

Review

Retina-like Imaging and Its Applications: A Brief Review

Qun Hao^{1,2}, Yu Tao^{1,3}, Jie Cao^{1,2,*}, Mingyuan Tang¹, Yang Cheng^{1,2}, Dong Zhou¹, Yaqian Ning¹, Chun Bao^{1,2} and Huan Cui¹

¹ Key Laboratory of Biomimetic Robots and Systems, Ministry of Education, Beijing Institute of Technology, Beijing 100081, China; qhao@bit.edu.cn (Q.H.); 3220185052@bit.edu.cn (Y.T.); 3120190638@bit.edu.cn (M.T.); yangcheng2007@163.com (Y.C.); 3120195343@bit.edu.cn (D.Z.); daydayupnyq@163.com (Y.N.); baochun@bit.edu.cn (C.B.); 3120190619@bit.edu.cn (H.C.)

² Yangtze Delta Region Academy, Beijing Institute of Technology, Jiaxing 314003, China

³ System Engineering Institute, Academy of Army Research, Chinese People's Liberation Army, Beijing 100039, China

* Correspondence: caojie@bit.edu.cn

Abstract: The properties of the human eye retina, including space-variant resolution and gaze characters, provide many advantages for numerous applications that simultaneously require a large field of view, high resolution, and real-time performance. Therefore, retina-like mechanisms and sensors have received considerable attention in recent years. This paper provides a review of state-of-the-art retina-like imaging techniques and applications. First, we introduce the principle and implementing methods, including software and hardware, and describe the comparisons between them. Then, we present typical applications combined with retina-like imaging, including three-dimensional acquisition and reconstruction, target tracking, deep learning, and ghost imaging. Finally, the challenges and outlook are discussed to further study for practical use. The results are beneficial for better understanding retina-like imaging.



Citation: Hao, Q.; Tao, Y.; Cao, J.; Tang, M.; Cheng, Y.; Zhou, D.; Ning, Y.; Bao, C.; Cui, H. Retina-like Imaging and Its Applications: A Brief Review. *Appl. Sci.* **2021**, *11*, 7058. <https://doi.org/10.3390/app11157058>

Academic Editor: Oscar Reinoso García

Received: 30 June 2021
Accepted: 27 July 2021
Published: 30 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: retina-like imaging; three-dimensional reconstruction; deep learning; ghost imaging

1. Introduction

Many applications benefit from bioinspired optical methods. Here, some typical examples are given for illustration. Artificial compound eye inspired from insects has an extremely large field of view (FOV), low aberration and distortion, high temporal resolution, and an infinite depth of field [1–3]. Many artificial compound eye imaging systems have been proposed and used in many applications, such as medical diagnosis [4], navigation [5], and egomotion estimation [6]. The optical systems inspired by lobster eye provide several advantages over traditional system, such as wide field and energy acquisition ability in high-energy radiation field due to the remarkable structure of spherical microchannel [7,8]. Therefore, the lobster eye optical system can be used in the imaging applications where the wavelength includes X-rays, light waves, and infrared bands [9].

The eye is an important organ for learning about the environment for humans. We can focus on near or far scenes by the use of our crystalline lens, that is, different depths. We can also focus on interesting targets in our FOV, that is, the different resolutions in retina. Liquid lenses, as inspired by the human eye, simulate the properties of a crystalline lens [10]. The space-variant pixels of imaging sensors also simulate the properties of retina. Imaging sensors, including two-dimensional (2D) or three-dimensional (3D) ones, play an important role in national defense and civilian use [11–13], such as smart monitoring, robotic navigation, mobile phones, AR/VR, and automatic driving. Therefore, imaging sensors based on retina-like mechanisms are suitable for applications that simultaneously require large FOVs, high resolutions, and real-time performance. The development of computing and electronics has facilitated the study on the imaging system based on bioinspired vision from theory to practical use.

The human eye differs from compound eye mainly in the space-variant distribution of photoreceptor cells in the retina [14]. The interested region is assigned by high resolution, while the uninterested region is assigned a low resolution. Meanwhile, the transformation from retina to cortical approximately conforms to log-polar law, which results in good feature of scaling and rotation invariance. Many researchers have explored the properties of retinas and developed retina-like imaging theory and systems. Active and passive are the two main kinds of retina-like imaging systems according to the use of light. This review focuses on the state-of-the-art studies on retina-like imaging, including the implementations, applications, and challenges. The results are beneficial for understanding the properties of retina-like imaging and enable the study on the high performances of imaging systems. The rest of the review is organized as follows. In Section 2, we start from the principle of retina-like imaging and describe the mathematical model and advantages of this imaging method. Then, we propose the development process of the retina-like imaging and introduce it in detail from the two aspects of software and hardware. In Section 3, we discuss in detail the advantages of retina-like imaging in practical applications through four representative applications. In Section 4, we further discuss the retina-like imaging and once again point out the future development trends and the strong potential application value of this method.

2. Principle and Methods

2.1. Principle and Properties

The principle of retina-like imaging includes two aspects, namely, space-variant sampling and log-polar transformation (LPT) [15], as shown in Figure 1. It is written as Equation (1) [16]:

$$\begin{cases} \theta = \arctan \frac{y}{x} \\ \zeta = \log_q \sqrt{x^2 + y^2} \end{cases} \quad (1)$$

where (x,y) are the Cartesian coordinates and (ζ,θ) are the log-polar coordinates.

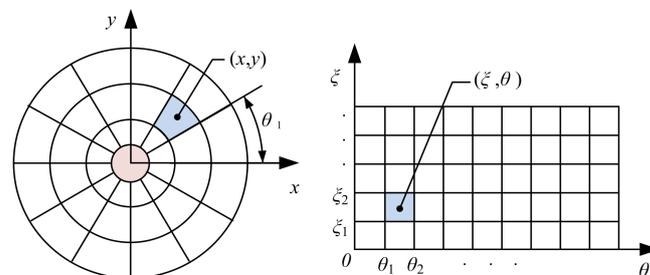


Figure 1. Mapping from Cartesian coordinates to log-polar coordinates.

Some methods of optimizations are added on the basis of log-polar coordinates to ensure the coverage rate of the FOV. Circle spots are used to sample the image, and the sampling between rings is tangent. Accordingly, the space-variant retina sampling model is obtained. The mathematical model is written as Equation (2) [17]:

$$\begin{cases} q = \frac{1+\sin(\pi/N)}{1-\sin(\pi/N)} \\ r_1 = \frac{r_0}{1-\sin(\pi/N)} \\ R_1 = \frac{r_0 \sin(\pi/N)}{1-\sin(\pi/N)} \\ r_i = r_1 \cdot q^{i-1} \\ R_i = q^{i-1} \cdot R_1 \\ r_{\max} = r_0 \cdot q^M \end{cases} \quad (2)$$

where M and N are the number of rings and number of sectors per ring of the retina-like imaging, respectively. r_0 is the radius of blind hole, q is the growth rate between rings, r_1

is the radius of the first ring, R_1 is the radius of the spot in the first ring, and r_{max} is the maximum radius of the FOV, which is tangent with the spots in the last ring, as shown in Figure 2.

