [19].

Based on the above analysis, panorama acquisition using a monocular camera system generally requires the UAV to possess stable hovering capabilities and relies on the consistent rotational performance of a mechanical platform. Furthermore, to ensure high-quality image stitching, additional redundant information or sufficiently large overlapping regions between images are necessary. This, in turn, can lead to increased energy consumption, which may be undesirable for UAVs with limited power resources

2.2 Multiple Cameras-based Panoramic Vision Platform

The multiple cameras-based panoramic vision platform consists of several cameras mounted along the same horizontal plane, surrounding the UAV. By combining images from different viewpoints, the FOV of the individual cameras can collectively cover all directions. The key to implementing such a platform lies in achieving precise synchronization across all cameras. Systems such as Google Jump or Stereo Panorama Modeler [20] employ electronic synchronization mechanisms to ensure simultaneous triggering, enabling the capture of 360° environmental information in real time. In Ref. [21], a system was developed using three wide-angle camera modules, each with a 132° FOV, allowing full 360° coverage. However, the device is relatively large, structurally complex, and challenging to manufacture and deploy. A fundamental aspect of multi-camera platforms is image stitching. In commercial software such as PanoramaStudio [22], PTGui [23], and Hugin [24], high-quality stitching requires significant user expertise and repeated practice, making them less accessible to non-professionals. Moreover, common artifacts—such as motion parallax, visible seams, and synchronization errors [18]—often appear in the final stitched panoramas, especially under dynamic or high-speed conditions.

In addition, weather conditions play a critical role in the performance of obstacle detection systems. For example, the Mavic 3 - a widely used UAV equipped with six fisheye cameras and two wide-angle cameras for omnidirectional obstacle avoidance - demonstrates certain limitations in adverse environments. Practical experience has shown that its obstacle avoidance capabilities are reduced in scenarios involving rain, snow, fog, or reflective surfaces such as water or ice. Moreover, low ambient light conditions may cause the vision system to malfunction or perform unreliably.

As mentioned above, both monocular and multi-camera systems share certain similarities. First, multiple images covering omnidirectional scene information must be acquired; Then, panoramic images are generated through image stitching algorithms. In both cases, when the time interval between adjacent frames is sufficiently short - or when the number of cameras is large - a significant amount of image processing is required. Consequently, this imposes high demands on the real-time processing capabilities of onboard computing platforms.

Therefore, for any panoramic vision-based obstacle avoidance system, future development should focus on mitigating the impact of adverse weather conditions, enhancing the quality and efficiency of image stitching, and reducing the technical barriers to implementation.

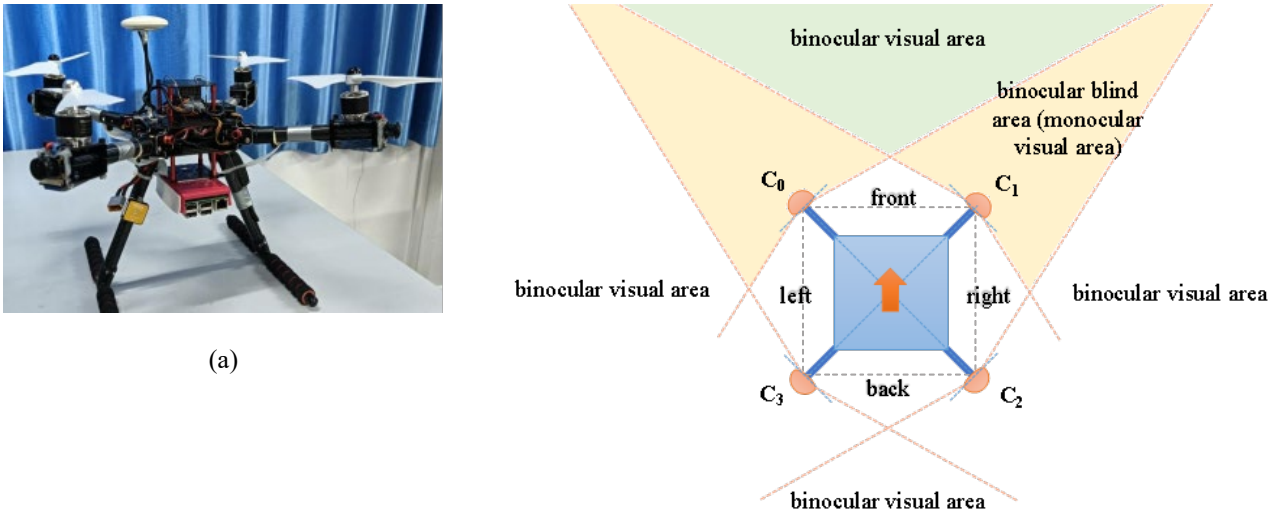
2.3 Fisheye Camera-based Panoramic Vision

Generally, a single camera-based panoramic vision system generally requires a rotating head to capture panoramic image sequences, which leads to a lack of synchronization in the acquired image sequences and makes it unsuitable for fast-moving UAVs. In contrast, a multi-camera panoramic system increases in cost and payload, and the calibration accuracy of the multi-view system decreases as the number of cameras increases. In practical applications, fisheye or wide-angle cameras are often used due to their large FOV [25]. With a larger FOV, fewer fisheye cameras are needed to capture horizontal 360°×vertical 180° omnidirectional information compared to conventional cameras. This not only facilitates omnidirectional obstacle detection but also reduces the payload requirements for UAVs. Recent studies on obstacle detection using panoramic vision in low-altitude rotorcraft UAVs have increasingly focused on wide-angle cameras [21] and fisheye cameras [26].

To construct a fisheye cameras-based panoramic vision system suitable for obstacle avoidance, researchers around the world have made numerous attempts. In systems such as AerialPro360 [27] and Ricoh Theta [28], two fisheye cameras are used to build the panoramic vision system. These configurations provide maximum stereo information along the median line (perpendicular to the baseline) but lack stereo information along the baseline. Moreover, the image resolution is often insufficient to meet the requirements of image processing algorithms. Ref. [29] proposed geometry calibration, photometric compensation and seamless stitching algorithms to address challenges such as fisheye distortion, luminosity inconsistency, and non-collocated optical centres. The work in Ref. [30] introduced a camera setup capable of capturing omnistereo videos using a minimum of three fisheye cameras without any moving parts. Ref. [31] employed a panoramic vision system composed of four fisheye cameras. Increasing the number of cameras helped reduce distortion, thereby improving obstacle detection performance. Furthermore, the study in Ref. [32] utilized six GoPro cameras (horizontal 220°× vertical 195°), each mounted under a separate rotor arm, to construct a surround-view system. The wide FOV of the GoPro cameras ensures sufficient overlapping regions between adjacent images, which is advantageous for generating artifact-free stitched panoramic images.

2.3.1 UAV Platform

To develop a panoramic vision obstacle detection system suitable for small quadrotor UAVs, we comprehensively consider factors such as the payload capacity, operating environment, task requirements, computational demands, and cost control. As a result, we selected the Raspberry PI 4.0B as the onboard processing unit, and mounted four fisheye cameras beneath the four propellers to construct a panoramic vision hardware platform. This configuration provides a FOV that covers the full 360° environment surrounding the UAV (Figure 1).



(a) (b)

Figure 1: (a) The UAV platform, (b) Binocular visual area.

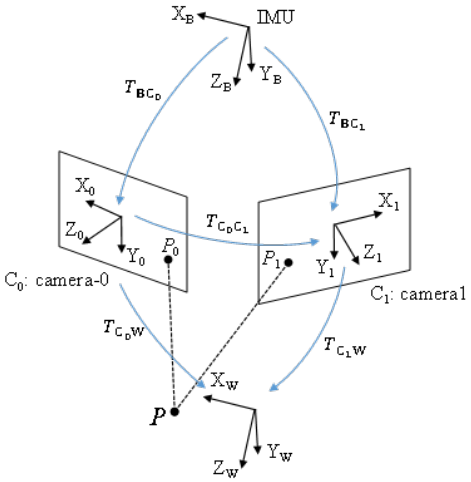


Figure 2. Transformation between Cameras and IMU.

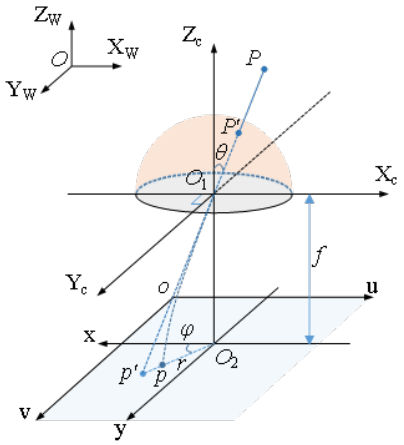


Figure 3. Imaging process of a fisheye camera [33].

As shown in Figure 1, the components highlighted in blue represent the quadrotor UAV. C₀ to C₃ are the four fisheye cameras, each with a FOV of 222°. Any two adjacent cameras share a certain overlapping visible area, referred to as the binocular visual area, allowing them to be approximately treated as a binocular vision system. Figure 2 illustrates the transformation relationships among the IMU (representing the UAV body), the fisheye cameras (using C_0 and C_1 as an example), and a spatial point P. $X_W Y_W Z_W$ is the world coordinate; $X_B Y_B Z_B$ is the IMU coordinate; $X_0 Y_o Z_o$and $X_1 Y_1 Z_1$ are, respectively, the camera coordinate of fisheye cameras C_0 and C_1. P_0 and P_1 are, respectively, the imaging point of a space point P in C_0 and C_1.

2.3.2 Projection Model of Fisheye Camera

Fisheye lenses are capable of capturing incident light from a much wider range and projecting it onto a size-limited imaging plane through varying degrees of refraction. The FOV of a fisheye camera can reach or even exceed 180°. As a result, unlike the pinhole projection model used in conventional cameras, the fisheye model must incorporate distortion parameters to accurately represent its imaging characteristics.

Figure 3 shows the distortion-based imaging process of a fisheye camera where $O_1 - X_C Y_C Z_C$: Camera coordinate system, $O_2 - xy$: Physical image coordinate system, $o - uv$: pixel coordinate system, P: A point in the world coordinate system $O_1 - X_W Y_W Z_W$, $p(x, y)$: The distorted imaging point P captured by the fisheye camera, $p'(x', y')$: The undistorted projection points of P using a pinhole camera model, θ: The incidence angle between the $Z_c - axis$ and the incoming ray from P, r: the radial distance from p to the principal point O_2, φ: The polar angle between the radial direction of p and the $x - axis$ and f: The focal length of the fisheye camera. The process can be described as follows:

- Firstly, according to the imaging process of pinhole model, a 3D point $P(X, Y, Z)$ is projected onto the image plane as a point $p'(a, b)$ without distortion. The polar coordinates of p' can be denoted as (r, φ), where r is the radial distance from the imaging point to the principal point, and φ is the corresponding polar angle. The angle of projected incidence ray,

denoted as θ, can be calculated accordingly.

- Due to lens distortion, the light ray exits at angle $\theta_d \neq \theta$, and the actual imaging point on the distorted image plane becomes $p'(x', y')$.
- Using the equidistant projection model, the relationship between the incident angle θ and the radial distance r on the image plane can be described as $r = f \cdot \theta$, where f is the focal length of the fisheye lens, and θ is the angle between the optical axis and the incoming light ray. To simplify the model or account for small-angle approximations, a Taylor series expansion can be applied to approximate the projection function:

$$\theta_d = \theta(1 + k_1\theta^2 + k_2\theta^4 + k_3\theta^6 + k_4\theta^8) \tag{1}$$

where, k_1, k_2, k_3, k_4 are high-order distortion coefficients, calculated by camera calibration.

- Let $r_d = \theta_d$, then the polar coordinates of p' are (θ_d, φ), and the corresponding Cartesian coordinates are

$$x' = \frac{\theta_d}{r}a, y' = \frac{\theta_d}{r}b \tag{2}$$

- To convert the point $p'(x', y')$ from the camera coordinate system to the pixel coordinate system, the camera intrinsic parameters are applied. The transformation is given by:

$$\begin{cases} u = f_x x' + c_x \\ v = f_y y' + c_y \end{cases} \tag{3}$$

where, f_x, f_y are the focal lengths in the x and y directions (in pixels). c_x, c_y are the coordinates of the principal point (optical centre) in the image.

3. KEY ALGORITHMS IN OBSTACLE DETECTION

As shown in Figure 4, the key algorithms involved in obstacle detection based on panoramic vision include system calibration, omnidirectional information acquisition, distortion correction, feature matching, fisheye image stitching, obstacle recognition and extraction, among others.

3.1 System Calibration

3.1.1 Parameters

According the Figures 2 and 3, in the process of projecting a 3-D point $P(X, Y, Z)$ onto a 2-D image, the extrinsic parameters $T_{C_0C_1} = (R, T)$ and intrinsic parameters $(k_0, k_1, m_u, m_v, u_0, v_0, k_2, k_3, k_4)$ must be calculated.