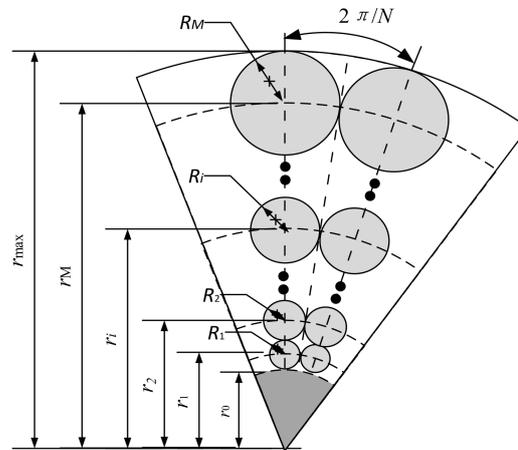


Figure 2. The sample model of retina-like imaging.

Retina-like imaging has three properties. First, it is beneficial for balancing among high resolution, large FOV, and real-time performance. The two axes of log-polar coordinates represent the distance from a certain point to an original point ζ and an angle θ from a certain point to the polar axis, respectively. When the ring number is low, the pixel is near the center of the FOV, and the resolution density is higher than the Cartesian coordinate that brings details to the target details. However, when the ring number is large, the resolution density becomes sparse, and it represents a lower resolution background area. Therefore, the retina-like structure allows for the efficient suppression of redundant data. However, this suppression image makes the pixels uncertain where a fuzzy edge detector would be necessary [18,19]. Second, rotation and scaling invariance are considered. The rotation and scaling of the target in the Cartesian coordinate system only affects the axes θ and ξ , respectively, as represented by Equation (3) [20]:

$$\begin{cases} \theta' = \theta + \alpha \\ \zeta' = \log_z(\sqrt{(qx)^2 + (qy)^2}) = \log_z(qr) \\ = \log_z(q) + \log_z(r) = \zeta + \log_z(r) \end{cases}, \quad (3)$$

where α stands for the rotation and q stands for the scale change. θ and ζ stand for the two axes of LPT, respectively. In this way, the relationship of rotation and scaling is rewritten by the shift in the θ and ξ directions, and it is called the rotation and scale change invariance [20]. Third, the adjustable optical axis is given attention. On the basis of the space-variant resolution, we always pay attention to the most interesting region. Meanwhile, we can transfer the interesting region according to different environments. With the development of the theory on retina-like imaging, retina-like imaging models do not always strictly conform to LPT, but the space-variant resolution is the typical common feature of retina-like imaging. Therefore, space-variant resolution is the kernel character and is illustrated in the following methods (Section 2.2) and applications (Section 3).

2.2. Methods

Many researchers have studied the realization of retina-like imaging from hardware and software. The realization based on hardware can be divided into two kinds of approaches: passive and active. Passive retina-like imaging methods selectively collect the target information, and active retina-like imaging methods explore the target information purposively. The two methods realize the retina-like imaging with specially designed

sensors. However, the software-based method uses LPT algorithm directly on the Cartesian imaging sensors to simplify the processing procedure.

The passive method of retina-like imaging has been used since 1990. Tistarellia et al. [21] designed a retina-like CCD sensor to search for targets in a wide FOV. A total of 30 rings with 64 sectors are present per ring. The size of these 1920 pixels increases from 30 μm to 412 μm . In 2000, Sandini et al. [22] realized the retina-like CMOS sensors with 8013 pixels located with log-polar algorithm and 845 pixels in the fovea. Compared with CCD, CMOS is more controllable in data storage and easier to interface with micro-processors; thus, it is more suitable for simulating the acquisition of human eye retina information. In 2014, our group [23] used nonuniformity lens array to collect light on the light-sensitive materials with log-polar configuration. This method shifts the difficulty to the lens array design, which makes it more flexible to adjust to different situations. In 2016, the camera array [24] and dual-channel foveated imaging systems [25] were proposed by Santacana G et al. and Xu C et al., respectively. A camera array formed by prisms and cameras combines the advantages of compound eyes and realizes high-resolution fovea and low-redundancy periphery. It reconstructs the super resolution for imaging the center with repeated stack of multiple cameras. The all-reflective, dual-channel foveated imaging system realizes the retina-like variable-resolution imaging with double barrels and double sensors. It is a combination of two different focus imaging systems with common light channels. Compared with a divided camera, it has a compactable structure, but subsequent image processing is needed to form the final image. This method was improved in 2017 by Guillem et al. [26]. An imaging system with dual-aperture optics that superimpose two images on a single sensor was designed to enable arbitrary magnification ratios using a relatively simple system. This method fully utilizes the sensor's pixels, which is important in applications where the cost of each pixel is high. In the same year, a micro-lens array was used to form the foveated imaging [27]. A 3D-printed micro-lens with different focus lengths in a 2×2 arrangement was located directly on the CMOS sensors. It achieves a full FOV of 70° with an increasing angular resolution of up to two cycles/deg FOV in the center of the image. Different from the methods above, a foveated imaging system with reflective spatial light modulator was proposed by Wang et al. [28]. It corrects the aberration of the region of interest (ROI) while the resolution of the other area is still very low. This method adjusts the fovea location with different phase patterns. In 2019, Wang et al. [29] proposed a foveated imaging system with a liquid crystal lens. This method modulates light selectively with a liquid crystal lens of variable focal length. In 2020, Cao, J.J. et al. [30] realized the bioinspired, zoom eye-enabled, variable-focus imaging with a deformable polydimethylsiloxane lens array. Similar to the lens of human eyes, it images the target with different distances clearly. Its FOV reaches 180° , and its focal length can be tuned from 3.03 mm to infinity with an angular resolution of 3.86×10^{-4} rad.

Although the passive methods effectively select and receive target information, the efficiency can be further improved if the required information can be extracted directly. Therefore, the active methods of retina-like imaging have been proposed. In 2017, adaptive, foveated, single-pixel imaging was proposed by David et al. [31]. Given that traditional single-pixel imaging requires a large number of measurements for image reconstruction, the target of interest is reconstructed in high-definition with the help of the human eye gaze function, which improves the reconstruction efficiency. Our group combined the retina-like structure with light detection and ranging (LiDAR). The space-variant scanning based on a micro-electro-mechanical system (MEMS) [17] and an optical phased array [32] were proposed. Compared with traditional LiDAR imaging, this retina-like structure maintains the high resolution in the fovea while enlarging the FOV. Importantly, these methods are simply an improvement on the scanning strategy and need no additional devices.

For software-based method, the algorithm model construction inspired by the human retina has attracted the attention of many researchers. In 2008, Paolo et al. [33] proposed a 3D reconstruction method based on the human retina, and it relies on the rotation and scaling invariance of log-polar transform to rebuild the view of room. In the same year,

Wong et al. [34] used log-polar transform to realize panoramic imaging with high efficiency based on FPGA. It decomposes the 256×256 omnidirectional image into a 128×128 panoramic image in the logarithmic polar coordinate. In 2016, Cheung et al. [35] used the visual attention mechanism of the human eye to solve the problem of high-precision recognition of small targets. Compared with the fixed sampling model, this method reduces the error rate of recognition by half.

Figure 3 shows the summary of retina-like imaging methods. The software methods are the least difficult to implement and the most tightly integrated with the other features, such as attention mechanism or omnidirectional imaging. However, more computation cost is required because this method is a further processing of the redundant information. Meanwhile, the information has been screened in the imaging process of the hardware methods, including passive and active methods. In this way, imaging has low redundancy while having increased technical difficulty. Through vertical comparison, more functional retina-like imaging becomes the central issue of bionic imaging. Especially after 2017, fovea gaze control, lens-like longitudinal focus adjustment, and multi-center imaging are combined in the retina-like imaging and achieve good effect.