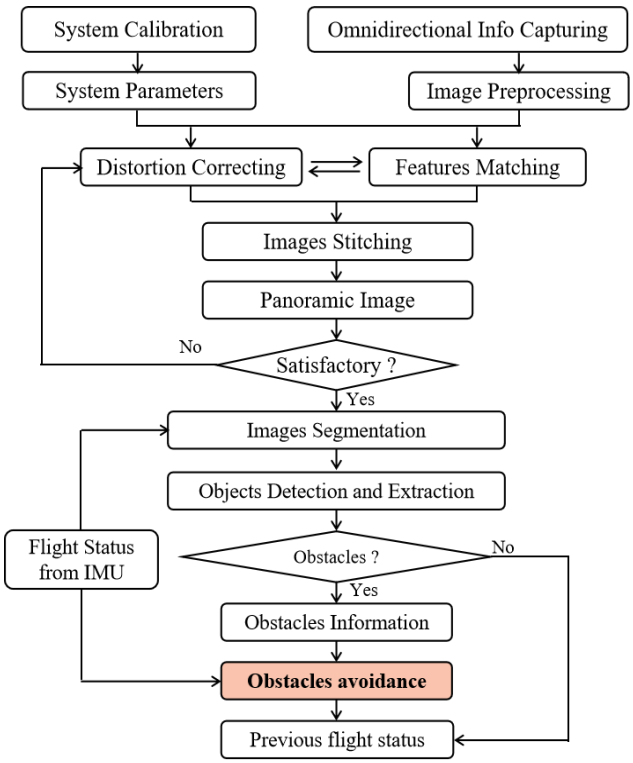


Figure 4. Flowchart of obstacle avoidance.

For the extrinsic parameters, R is a 3×3 rotation matrix and $T = (t_x, t_y, t_z)^T$ is a translation vector. Considering the similarities with a binocular system, R and T can be obtained through binocular calibration [34]. Additionally, the extrinsic parameters, including $R_{BC_i}$and t_{BC_i}, can be calibrated via Camera-IMU calibration. The intrinsic parameters are used to map the camera coordinates onto the image plane. From the perspective of automatic calibration, the first five terms of the Taylor expansion of r with respect to θ are generally used to approximate the radially symmetric part of the fisheye camera model:

$$r(\theta) \approx k_0\theta + k_1\theta^3 + k_2\theta^5 + k_3\theta^7 + k_4\theta^9 \tag{4}$$

From this formula, the radially symmetric part of the fisheye camera model contains the five parameters: k_0, k_1, k_2, k_3, k_4, which are the radical distortion parameters that need to be calibrated.

In the pixel coordinate system, the origin is located at the top-left corner of the image plane. The imaging point $p = (u, v)$ can then be calculated as:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{bmatrix} m_u & 0 \\ 0 & m_v \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \tag{5}$$

where $(u_0, v_0)^T$ is the principal point, and m_u, m_v are the number of pixels per unit distance in the horizontal and vertical directions, respectively.

3.1.2 Parameters Calibrating

That is, obtaining the intrinsic and extrinsic parameters of a fisheye camera. Fisheye cameras are widely used in machine vision due to their wide FOV. However, the distortion in fisheye images pose challenges to calibration accuracy [35]. The work in Ref. [36] proposed a precise mathematical imaging model and implemented an adaptive automatic corner detection algorithm, addressing some limitations of traditional methods, such as original constraints and the need for user intervention. Building upon high-precise calibration of the intrinsic and external parameters of fisheye cameras with FOVs greater than 180°, Ref. [37] defined a loss function based on camera projection to effectively optimize extrinsic parameters (tilt and roll angle) and intrinsic parameters (focal length). For setups involving three or more fisheye cameras, considering the error caused by distortion in fisheye images, Ref. [38] proposed calibration and stereo measurement methods that introduce optimal baseline weights to improve distance accuracy and ensure a more uniform error distribution.

In practice, the larger the FOV of a fisheye camera, the more serious the imaging distortion becomes, particularly for objects near the edges of the image. In a panoramic vision system composed of fisheye cameras, high-quality, seamlessly stitched panoramas can only be obtained after accurate calibration and distortion correction [36]. These panoramas are essential for the reliable localization and extraction of omnidirectional obstacles.

To calibrate the parameters of our panoramic vision system, we followed the process outlined:

- Step 1. Monocular Calibration:
  The intrinsic and distortion parameters of each fisheye camera were obtained individually. A checkerboard calibration plate was placed at various positions and angles within the FOV of each camera to capture a sufficient number of calibration images. An appropriate fisheye model and distortion model were selected. In our case, we adopted the intrinsic and distortion parameters described above as the model parameters. The Camera Calibration Toolbox in OpenCV was used to compute the intrinsic parameter matrix and distortion coefficients for each fisheye camera.
- Step 2. Extrinsic Calibration of Multiple Cameras:
  To enable panoramic stitching, the relative pose—comprising the rotation matrix R and transfer vector T—between the four fisheye cameras were calculated. The calibration board was positioned within the overlapping regions of the FOVs of the adjacent cameras, and multiple images were captured simultaneously. Feature points on the calibration board were extracted and matched across overlapping FOVs.

A binocular calibration tool was then used to estimate the relative poses between adjacent camera pairs.

3.2 Image Preprocessing

In natural environments affected by fog, rain, or snow, imaging quality deteriorates significantly, which severely impacts the performance of visual perception algorithms. Addressing the influence of such harsh environmental conditions has long been a key challenge in the field of machine vision.

Studies in Refs. [39], [40], and [41] have proposed algorithms to enhance the quality of real-time images. By collecting the data under adverse weather conditions, corresponding neural network models can be trained to improve image robustness. Additionally, UAV dynamics can serve as auxiliary information for self-state perception [42], further enhancing the effectiveness of visual perception under challenging environmental conditions.

3.3 Distortion Correcting

In practice, distortion in fisheye images primarily includes radical distortion and tangential distortion. These distortions can lead to significant geometric deformation of object images, making distortion correction essential to eliminate visual degradation caused by such deformations.

For simple correction algorithms, polynomial models [43] can be used to fit the radical distortion parameters. The work in Ref. [44] proposed a "smile"-based lens distortion correction (SLDC) method, which utilizes the Hough transform to reconstruct distorted lines using key points along the smile-shaped Hough pattern. Ref. [36] extracted feature points from a

calibration plate with a known pattern and completed the distortion computation and correction accordingly. Other commonly used distortion correction models include the Equidistant Projection Model [45], the Double Sphere Model [46], methods based on deep learning [47]. In practical applications, the severity of distortion determines the most suitable correction model and parameters.

In our study, the distortion correction process for fisheye images is as follows:

- Step 1. Select the distortion model: We adopted the Isometric model, expressed as $r = f \cdot \theta$.
- Step 2. Select the projection model after correction: We used a planar projection model to produce a panoramic image resembling a normal camera view after correction.
- Step 3. Compute the mapping matrix: We employed the *cv2.InitUndistortRectifyMap*( ) function from OpenCV to calculate the mapping matrix based on the chosen distortion and projection models.
- Step 4. Apply the correction mapping: The original fisheye image was transformed into the corrected image using the computed mapping.