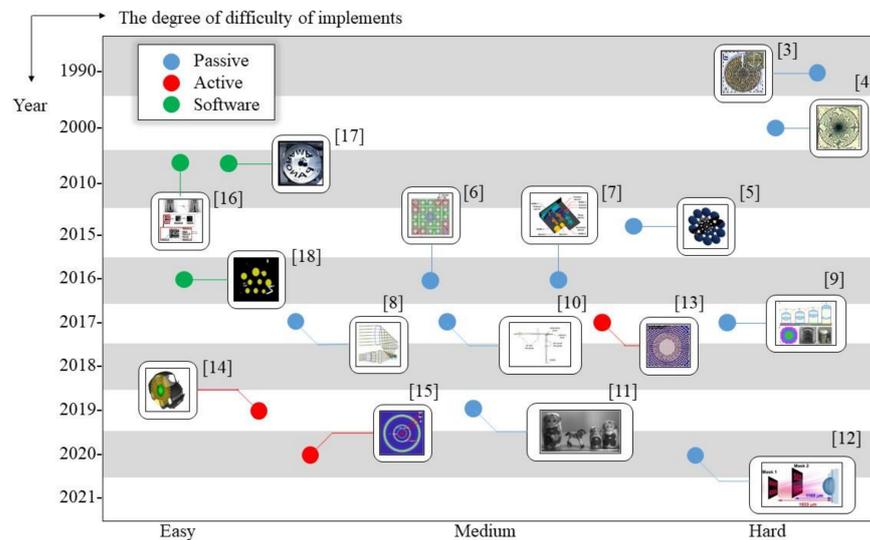


Figure 3. Implementation methods of retina-like imaging.

The comparison of these methods is shown in Table 1. Passive methods, which require more improvement on the imaging devices, have the highest difficulty. Meanwhile, active methods also compress the data volume, and the traditional devices are still suitable for retina-like imaging. Therefore, the difficulty in the space-variant scanning strategy is the requirement of more cooperation between imaging elements. The software methods, which are a processing of the imaging, have simple structure and can be flexibly used in different situations. The disadvantage of poor real-time performance will be alleviated with the improvement in computing power.

Table 1. Comparison of different methods.

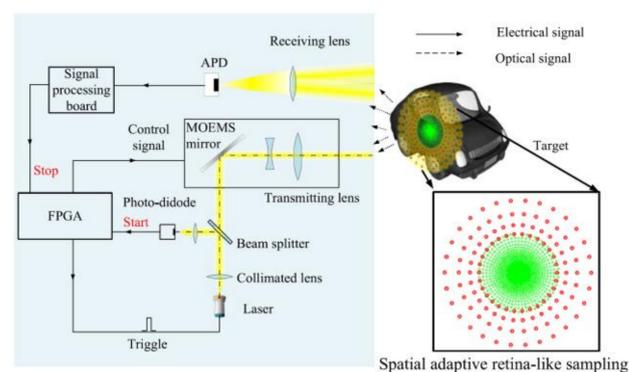
Implementation Methods	Hardware		Software
	Passive	Active	
Characteristics	Selectively receive image information	Purposeful information collection	Data compression after imaging
Advantages	Realize local high resolution, good scalability	Good real-time performance, high sensitivity, low data volume	Simple structure, easy to dock with existing equipment
Disadvantages	More improvement on imaging devices	More equipment cooperation	Poor real-time performance
Level of difficulty	Difficult	Medium	Easy

3. Applications

The above-mentioned methods have been used in many applications. For example, 2D or 3D imaging methods are the direct applications of retina-like sampling. With the development of deep learning and novel imaging methods, retina-like mechanisms are combined and used to obtain better performances than traditional methods. Here, we select several typical applications combined with retina-like imaging for illustration.

3.1. Three-Dimensional Acquisition and Reconstruction

The technology of 3D image acquisition and reconstruction describes the real scenario into the mathematical model, which conforms to the logic expression of computers, and these models play an auxiliary role in many research fields, including historical preservation [36], game development [37], architectural design [38], and clinical medicine [39,40]. The emphasis of 3D reconstruction technology is to obtain the target scenario or the depth and texture information of the object. The methods of retina-like technology result in a great number of interests and considerable attention owing to it fully combining the advantages of variable resolution, anti-rotation, and scale transformation. Figure 4 shows the retina-like LiDAR. It obtains the 3D figure using the time-of-flight principle, and the retina-like adaptive spatial invariance scanning method makes it more flexible.

**Figure 4.** Retina-like LiDAR based on MEMS.

The 3D reconstruction is conducted by the use of LPT [33], and the original input image (Figure 5a) is divided into different regions (Figure 5b) according to linear detection in log-polar coordinates. The main line of (Figure 5a) is found through image processing, and it is divided into front wall, right wall, ceiling, left wall, floor, and right wall. Then, the obtained line is mapped to the inverse log-polar coordinates and fused with the original image. The optical center of the image is acquired via finding the vanishing point (i.e., the intersection points of the line). Finally, the depth information of the image is extracted according to the position of the optical center, and the 3D image reconstruction is realized.

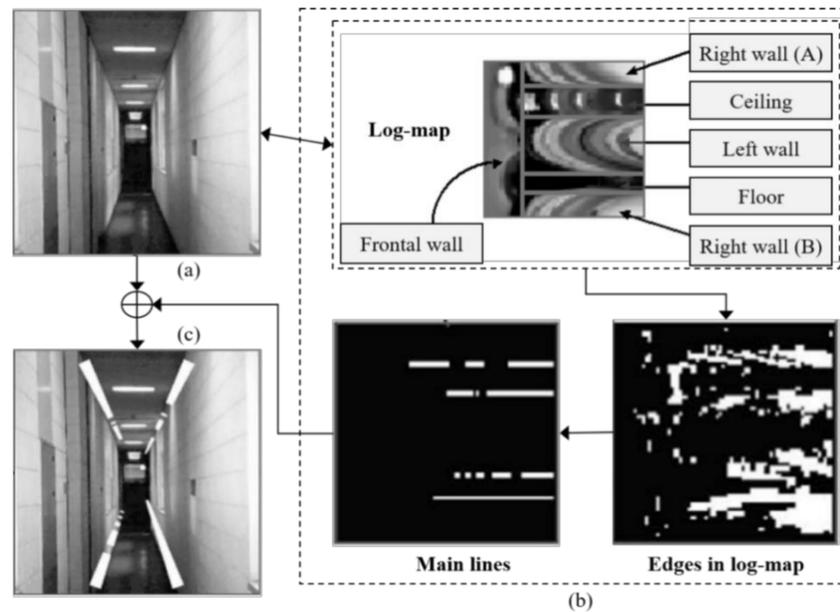


Figure 5. Information extraction of the image depth. (a) The original input image. (b) Linear detect process with LPT. (c) The output image.