3.4 Distortion Correcting

3.4.1 Choice of Overlapping Regions

Since feature detection and matching are computationally intensive processes, the appropriate selection of overlapping regions between adjacent images can significantly improve the efficiency of image stitching. Taking spherical panoramas composed of fisheye cameras as the research context, Ref. [48] adopted a 16% overlap threshold. The work in Ref. [49] suggested a threshold of 15%. While Ref. [50] preferred an overlap in the range of 20%-25%. Considering both the FOV and full-coverage distance, Ref. [51] estimated the optimal threshold to be 32%. Although these methods vary in the choice of overlap percentage, all successfully produced panoramic images.

However, in UAV-based systems, constraints such as payload and power consumption necessitate an optimal and efficient definition of the overlap threshold. In our study, we captured four fisheye images from four fisheye cameras almost at the same time. Naturally, overlapping regions exist between any two adjacent images. Given the significant distortion present in fisheye images, and in order to obtain more reliable feature points (i.e., potential matching points), we selected a 40% overlap threshold, as illustrated in Figure 5. In the figure, starting from the left and right boundaries of the circular fisheye image, 40% of the horizontal imaging region is selected as the overlapping area for feature detection and matching.

3.4.2 Feature Matching

In feature matching algorithms, commonly used methods include SIFT [52], SURF [53], and ORB [54]. SIFT extracts features by applying a Gaussian filter and constructing Gaussian and Difference-of-Gaussian (DoG) functions. Although it offers high accuracy and robustness, its high-dimensional descriptors and computational complexity result in limited real-time performance. Ref. [52] proposed a method for extracting distinctive invariant features from images, which enables reliable matching between different views of an object or scene, and further demonstrated its application in object recognition. The work in Ref. [19] introduced a fisheye image stitching method based on colour calibration tools, effectively eliminating luminance differences between fisheye lenses. In panoramic vision systems, feature point-based matching and stitching methods [55] tend to be computationally expensive, while region-based methods [56] offer better real-time performance with reduced computational load [57].

Additionally, during the image stitching process, matching errors may accumulate progressively [58], leading to air facts such as waviness, tilt, or warping in the resulting panoramic image. Therefore, reducing matching errors is crucial for improving the overall stitching quality [59]. Figure 6 show a case where SURF is used to detect and match the feature points between adjacent images

Figure 5. Selection of overlapping regions.

Figure 6. Feature point matching.

Figure 7. Image stitching.

3.5 Image Stitching

After feature matching is performed within the overlapping regions, the matched feature points in adjacent images are identified and used for stitching, resulting in a seamless panoramic image as shown in Figure 7. The grey areas in the sky are caused by interference from the propellers, and are primarily located along the seams of adjacent images. Due to differences in focus or exposure settings between cameras, preprocessing the original images to ensure consistent resolution and brightness is beneficial for improving both real-time performance and stitching quality.

Depending on the hardware platform, 360° panoramas are typically obtained through either traditional image stitching [60-64] or fisheye image stitching methods [65-67]. For the former, conventional feature descriptors, such as SIFT, SURF, or ORB perform well with standard images but struggle to handle the significant distortion present in fisheye images. Ref. [65] introduced a method for aligning dual-fisheye images using interpolation grids based on rigid moving least squares, refining control points before alignment and achieving good results in case with mineral parallax. Ref. [66] proposed a stitching model that learns image distortion and incorporates professional image quality estimation to enhance stitching results. Similarly, Ref. [67] employed an image quality estimation approach to accelerate the stitching process by a factor of six. Ref. [68] proposed an Attention-based Nonlinear Activation Free Network to effectively deblur fisheye images.

Regardless of the specific method, image fusion remains the key step in the stitching process [69][70]. Most traditional stitching techniques rely on weighted average fusion, which often results in blurred transitions in the overlapping regions and fails to preserve complete object contours near stitching boundaries [71]. To address this, Ref. [59] presented an improved Laplacian multi-resolution fusion method based on the Min-Cut/Max-Flow algorithm, which achieved high-quality, seamless panoramic results. Ref. [72] introduced an algorithm based on discrete cosine transform (DCT) standards to fuse multi-focus images in wireless visual sensor networks. Furthermore, Ref. [73] proposed a weighted fusion algorithm based on the arc function, which provided smoother transitions, especially in the central region of stitched images.

4. OMNIDIRECTIONAL OBSTACLE DETECTION

Through the above processing steps, high-quality panoramic images have been successfully obtained. As the next stage, to ensure the safe flight of low-altitude UAVs in complex environments [74], the ability to recognize obstacles quickly and accurately becomes crucial. Currently, there are two primary approaches for achieving obstacle perception in low-altitude UAV systems.

4.1 Visual Ranging Methods

To obtain scene depth in the coordinate system of visual cameras, both active and passive vision techniques can be employed. Common approaches include structural light systems [75] and binocular disparity estimation [76]. Using depth data acquired from these methods, a local 3D obstacle map can be reconstructed, enabling obstacle localization and flight path planning. However, these techniques often suffer from limited accuracy and depth range. Moreover, visual algorithms tend to be computationally intensive, making real-time processing a significant challenge.

To address this, parallel computation using GPU hardware has been explored. Ref. [77] improved algorithm efficiency significantly by offloading obstacle image processing to GPU resources. Ref. [78] developed an optical detection system that effectively detects front-facing obstacles, providing information on distance, azimuth, and width in unknown environments. Ref. [79] proposed a depth perception method that combines omnidirectional imagery with structured light encoding to obtain full-scene depth information, while Ref. [80] presented a novel omnidirectional structured light system with a flexible configuration that overcomes the rigidity of traditional setups.

In terms of passive techniques, Ref. [81] introduced a motion parallax-based depth estimation algorithm, which utilizes both image sequences and known camera motion parameters. Ref. [82] developed BidNet, a binocular image dehazing Network based on deep learning, which enhances the visibility of stereo image pairs without relying on time-consuming disparity estimation. Ref. [83] proposed OmniVidar, a system employing four wide-FOV cameras to achieve high-resolution and high-speed depth estimation. Ref. [84] presented a cascade architecture using a dynamic spherical sweeping method to refine depth estimates progressively from coarse to fine. Lastly, Ref. [85] introduced a two-stage omnidirectional depth estimation framework that utilizes spherical feature learning to handle panoramic distortion, delivering robust results in both synthetic and real-world scenarios.

4.2 Optical Flow Methods

When objects in the environment move relative to a visual system, the system perceives changes over time. By processing this temporal information through neural networks or algorithmic models, the Time to Collision (TTC)—the time remaining before an impact occurs—can be estimated [86]. This concept forms the basis of TTC-based obstacle avoidance algorithms, which have been successfully applied in assisted driving systems. For instance, Mobile Eye products [87] utilize similar technologies to provide real-time collision warnings.

Applying TTC-based methods to low-altitude UAVs, however, presents two primary challenges:

4.2.1 Optical Flow Calculation Quality and Efficiency

Accurate TTC estimation relies heavily on reliable optical flow data. While the Lucas-Kanade method [88] is computationally efficient and widely used, it only calculates optical flow at sparse feature points, potentially missing important motion cues from obstacles, particularly in texture-less regions. On the other hand, dense optical flow algorithms—which provide pixel-level motion estimation—require intensive optimization procedures [89][90], making them unsuitable for real-time applications on CPU-constrained UAV platforms.