Lee et al. [41] proposed a robust adaptive focusing measurement operator based on the biological inspiration, data selection, and edge in-variance characteristics of LPT, which was widely used in 3D shape restoration. The signal-to-noise ratio (SNR) of the image is improved by the use of LPT, and a better corresponding focus plane is obtained, which determines 3D shapes more accurately. Experiments on simulated target and real target image sequences indicated that the robust adaptive focusing measurement operator is effective in the presence of various types of noise, including high noise variance or strong noise density. Image registration is a necessary part of 3D reconstruction. Masuma et al. [42] studied a multi-modal image registration algorithm based on LPT, which conducts in-plane translation, rotation, and out-plane translation by using LPT. Large initial displacement registration of 3D CT images and 2D single-plane fluoroscopic images is successfully performed. Ravi et al. [43] built a scale, rotation, and translation-invariant descriptor based on the LPT of Gabor filter derivatives for the automatic co-registration of 3D multi-sensor point clouds. The co-registration process using data from a study area in Toronto, Ontario, Canada, which consists of a building surrounded by vegetation, bare land, paved roads, and parking lots, is shown in Figure 6. The source and target point clouds are collected from unmanned aerial vehicles and mobile laser scanning, respectively. After extracting multi-scale keypoints from 3D point cloud data, the keypoint descriptors with constant scaling, rotation, and translation are generated on the basis of the scaling and rotation invariance of log-polar coordinate strategy and derivative mapping calculated from local height blocks around the keypoints. Then, the keypoints of the height map are matched, and the registration results are obtained.

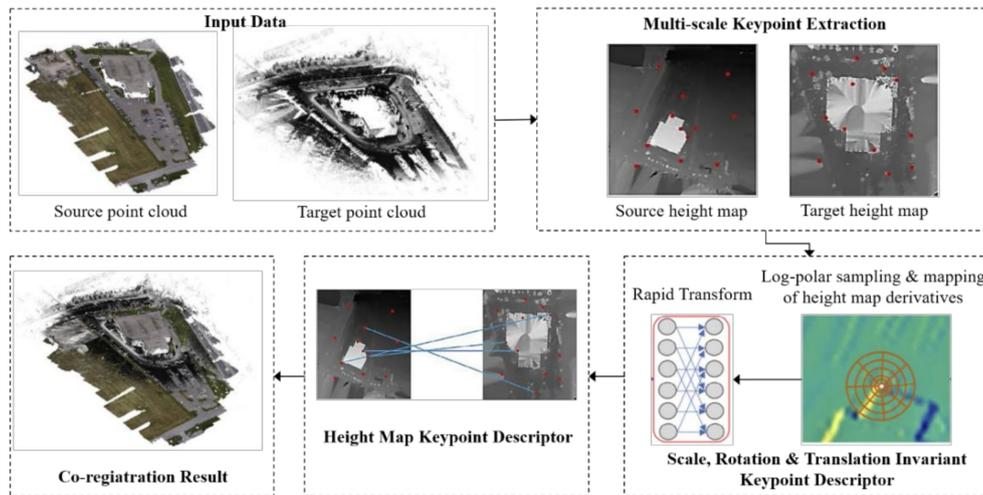


Figure 6. Overview of the height map image point matching approach for co-registering 3D multi-sensor point clouds.

Takeshi Masuda [44] used log-polar height map (LPHM) to establish the corresponding relationship for achieving the coarse registration of multiple range images. This map represents the shape of a local surface as a height mapping of the log-polar coordinate system relative to the tangent plane. Figure 7 illustrates an overview of the LPHM-based registration method. For the input range images, the corresponding log-polar depth maps are established via LPHM. Subsequently, the paired range images are registered in a robust manner by the RANSAC algorithm, and the incorrect correspondences are eliminated as outliers. Finally, the fine registration is completed on the basis of the coarse registration results.

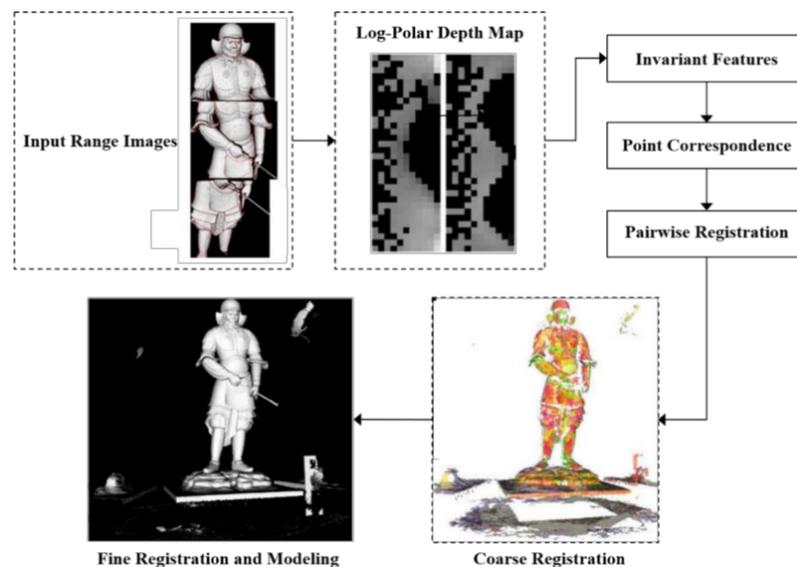


Figure 7. Overview of the registration method based on LPHM.

In the application of 3D object reconstruction based on a skeleton, a retina-like descriptive operator algorithm is constructed on the basis of the attention mechanism of human eyes and the ability to deal with complex tasks of visual scenes. The retina-like descriptive operator algorithm improves the performance of artificial bee colony algorithm in 3D object reconstruction. The description operator of retina-like mechanism also cooperates with digital elevation model (DEM). The purpose is using the variable resolution characteristics of the human retina to achieve high resolution while increasing the efficiency of 3D image

reconstruction as much as possible. Liu et al. [45] presented a continuative variable resolution DEM (cvrDEM) for the representation of 3D terrain, which has the characteristics of retina-like high and variable resolutions. The final cvrDEM product with resolution varying from 0.004 m to 0.067 m is displayed in Figure 8a. Figure 8c is the grid DEM used for performance comparison, which has a resolution of 0.03 m per pixel. Figure 8b,d are enlarged views of the area outlined by the rectangle in Figure 8a,c, respectively.

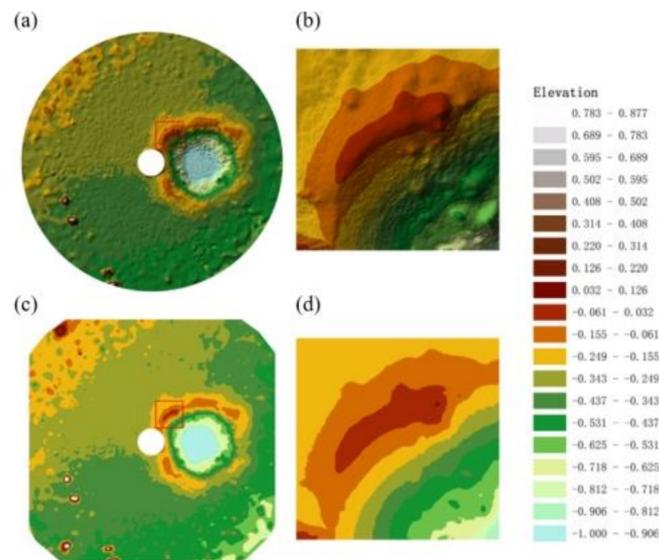


Figure 8. (a) cvrDEM, (b) enlarged view of the area outlined by a rectangle in (a), (c) grid DEM, and (d) enlarged view of the area outlined in (c).

3.2. Target Tracking

The retina-like mechanism has a notable tracking effect, which is also suitable for the field of computer vision [46,47]. Li et al. [48] introduced LPT into the target tracking field. They used the scale transformation characteristics of LPT for scale estimation without the requirement for multi-resolution pyramids. This method solves the problem of poor tracking performance for objects with obvious scale or appearance transformations under the Cartesian coordinates. Similarly, the high-resolution feature of the fovea that is similar to the human eye has also been widely used in the task of target grasping. Yamaguchi et al. [49] used a mechanism of high resolution in fovea combined with low resolution in the periphery. This method separates the target that needs to be grasped from the complex background, which can better eliminate the effect of the rotation [50].

Several works have also been conducted to enhance the robustness in target tracking. The general approach is the correlated filter tracking. The conventional correlation filter-based tracking approaches only track targets with boxes parallel to the coordinate axis. Similarity transformation for rotated targets has been rarely explained. Unlike conventional methods, Li et al. [51] converted the 4-degree of freedom (DoF) problem of target tracking into the 2-DoF problem. They applied LPT in correlated filter tracking and proposed a target tracking algorithm for large displacements. This method is beneficial to improve the robustness of target tracking.

Sharif et al. [52] introduced a novel approach that combined scale-invariance feature transform (SIFT) and LPT. SIFT solves the problem of keypoint matching between frames, and LPT eliminates the error effect caused by rotation. This kind of stability requirement is also important in the recognition of traffic signs. Ellahyani et al. [53] used LPT to recognize traffic signs. They first used the mean shift clustering method to preprocess the road sign images and then used the random forest for classifying the target. Finally, they used a combination of LPT and cross-correlation for identification. The addition of LPT effectively reduces the detection error of road signs and decreases the effect of reduced accuracy

caused by rotation. The approach proposed by Gudigar et al. [54] was similar to that of Ellahyani et al. They first selected ROI and then used LPT on the target in the ROI. Finally, they applied SVM classification on the fovea target. The overview of this approach is shown in Figure 9.

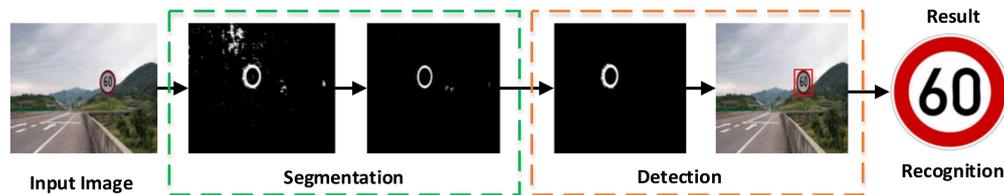


Figure 9. Overview of the approach proposed by Gudigar et al.

The retina-like mechanism also plays an important role for fast-moving objects. High-speed target detection is required in various scenarios, such as sports events and traffic detection. Conventional cameras generally record visual information during the entire exposure time through the shutter. By contrast, Zhao et al. [55] proposed a spike camera method based on the retina-like mechanism. This method continuously monitors the light intensity value of fast-moving targets and records the continuous spikes emitted by each pixel. By performing independent operations on each spike sequence, they improved the effect of frame loss caused by high-speed moving targets within the exposure time. On this basis, Zhu et al. [56] proposed a different spike camera that combines dynamic and static information. They used the retina-like principle to record continuous spike data and reconstruct images. This structure not only accurately reconstructs static targets but also has a good capture effect on high-speed moving targets. They also constructed a new spike dataset to provide a basis for subsequent research.

3.3. Deep Learning

The retina-like mechanism has also been fully applied to the field of deep learning. When humans need to pay attention to a certain area, the light is concentrated to that area by turning their eyes to obtain a high-resolution image. Several works based on this mechanism have been used to solve various problems. Itti et al. [57] proposed the human visual attention mechanism model, and thus, attention was introduced into the computer vision field. In 2017, Cheung et al. illustrated the fovea sampling grid of the biological primate retina through a deep learning computational model. This method is contrary to the previous methods and ideas of designing machine learning models based on biological results. The model is trained on the classification task and uses the least number of fixations in the visual scene under background disturbance. The tiling attributes appearing in the retina sampling grid of the trained model are obtained. They found that this lattice is similar to the primate retina. The eccentricity depends on the sampling lattice. The high-resolution area of the fovea is surrounded by the low-resolution periphery. Under certain conditions, these emerging characteristics are amplified or eliminated.

Many state-of-the-art methods of attention-based classification are available for existing deep learning networks. However, these methods have a general limitation, that is, a large amount of data training is required to improve the accuracy of the model. Dai et al. [58] proposed a guided attention recurrent network (GARN) based on retina-like scanning, as shown in Figure 10, to solve this problem. In this way, multiple ROIs are trained by scanning one image. The instructive ROI selection jointly determines the label category. This kind of guided multi-attention is trained with a small dataset to achieve high accuracy.

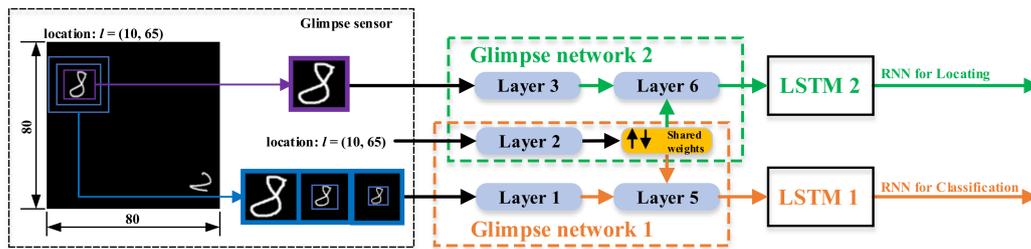


Figure 10. GARN overview. The proposed GARN model consists of two RNNs: one for locating ROIs, and the other for classification.

Recurrent Neural Networks (RNNs): one for locating ROIs, and the other for classification. The different scales of the glimpse sensors are similar to the variable resolution characteristics of the human eye. Glimpse networks output features with variable resolution. And the guided attention is added to feature maps by Long Short-Term Memory (LSTM) networks.

In 2020, Xia et al. [59] proposed a novel peripheral fovea multi-resolution driving model based on the retina-like structure. This model predicts the speed of the car through the video of a dashcam. The overview of the peripheral fovea multi-resolution driving model is shown in Figure 11. The peripheral vision module processes complete video frames with low resolution, and the foveal vision module selects subregions and uses high-resolution input from these regions to improve its driving performances. They trained the fovea selection module under the supervision of the driver’s line of sight. Adding high-resolution input from the predicted gaze position of a human driver will greatly improve the driving accuracy of the model. The multi-resolution fovea model is better than a single-resolution edge model with the same amount of floating-point operations. The driving model based on fovea multi-resolution achieves higher performance than conventional methods.

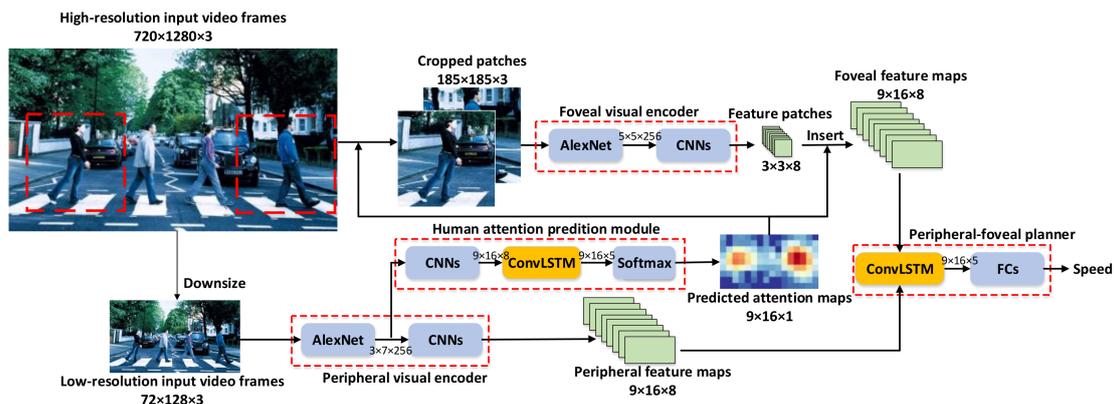


Figure 11. Overview of the peripheral fovea multi-resolution driving model.