Recent advances in CNN-based dense optical flow estimation have outperformed traditional methods in terms of accuracy and robustness. These models can operate in real time when supported by high-performance GPU hardware. However, the deployment of such deep learning models on small UAVs is currently limited by the capabilities of onboard AI computing platforms. Ref. [91] introduced a phase-based method for optical flow estimation on omnidirectional images, demonstrating improved robustness and accuracy over conventional approaches. Ref. [92] proposed a wavelet-based method leveraging a decomposition of the Brightness Change Constraint Equation to efficiently compute optical flow in omnidirectional image sequences, effectively addressing the challenges of locally non-uniform flows.

4.2.2 Accurate TTC Estimation from Optical Flow

Once optical flow is computed, the next challenge is to accurately derive TTC values. Two main approaches are commonly used. Ref. [93] estimates TTC under the forward flight assumption, where the UAV is presumed to be moving straight ahead. Ref. [94] proposes calculating TTC by analysing the area scaling of local image regions, a method that is less dependent on the UAV's specific flight mode. While the second approach provides greater flexibility, both are approximate methods and often introduce significant estimation errors, especially in complex or dynamic environments.

4.3 Research and Application Status

With the continuous advancement of hardware technologies, the performance and computational efficiency of obstacle perception algorithms have significantly improved. Various approaches have emerged, each with unique strengths and limitations:

4.3.1 Obstacle Recognition Algorithms

Typically, there are three categories of obstacle recognition algorithms:

1. Traditional Feature-Based Methods
   An early approach involves the use of morphological Haar wavelets [95] to locate candidate obstacle regions, followed by the extraction of HOG features. These features are then classified using a Support Vector Machine (SVM). While this method provides a structured pipeline, the process of feature extraction is complex and prone to capturing ineffective or irrelevant features, which may reduce recognition accuracy.

2. Region-Based Convolutional Neural Networks (R-CNN)
   The second class of methods is based on region proposals, with the R-CNN series [96] being a representative example. Faster R-CNN [97], in particular, has demonstrated excellent performance in detecting small, overlapping, and partially occluded aerial targets under challenging conditions such as lighting variations and cluttered backgrounds. This makes it highly suitable for UAV applications.

3. One-Stage Detection Models (Without Region Proposal)
   A more recent trend is the development of single-shot detection models, such as the YOLO series [98], which bypass region proposal steps and perform object detection in a single forward pass. These models offer high speed and efficiency, making them well-suited for real-time onboard processing on UAV platforms.

4.3.2 Collaborative Perception in Multi-UAV Systems

In multi-UAV deployments, collaborative perception poses a significant challenge, particularly in terms of efficient communication and mutual localization. Each UAV must be aware of the positions and task progress of others in the formation. Several approaches to relative pose estimation have been proposed:

- Common Visual Feature Matching: Identifying shared environmental features visible to multiple UAVs [99].
- Visual Markers or Identification Tags: Placing markers or tags on UAV fuselages for easier recognition [100].
- Key Point Regression: Estimating the relative pose by regressing the key parts of other UAVs within an image [101].

Figure 8. Obstacles detection.

Despite their potential, these above methods face challenges related to accuracy, robustness, and real-time reliability, especially in dynamic or cluttered environments. Figure 8 shows the result of obstacle detection by YOLOv8. Based on the UAV's flight height, several potential obstacles are extracted from the panoramic image. It can be observed that there are cases of misjudgements, duplication and other issues in the extracted obstacles. Improving the accuracy of obstacle extraction will be one of our future research directions. Further research is required to enhance their performance and adaptability in practical scenarios.

5. CONCLUSION

5.1 Challenges

In summary, the panoramic vision-based obstacle avoidance system using fisheye cameras faces the following challenges:

1. Trade-off between perceptual range and resolution.
   Panoramic vision systems are designed to cover a 360-degree FOV, often using fisheye lenses or multiple camera arrays. However, this wide coverage often results in reduced resolution and a loss of detail, making it difficult to perceive distant obstacles accurately. On the other hand, using high-resolution cameras significantly increases the computational complexity of image processing, thereby affecting real-time performance.

2. Sensitivity to lighting conditions and adverse weather.
   Variations in lighting—such as strong light, shadows, and night flight—as well as adverse weather conditions—such as rain, fog, and strong winds—pose significant challenges to the stability and reliability of panoramic vision systems. For example, the Mavic 3, which uses omnidirectional obstacle sensing, performs poorly when encountering uniformly coloured surfaces, rapid light changes, or small-sized obstacles.

3. High computational demand.
   Panoramic vision involves extensive image processing tasks, including panoramic image stitching, obstacle detection, and depth estimation. These processes require substantial computational resources. However, the onboard computing power of small rotary-wing UAVs is limited, making it difficult to support complex image processing and deep learning inference. The delays caused by such processing make it challenging for the system to react promptly to dynamic obstacles.

4. Sensor failure and system robustness.
   During flight, cameras may experience failure due to vibration, physical damage, or occlusion. Such failures can compromise the robustness and safety of the obstacle avoidance system. Ensuring continued obstacle avoidance functionality in the event of partial sensor failure remains a critical and unresolved challenge

Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions.

5.2 Development Direction

It can be concluded that the development directions of panoramic vision systems based on fisheye cameras for UAVs are as follows:

1. Low-power, high-efficiency panoramic visual perception.
   The adoption of lightweight deep learning models, such as Tiny-YOLO and MobileNet, enables real-time obstacle detection under constrained computational resources. Additionally, the integration of transformer architectures or reinforcement learning algorithms can enhance the system's capability of dynamic scene prediction.

2. Improved adaptability to diverse environmental conditions.
   With the rapid development of machine vision hardware, high-resolution and wide dynamic range cameras are becoming more accessible, allowing better adaptation to complex lighting environments. Moreover, the use of novel vision sensors, such as event cameras, further enhances perception performance in low-light or high-speed scenarios.

3. Integration of multi-sensor intelligent systems.
   By combining visual sensors with liDAR, infrared sensor, IMUs and other sensory modalities, and applying data fusion techniques, such as Kalman filter or particle filter, the system can achieve higher fault tolerance, robustness, and overall reliability through sensor redundancy.

4. Enhance autonomy and environmental adaptability.
   Utilizing online learning and adaptive algorithms allow UAVs to dynamically adjust their perception and obstacle avoidance strategies in real-time based on environmental changes. The development of simulation platforms, such as AirSim and Gazebo enables effective training and testing in realistic virtual environments.

5. Cross-field technological collaboration, for intelligent and cooperative flight.
   The integration of advanced technologies from fields such as autonomous driving and robotics—including reinforcement learning, SLAM, and VIO (Visual-Inertial Odometry)—will further strengthen system intelligence. Additionally, leveraging 5G communication and edge computing can enable cloud-based collaborative obstacle avoidance and real-time path optimization.

In summary, panoramic vision-based obstacle avoidance systems using fisheye cameras in low-altitude rotary-wing UAVs are evolving toward greater intelligence, lightweight design, and strong environmental adaptability. Despite existing challenges, ongoing advancements in sensor technology, computational capabilities, and artificial intelligence are expected to drive significant progress in this field, making it increasingly vital to future UAV applications.

Based on the review presented in this paper, we will carry out practical research on real-time obstacle detection in unknown environments. With the continuous advancement of hardware and algorithmic development, it is anticipated that, in the near future, lightweight and flexible Low-altitude Rotorcraft UAVs will be equipped with intelligent "eyes" capable of perceiving and navigating complex environments autonomously

ACKNOWLEDGEMENT AND FUNDING

The authors receive no financial support for the research, authorship, and publication of this article.

DECLARATION OF CONFLICTING INTERESTS

The authors declare no potential conflicts of interest with respect to the research and publication of this article.

REFERENCES

[1] H. Y. Lee, H. W. Ho and Y. Zhou, Deep learning-based monocular obstacle avoidance for unmanned aerial vehicle navigation in tree plantations: faster region-based convolutional neural network approach, *Journal of Intelligent & Robotic Systems*, 101(5), 2021, 1-27.

[2] Y. Yu, W. Tingting, C. Long and Z. Weiwei, Stereo vision-based obstacle avoidance strategy for quadcopter UAV, *30th Chinese Control and Decision Conference*, Shenyang, China, 2018, 490-494.

[3] T. Hinzmann, C. Cadena, J. Nieto and R. Siegwart, Flexible trinocular: non-rigid multi-camera-IMU dense reconstruction for UAV navigation and mapping, *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Macau, China, 2019, 1137-1142.

[4] F. Liang, S. Kevin, K. Kunze and Y. S. Pai, PanoFlex: Adaptive panoramic vision to accommodate 360 field-of-view for humans, *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*, Parramatta, Australia, 2019, 1-2.

[5] M. Nieuwenhuisen, D. Droeschel, J. Schneider, D. Holz, T. Läbe and S. Behnke, Multimodal obstacle detection and collision avoidance for micro aerial vehicles, *2013 European Conference on Mobile Robots*, Barcelona, Spain, 2013, 7-12.

[6] R. Aggarwal, A. Vohra, A. M. Namboodiri, Panoramic stereo videos with a single camera, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, 2016, 3755-3763.

[7] H. Z. Yuan, B. P. Wang, J. Zhang and H. Li, A novel method for geometric correction of multi-cameras in panoramic video systems, *2010 International Conference on Measuring Technology and Mechatronics Automation*, Changsha, China, 2010, 248-251.

[8] O. Zia, J. H. Kim, K. Han and J. W. Lee, 360 panorama generation using drone mounted fisheye cameras, *2019 IEEE International Conference on Consumer Electronics*, Las Vegas, USA, 2019, 1-3.

[9] B. Akdemir, A. N. Belbachir and L. M. Svendsen, Real-time vehicle localization and tracking using monocular panomorph panoramic vision, *2018 24th International Conference on Pattern Recognition, Beijing*, China, 2018, 2350-2355.

[10] Y. Zhang, X. Xu, N. Zhang and Y. Lv, A semantic SLAM system for catadioptric panoramic cameras in dynamic environments, *Sensors*, 21(17), 2021, 5889.

[11] I. Stamenov, A. Arianpour, S. J. Olivas, I. P. Agurok, A. R. Johnson, R. A. Stack, R. L. Morrison and J. E. Ford, Panoramic monocentric imaging using fiber-coupled focal planes, *Optics Express*, 22(26), 2014, 31708-31721.
[12] C. Geyer and K. Daniilidis, Catadioptric camera calibration, *Proceedings of the 7th IEEE International Conference on Computer Vision*, Kerkyra, Greece, 1999, 398-404.
[13] S. Yi and N. Ahuja, An omnidirectional stereo vision system using a single camera, *18th International Conference on Pattern Recognition*, Hong Kong, China, 2006, 861-865.
[14] K. Tanaka and S. Tachi, Tornado: Omnistereo video imaging with rotating optics, *IEEE Transactions on Visualization and Computer Graphics*, 11(6), 2005, 614-625.
[15] C. Richardt, Y. Pritch, H. Zimmer and A. Sorkine-Hornung, Megastereo: Constructing high-resolution stereo panoramas, *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Portland, USA, 2013, 1256-1263.
[16] N. K. E. L. Abbadi, S. A. Al Hassani and A. H. Abdulkhaleq, A review over panoramic image stitching techniques, *Journal of Physics: Conference Series*, 1999(1), 2021, 012115.
[17] C. Arth, M. Klopschitz, G. Reitmayr and D. Schmalstieg, Real-time self-localization from panoramic images on mobile devices, *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, Basel, Switzerland, 2011, 37-46.
[18] S. Peleg, M. Ben-Ezra and Y. Pritch, Omnistereo: Panoramic stereo imaging, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, 2001, 279-290.
[19] W. Ye, K. Yu, Y. Yu, and J. Li, Logical stitching: A panoramic image stitching method based on color calibration box, *2018 14th IEEE International Conference on Signal Processing*, Beijing, China, 2018, 1139-1143.
[20] A. S. Amini, M. Varshosaz and M. Saadatseresht, Evaluating a new stereo panorama system based on stereo camera, *International Journal of Scientific Research in Inventions and New Ideas*, 2(1), 2014, 1-10.
[21] Y. Pritch, M. Ben-Ezra and S. Peleg, Optics for omnistereo imaging, *Foundations of Image Understanding*, 2011, 447-467.
[22] PanoramaStudio, https://www.tshsoft.com/en/index, 2025 (accessed 19.01.2025).
[23] PTGui, https://ptgui.com/, 2025 (accessed 19.01.2025).
[24] Hugin-Panorama photo stitcher, https://hugin.sourceforge.io/, 2025 (accessed 19.01.2025).
[25] Y. Hou, L. P. Niu, Y. M. Zhao and S. Z. Lan, Fisheye images correction based on different angle of views, *2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference*, Chongqing, China, 2020, 854-856.
[26] G. Krishnan and S. K. Nayar, Cata-fisheye camera for panoramic imaging, *2008 IEEE Workshop on Applications of Computer Vision.* Copper Mountain, CO, USA, 2008, 1550-5790.
[27] Aerial Pro 360, https://diydrones.com/members/AerialPro360, 2025 (accessed 19.01.2025).
[28] RICOH360, https://www.ricoh360.com/theta/, 2025 (accessed 19.01.2025).
[29] I. C. Lo, K. T. Shih and H. H. Chen, Efficient and accurate stitching for 360° dual-fisheye images and videos, *IEEE Transactions on Image Processing*, 31, 2022, 251-262.
[30] V. Chapdelaine-Couture and S. Roy, The omnipolar camera: A new approach to stereo immersive capture, *IEEE International Conference on Computational Photography*, Cambridge, MA, USA, 2013, 1-9.
[31] H. Cheng, C. Xu, J. Wang and L. Zhao, Quad-fisheye image stitching for monoscopic panorama reconstruction, *Computer Graphics Forum*, 41(6), 2022, 94-109.
[32] O. Zia, J. H. Kim, K. Han and J. W. Lee, 360 panorama generation using drone mounted fisheye cameras, *2019 IEEE International Conference on Consumer Electronics*, Las Vegas, NV, USA, 2019, 1-3.
[33] J. Kannala and S. S. Brandt, A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 28(8), 2006, 1335-1340.
[34] Q. Fu, K. Y. Cai and Q. Quan, Calibration of multiple fish-eye cameras using a wand, *IET Computer Vision*, 9(3), 2015, 378-389.
[35] Fisheye Calibration Basics, https://www.mathworks.com/help/vision/ug/fisheye-calibration-basics.html, 2025 (accessed 19.01.2025).
[36] S. Chan, X. Zhou, C. Huang, S. Chen and Y. F. Li, An improved method for fisheye camera calibration and distortion correction, *2016 International Conference on Advanced Robotics and Mechatronics*. Macau, China, 2016, 579-584.
[37] N. Wakai and T. Yamashita, Deep single fisheye image camera calibration for over 180-degree projection of field of view, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, 1174-1183.
[38] N. Wakai, T. Azuma and K. Nobori, Multiple fisheye camera calibration and stereo measurement methods for uniform distance errors throughout imaging ranges, *17th International Conference on Machine Vision and Applications*, Aichi, Japan, 2021, 1-5.
[39] G. H. Babu and N. Venkatram, A survey on analysis and implementation of state-of-the-art haze removal technique, *Journal of Visual Communication and Image Representation*, 72, 2020, 102912.
[40] H. Wang, Q. Xie, Q. Zhao and D. Meng, A model-driven deep neural network for single image rain removal, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, 3103-3112.
[41] M. Li, X. Cao, Q. Zhao, L. Zhang and D. Meng, Online rain/snow removal from surveillance videos, *IEEE Transactions on Image Processing*, 30, 2021, 2029-2044.
[42] R. Wang, D. Zou, C. Xu, L. Pei, P. Liu and W. Yu, An aerodynamic model-aided state estimator for multi-rotor UAVs, *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vancouver, BC, Canada, 2017, 2164-2170.
[43] H. Zhu, X. Yin and J. Zhou, A cubic polynomial model for fisheye camera, *14th International Conference on Human-Computer Interaction: Design and Development Approaches*, Orlando, FL, USA, 2011, 684-693.