The retina-like attention mechanism has also shown its unique advantages for practical uses in medical image processing. Hayashi et al. [60] used the retina-like sequence scanning method to segment medical images. This method changes the position of the center point of the retina to perform higher-resolution processing on the subregions of the medical image that are difficult to classify. The accuracy of the overall segmentation is improved by increasing the segmentation accuracy of each subregion.

Different from target tracking, the task of object detection and recognition is not only to find the target but also to classify it. The emergence of the retina-like mechanism makes the effect of object detection and recognition more remarkable.

Emmanuel et al. [61] defined retina-like search as a strategy for the pursuit of target detection accuracy. They divided the target detection task into two parts, as shown in Figure 12, in which one is searching for the location of the target, and the other is determining whether the target is in the searched location. They first transformed the image into log-polar coordinates to obtain the rough position of the target. Then, they trained the high-resolution image of the fovea to obtain accurate information about the target type.

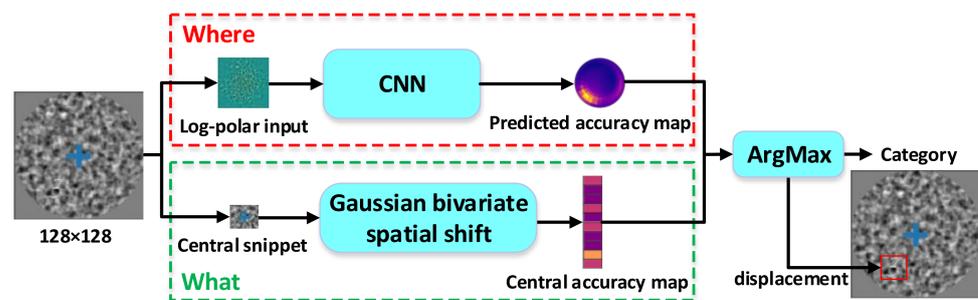


Figure 12. Computational graph of the method proposed by Emmanuel et al.

In 2019, Azevedo et al. [62] proposed a focusing algorithm based on human-like eyes named augmented-range vehicle detection system (ARVDS), which uses a deep convolutional neural network (DCNN) to detect vehicles at different image scales. This method captures the image by the front camera of the self-driving car and obtains slices of the image according to the waypoint projection of the self-driving car in the image. They simulate the way humans look forward while driving. These image slices are enlarged and sent to DCNN. Then, DCNN focuses on them and detects vehicles at a long distance. Compared with detecting the entire image at multiple scales, this method requires less processing power consumption. The ARVDS algorithm improves the average accuracy of long-distance vehicle detection from 29.51% of a single complete image to 63.15%. In 2020, Kim et al. [63] introduced the spiking neural network (SNN), which simulates the human eye, into the field of target detection and recognition. They also proposed the Spiking-YOLO network. This method is different from the method proposed by Cao et al. [64] to convert the CNN model into SNN, which is the method of directly designing the structure of SNN. In this method, two optimization tricks are proposed at the same time: the channel normalization and signed neuron imbalance threshold methods. Both methods provide fast and accurate deep SNN information transmission. The aforementioned method has achieved performance equivalent to Tiny-YOLO, but the power consumption is extremely low, and the accuracy of 51.83% is achieved on the pattern analysis, statistical modeling, and computational learning visual object class dataset.

Although the existing deep learning methods play a great role in image classification tasks, deep learning methods also have limitations for some rotated objects. However, such problems are solved in LPT. Esteves et al. [65] introduced the LPT method into the structure of CNNs and proposed the polar transformer network. As shown in Figure 13, this structure is different from the conventional CNN. The polar origin predictor and polar transformer modules are added before the classification network. The two modules eliminate the effect of rotation and scale transformation on accuracy.

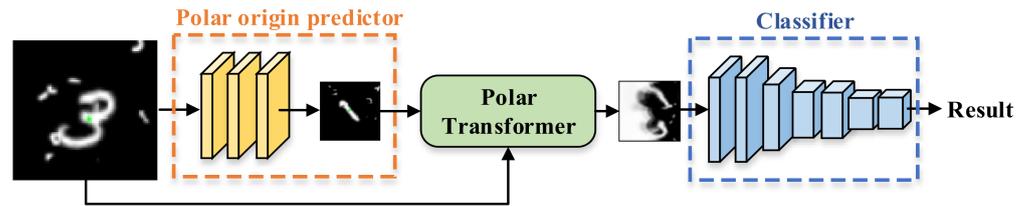


Figure 13. Architecture of polar transformer network.

3.4. Ghost Imaging

Ghost imaging (GI), which is a novel imaging technology, has attracted much attention due to its advantages, such as low cost, wide spectrum range, and robustness to light scattering [66,67]. GI provides scene information by correlating the modulate light patterns generated by the pseudo-thermal source, digital micromirror device (DMD), or other types of spatial light modulators [68] with the intensity of the light from the target scene. However, GI using traditional light modulation patterns, such as random ones, cannot balance imaging efficiency and quality [4,69]. Some researchers have conducted studies on the retina-like patterns of GI as inspired by the human eye to improve the performances of GI.

A strategy that exploits the spatiotemporal redundancy was presented by Phillips et al. [31]. This strategy rapidly records the detail information of quickly moving features in the scene and accumulates details of slower evolving areas over several consecutive frames. The retina-like patterns are created by reformatting each row of the Hadamard matrix into a 2D grid of spatially variant pixel size, as shown in Figure 14D,E. Figure 14C,F show the imaging results of traditional patterns and retina-like patterns, and using a retina-like structure, enhances the detail in the foveal region at the expense of a lower resolution in the periphery. They also realized the dynamic imaging of single fovea area and dual fovea through experiments. This proposed method is a novel approach for compressive sensing and improves the performances of GI.

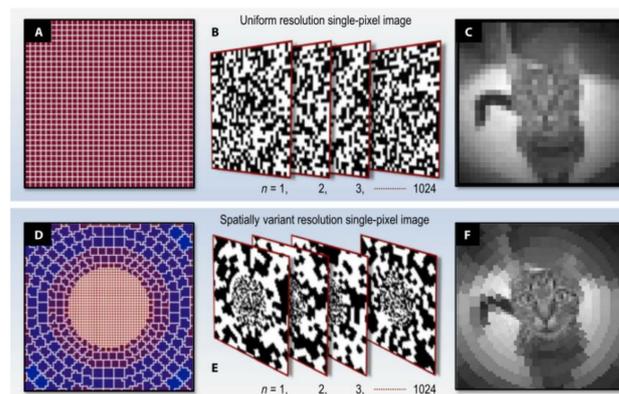


Figure 14. GI with retina-like patterns. (A) Uniform grid with 1024 pixels. (B) Examples of a complete 1024 pattern. (C) The reconstructed image with uniform grid. (D) Spatially variant pixel grid with 1024 pixels. (E) Examples of a complete 1024 pattern with spatially variant grid. (F) The reconstructed image with spatially variant grid.