[44] Y. Chang, D. Bailey and S. L. Moan, Lens distortion correction by analysing peak shape in Hough transform space, *2017 International Conference on Image and Vision Computing New Zealand*, Christchurch, New Zealand, 2017, 1-6.
[45] G. Zhou, H. Li, R. Song, Q. Wang, J. Xu and B. Song, Orthorectification of fisheye image under equidistant projection model, *Remote Sensing*, 14(17), 2022, 4175.
[46] V. Usenko, N. Demmel and D. Cremers, The double sphere camera model, *2018 International Conference on 3D Vision*, Verona, Italy, 2018, 552-560.
[47] J. Xu, D. W. Han, K. Li, J. J. Li and Z. Y. Ma, A comprehensive overview of fish-eye camera distortion correction method, *arXiv preprint*, arXiv:2401.00442, 2023.
[48] J. Kopf, M. Uyttendaele, O. Deussen and M. F. Cohen, Capturing and viewing gigapixel images, *Association for Computing Machinery Transactions on Graphics*, 26(3), 2007, 93-es.
[49] 8 Guidelines to taking panoramic photos with any camera, https://digital-photography-school.com/8-guidelines-to-taking-panoramic-photos-with-any-camera/, 2025 (accessed: 19.01.2025).
[50] 360 DJI Drone Panorama DroneBlocks, http://www.hdrpano.ch/index_htm_files/DroneBlocks.pdf, 2025 (accessed: 19.01.2025).
[51] A. Akin, O. Cogal, K. Seyid, H. Afshari, A. Schmid and Y. Leblebici, Hemispherical multiple camera system for high resolution omni-directional light field imaging, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 3(2), 2013, 137-144.
[52] D. G. Lowe, Distinctive image features from scale-invariant keypoint, *International Journal of Computer Vision*, 60(2), 2004, 91-110.
[53] H. Bay, T. Tuytelaars and L. Van Gool, SURF: Speeded up robust features, *9th European Conference on Computer Vision*, Graz, Austria, 2006, 404-417.
[54] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, ORB: An efficient alternative to SIFT or SURF, *2011 International Conference on Computer Vision*, Barcelona, Spain, 2011, 2564-2571.
[55] J. Zaragoza, T. J. Chin, M. S. Brown and D. Suter, As-projective-as-possible image stitching with moving DLT, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, 2339-2346.
[56] Z. Wang, B. Fan, G. Wang and F. Wu, Exploring local and overall ordinal information for robust feature description, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(11), 2016, 2198-2211.
[57] B. He and S. Yu, Parallax-robust surveillance video stitching, *Sensors*, 16 (1), 2015, 7.
[58] C. C. Lin, S. U. Pankanti, K. Natesan Ramamurthy and A. Y. Aravkin, Adaptive as-natural-as possible image stitching, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA, 2015, 1155-1163.
[59] Image seamless stitching and straightening based on the image block, https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/iet-ipr.2017.1064, 2025 (accessed: 19.01.2025).
[60] F. Perazzi, A. Sorkine-Hornung, H. Zimmer, P. Kaufmann, O. Wang, S. Watson and M. Gross, Panoramic video from unstructured camera arrays, *Computer Graphics Forum*, 34(2), 2015, 57-68.
[61] W. Y. Lin, S. Liu, Y. Matsushita, T. T. Ng and L. F. Cheong, Smoothly varying affine stitching, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Colorado Springs, CO, USA, 2011, 345-352.
[62] W. Jiang and J. Gu, Video stitching with spatial-temporal content preserving warping, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Boston, Massachusetts, USA, 2015, 42-48.
[63] Y. Nie, T. Su, Z. Zhang, H. Sun and G. Li, Dynamic video stitching via shakiness removing, *IEEE Transactions on Image Processing*, 27(1), 2018, 164-178.
[64] A. Hamza, R. Hafiz, M. M. Khan, Y. Cho and J. Cha, Stabilization of panoramic videos from mobile multi-camera platform, *Image and Vision Computing*, 37, 2015, 20-30.
[65] T. Ho, I. D. Schizas, K. R. Rao and M. Budagavi, 360-degree video stitching for dual-fisheye lens cameras based on rigid moving least squares, *2017 IEEE International Conference on Image Processing*, Beijing, China, 2017, 51-55.
[66] J. Li, Y. Zhao, W. Ye, K. Yu and S. Ge, Attentive deep stitching and quality assessment for 360° omnidirectional images, *IEEE Journal of Selected Topics Signal Processing*, 14(1), 2019, 209-221.
[67] A. Utter, Dual-Fisheye Image Stitching Tool, https://github.com/ooterness/DualFisheye, 2025 (accessed: 19.01.2025).
[68] J. Hao, J. Xie, J. Zhang and M. Liu, A stronger stitching algorithm for fisheye images based on deblurring and registration, *IEEE Sensors Letters*, 7(10), 2023, 1-4.
[69] C. Anita, Image fusion methods and applications: A review, *Journal of Innovation and Technology*, 14, 2023, 1-8.
[70] H. Kaur, D. Koundal and V. Kadyan, Image fusion techniques: A survey, *Archives of computational methods in Engineering*, 28(7), 2021, 4425-4447.
[71] F. Zhang and F. Liu, Parallax-tolerant image stitching, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, USA, 2014, 3262-3269.
[72] K. S. Krishnendu, Multi-focus image fusion based on spatial frequency (SF) and consistency verification (CV) in DCT domain, *arXiv e-prints,* arXiv: 2305.11265, 2023.
[73] L. Xue, J. Zhu, H. Zhang and R. Liu, A high-quality stitching algorithm based on fisheye images, *Optik*, 238, 2021, 166520.
[74] L. Zhao, *3D Obstacle Avoidance for Unmanned Autonomous System (UAS)*, Master of Science in Engineering, University of Nevada, Las Vegas, USA, 2015.
[75] J. P. Angelo, S. J. Chen, M. Ochoa, U. Sunar, S. Gioux and X. Intes, Review of structured light in diffuse optical imaging, *Journal of Biomedical Optics*, 24(7), 2019, 071602-071602.
[76] S. Ryoka and O. Hirotsugu, Binocular disparity estimation algorithm using multiple spatial frequency information and a neural network, *ALife Robotics*, 28, 2023, 536-540.