A model of 3D GI combined with retina-like structure (R-3DCGI) is described for improving imaging efficiency by our group, as shown in Figure 15 [70]. A signal generator (SG) triggers a pulsed laser (PL). Random speckles are produced by the DMD that are illuminated by the PL, which is used to illuminate a target. The one-dimensional time-resolved total intensity distribution of a reflected or scattered light from the target is collected by a receiving lens (RL) and a time-resolved bucket detector (TBD). The 3D

imaging is obtained by combining images from different depth slices. However, the retina-like structure of R-3DCGI is based on logarithmic polar coordinates, as shown in Figure 16. Some retina-like properties, such as invariant scaling and rotation, have been realized in R-3DCGI. The other advantage is a higher imaging efficiency than traditional 3D GI for using a retina-like structure.

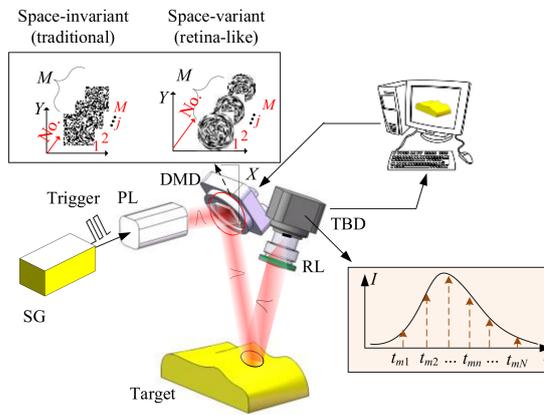


Figure 15. Principle of R-3DCGI.

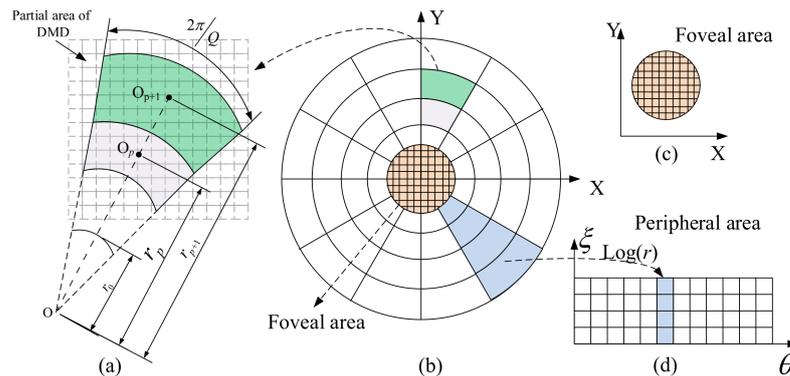


Figure 16. Retina-like structure on the DMD. (a) Details of the peripheral area of the retina-like structure. (b) Retina-like structure including the foveal and peripheral areas. (c) Foveal area. (d) Peripheral sampling.

A foveal GI based on deep learning to realize the intelligent selection of the ROI for foveal imaging was proposed by Zhai et al. [71]. Selecting the ROI intelligently by applying generative adversarial networks based on SSD architecture improves the imaging quality with higher PSNR of ROI than that of uniform-resolution GI can be achieved.

Gao et al. [72] proposed a novel compressive GI called R-CSGI, as inspired by human consciousness controlling the eyes to acquire the ROI. This method uses prior imaging information from fast Fourier single-pixel imaging to achieve a better visual effect and higher imaging quality. The principle of R-CSGI is shown in Figure 17. A sequence of fast Fourier basis patterns is illuminated on the object, and the image is reconstructed by the light intensity collected by a single-pixel detector. The ROI is then generated by the reconstructed image. The real-value random patterns used to perform compressive sensing reconstruction are generated according to the ROI. The advantage of this method is that the imaging quality obtained by compressive sensing technique can be effectively enhanced because the ROI is found previously.

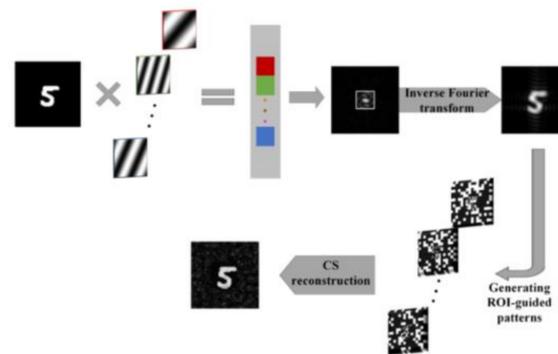


Figure 17. Principle of the R-CSGI.

A method based on parallel architecture with retina-like structure was proposed in our previous work [73], as shown in Figure 18. Retina-like patterns are divided into various blocks, and this way is different from the previous work that applied whole retina-like patterns to sampling. This method calculates the data of each of block rather than the whole image, and this way improves the efficiency of the reconstruction algorithm and the retina-like pattern enhances the imaging quality of the ROI at the same time.

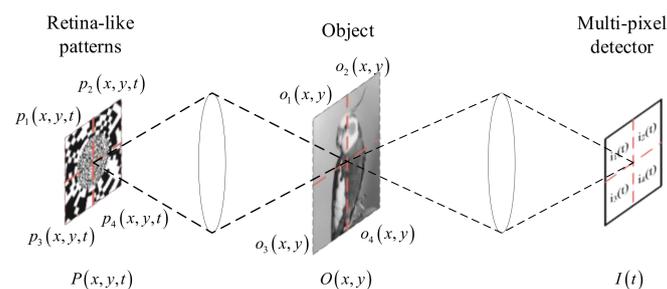


Figure 18. Measurement principle of the parallel retina-like computational GI.

The above-mentioned studies were mostly based on the structural characteristics of the retina. The fact that human eyes have a poorer spatial resolution to blues than reds and greens was utilized by Qiu et al. [74]. They proposed to use an ultra-low sampling ratio to sample the blue component of color images. The results showed that 95% of the measurements can be reduced in the acquisition of the blue component of nature images with the image size of 256×256 pixels. This method is an alternative approach to realize real-time and full-color imaging.

4. Discussion

We learn the methods of retina-like imaging, including hardware and software, from Section 2. Meanwhile, the view of practical use in Section 3 indicates that various applications use the properties of retina-like imaging. Although the theory of retina-like imaging is relatively good, some challenges exist in methods and applications. Here, we select several typical challenges for discussion. First is the challenge of the optimization of multi parameters of retina-like imaging. According to Equation (2), the parameters of retina-like imaging include rings (M), sectors (N), the radius of the blind hole (r_0), and the maximum radius of the outmost ring (r_{max}). These parameters are correlated and designed by practical use. For example, from the view of electro-based method, fill factor (FF) should be given more consideration due to achieving high optical efficiency. Meanwhile, the pixel crosstalk noise should be considered due to the space-variant size of photosurfaces, leading to low SNR. A space-variant lens array has been proposed to increase the distance between the neighbor pixels [23] for solving the above-mentioned issues. As a result, the size of

each pixel is the same and is beneficial of fabricating. However, the optical aperture is supposed to be circle, and the maximum of FF based on space-variant lens is $\pi/4$ due to the physical limitations of lens. Therefore, other structures of lens array have been proposed to increase FF [75]. Therefore, optimizing or balancing the retina-like parameters is important to improve the performances for different practical uses.