[77] C. Nian, W. Kaihua and W. Wenjie, Obstacle detection system of plant protection UAVs based on structural light, *Journal of Applied Optics*, 39(3), 2018, 343-348.
[78] K. H. Wu and W. J. Wang, Detection method of obstacle for plant protection UAV based on structured light vision, *Opto-Electronic Engineering*, 45(4), 2018, 170613.
[79] T. Jia, B. N. Wang, Z. X. Zhou and H. Meng, Scene Depth perception based on omnidirectional structured light, *IEEE Transactions on Image Processing*, 25(9), 2016, 4369-4378.
[80] C. Paniagua, L. Puig and J. J. Guerrero, Omnidirectional structured light in a flexible configuration, *Sensors*, 2013, 13, 13903-13916.
[81] M. Mansour, P. Davidson, O. Stepanov and R. Piché, Relative importance of binocular disparity and motion parallax for depth estimation: A computer vision approach, *Remote Sensing*, 11(17), 2019, 1990.
[82] Y. Pang, J. Nie, J. Xie, J. Han and X. Li, BidNet: Binocular image dehazing without explicit disparity estimation, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, 5931-5940.
[83] S. Xie, D. Wang, Y. H. Liu, OmniVidar: Omnidirectional depth estimation from multi-fisheye images, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, Canada, 2023, 21529-21538.
[84] P. Wang, M. Li, J. Cao, S. Du and Y. Li, CasOmniMVS: Cascade omnidirectional depth estimation with dynamic spherical sweeping, *Applied Sciences*, 14(2), 2024, 517.
[85] M. Li, X. Jin, X. Hu, J. Dai, S. Du and Y. Li, MODE: Multi-view omnidirectional depth estimation with 360° cameras, *European Conference on Computer Vision*, 2022, 197-213.
[86] J. M. Galbraith, G. T. Kenyon and R. W. Ziolkowski, Time-to-collision estimation from motion based on primate visual processing, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8), 2005, 1279-1291.
[87] Mobileye, https://www.mobileye.com/, 2025 (accessed 19.01.2025).
[88] S. Baker and I. Matthews, Lucas-Kanade 20 years on: A unifying framework, *International Journal of Computer Vision*, 56(3), 2004, 221-255.
[89] B. K. P. Horn and B. G. Schunck, Determining optical flow, *Artificial intelligence*, 17(1-3), 1981, 185-203.
[90] M. Menze, C. Heipke and A. Geiger, Discrete optimization for optical flow, *37th German Conference on Pattern Recognition (GCPR 2015)*, Aachen, Germany, 2015, 16-28.
[91] B. Alibouch, A. Radgui, M. Rziza and D. Aboutajdine, Optical flow estimation on omnidirectional images: An adapted phase based method, *5th International Conference on Image and Signal Processing (ICISP 2012)*, Agadir, Morocco, 2012, 468-475.
[92] C. Demonceaux and D. Kachi-Akkouche, Optical flow estimation in omnidirectional images using wavelet approach, *2003 Conference on Computer Vision and Pattern Recognition Workshop*, Madison, WI, USA, 7, 2003, 76-76.
[93] Q. Quan, *Introduction to Multicopter Design and Control*, Singapore: Springer, 2017, 978-981.
[94] G. Yang and D. Ramanan, Upgrading optical flow to 3D scene flow through optical expansion, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, 1334-1343.
[95] J. Chen, J. L. Niu and D. H. Chen, Research on intelligent wheelchair obstacle avoidance based on AdaBoost, *Applied Mechanics and Materials*, 312, 2013, 685-689.
[96] R. Girshick, J. Donahue, T. Darrell and J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, Columbus, Ohio, 2014, 580-587.
[97] A. Salvador, X. Giró-i-Nieto, F. Marqués and S. Satoh, Faster R-CNN features for instance search, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Las Vegas Nevada, 2016, 9-16.
[98] W. Wu, H. Liu, L. Li, Y. Long, X. Wang, Z. Wang, J. Li and Y. Chang Y, Application of local fully convolutional neural network combined with YOLOv5 algorithm in small target detection of remote sensing image, *PLOS One*, 16(10), 2021, e0259283.
[99] D. Zou and P. Tan, Coslam: Collaborative visual slam in dynamic environments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(2), 2012, 354-366.
[100] X. Yan, H. Deng and Q. Quan, Active infrared coded target design and pose estimation for multiple objects, *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Macau, China, 2019, 6885-6890.
[101] M. Pavliv, F. Schiano, C. Reardon, D. Floreano and G. Loianno, Tracking and relative localization of drone swarms with a vision-based headset, *IEEE Robotics and Automation Letters*, 6(2), 2021, 1455-1462.