Second, improving the performance of retina-like GI is also a challenge. From the view of the GI mechanism, the method based on retina-like patterns provides a potential method for balancing the trade-off between the resolution and imaging efficiency of GI. Compared with space-invariant patterns, combining the retina-like GI is efficient for improving the imaging quality and shortening the imaging time. For example, super-resolution reconstruction methods are efficient to improve the spatial resolution limited by the hardware device (e.g., the micro-mirror size of DMD), such as sub-pixel shifting. Meanwhile, the time cost is extremely reduced when the parallel GI structure is combined. The invariant rotation and scaling, as the main features of the retina-like structure, are used on the GI system to track moving objects, especially with the object that needs to vary the size and the movement forms on the same FOV. With the retina-like patterns, the change in target need not be considered during the imaging reconstruction. The complexity of data processing is also reduced, which improves the efficiency of the GI with moving objects.

The third challenge is deep learning combined with retina-like feature. Although the retina-like mechanism has been widely used in deep learning, it still has some limitations. For example, deep learning networks have a strong dependence on datasets, which restricts its application in many practical scenarios, such as defect detection, military target identification, and medical image processing. The mechanism of retina-like imaging mitigates the effect of redundant data. However, the way to select the information that needs to be focused on is challenging for deep learning. The application of retina-like attention mechanism to SNN is also challenging to study. SNN originated in neuroscience, and it also has an important directive value for the design of retina-like deep learning networks. Given the lightweight network parameters of SNN, it is more suitable for application in actual scenarios. The combination of retina-like architecture and SNN will be more promising.

5. Conclusions

The properties of the retina show good advantages in applications that simultaneously require large FOV, high resolution, and real-time performance. We review the methods and applications that use retina-like mechanisms. On the one hand, retina-like imaging is obtained by hardware or software. Software-based methods are easier to realize than hardware ones, but they have lower efficiencies. On the other hand, the applications including 3D acquisition and reconstruction, target tracking, deep learning, and GI have been introduced and their practical uses have been illustrated in detail. Compared with the traditional 3D acquisition and reconstruction technologies, the technologies that combine the characteristics of retina-like imaging have achieved better results, and corresponding methods are constantly being proposed. In the task of target tracking, the value of retina-like imaging is mainly reflected in the reduction of image noise. Moreover, for the target of rotation and scaling, the tracking of retina-like imaging is more robust. In the application of deep learning, the retina-like mechanism is mainly reflected in the following three aspects. First of all, the retina-like gaze can be treated as an attention mechanism in deep learning to enhance features. Secondly, in solving the problem of multi-scale object detection and recognition, the retina-like rotation and scale invariance also work well. Finally, the accuracy of image recognition of retina-like imaging has been greatly improved. Research on RGI has focused more on the adjustment of the location and size of the region of interest and its different applications in imaging. Most of the current retina-like structures are applied to the design of illumination patterns for GI, and investigating how to realize non-uniform sampling by designing devices may better enhance the performance of RGI. The discussions of challenges and the potential approaches make it easy to understand the

direction of the retina-like character, which will help in designing the good performances of imaging.

Author Contributions: Conceptualization, Q.H. and J.C.; methodology, Y.T.; investigation, resources and writing—original draft preparation, M.T., Y.C., D.Z., Y.N., C.B. and H.C.; writing—review and editing, M.T. and J.C.; project administration, Q.H. and Y.T.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported in part by the funding of foundation enhancement program under Grant 2019-JCJQ-JJ-273 and 2020-JCJQ-JJ-030, in part by the National Natural Science Foundation of China under Grant 61871031, Grant 61875012, and Grant 61905014.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

FOV	field of view
2D	two-dimensional
3D	three-dimensional
LPT	log-polar transform
LiDAR	light detection and ranging
MEMS	micro-electro-mechanical system
ROI	region of interest
SNR	signal-to-noise
LPHM	log-polar height map
SIFT	scale-invariance feature transform
DEM	digital elevation model
cvrDEM	continuative variable resolution DEM
DoF	degree of freedom
GARN	Guided attention recurrent network
RNNs	recurrent neural networks
LSTM	long short-term memory
ARVDS	augmented-range vehicle detection system
DCNN	deep convolutional neural network
SNN	spiking neural network
GI	ghost imaging
DMD	digital micromirror device
R-3DCGI	3D GI combining with retina-like structure
SG	signal generator
PL	pulsed laser
RL	receiving lens
TBD	time-resolved bucket detector
FF	fill factor

References

- Kim, J.J.; Liu, H.; Ashtiani, A.O.; Jiang, H. Biologically inspired artificial eyes and photonics. *Rep. Prog. Phys.* **2020**, *83*, 047101. [[CrossRef](#)] [[PubMed](#)]
- Lee, G.J.; Choi, C.; Kim, D.H.; Song, Y.M. Bioinspired artificial eyes: Optic components, digital cameras, and visual prostheses. *Adv. Funct. Mater.* **2017**, 1705202. [[CrossRef](#)]
- Wang, T.; Yu, W.; Li, C.; Zhang, H.; Xu, Z.; Lu, Z.; Sun, Q. Biomimetic compound eye with a high numerical aperture and anti-reflective nanostructures on curved surfaces. *Opt. Lett.* **2012**, *37*, 2397–2399. [[CrossRef](#)] [[PubMed](#)]
- Tanida, J.; Mima, H.; Kagawa, K.; Umeda, M. Application of a compound imaging system to odontotherapy. *Opt. Rev.* **2015**, *22*, 322–328. [[CrossRef](#)]
- Leitel, R.; Brückner, A.; Buß, W.; Viollet, S.; Pericet-Camara, R.; Mallot, H.; Bräuer, A. Curved artificial compound-eyes for autonomous navigation. In Proceedings of the Spie Photonics Europe Conference, Brussels, Belgium, 13–17 April 2014.

6. Neumann, J.; Fermuller, C.; Aloimonos, Y.; Viollet, S.; Bruer, A. Compound eye sensor for 3D ego motion estimation. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots & Systems, Sendai, Japan, 28 September–2 October 2004.
7. Li, L.; Zhang, C.; Jin, G.; Yuan, W.; Zhao, D. Study on the optical properties of Angel Lobster eye X-ray flat micro pore optical device. *Opt. Commun.* **2021**, *483*, 126656. [[CrossRef](#)]
8. Sveda, L.; Hudec, R.; Pina, L.; Semencova, V.; Inneman, A. Lobster eye: Technology and imaging properties. In Proceedings of the SPIE—The International Society for Optical Engineering Conference, Prague, Czech Republic, 20–23 April 2009; Volume 7360.
9. Inneman, A.V.; Hudec, R.; Pina, L.; Gorenstein, P. Lobster eye x-ray optics. In Proceedings of the SPIE—The International Society for Optical Engineering, Denver, CO, USA, 18–23 July 1999; pp. 72–79.
10. Cheng, Y.; Cao, J.; Tang, X.; Hao, Q. Optical zoom imaging systems using adaptive liquid lenses. *Bioinspir. Biomim.* **2021**, *16*, 041002. [[CrossRef](#)]
11. Shen, H.; Sun, W.; Fu, J. Multi-view online vision detection based on robot fused deposit modeling 3D printing technology. *Rapid Prototyp. J.* **2019**, *25*, 343–355. [[CrossRef](#)]
12. Yin, M.; Liu, W.; Zhao, X.; Guo, Q.W.; Bai, R.; F. Image denoising using trivariate prior model in nonsubsampling dual-tree complex contourlet transform domain and non-local means filter in spatial domain. *Opt. Int. J. Light Electron Opt.* **2013**, *124*, 6896–6904. [[CrossRef](#)]
13. Almagambetov, A.; Casares, M.; Velipasalar, S. Autonomous tracking of vehicle taillights from a mobile platform using an embedded smart camera. In Proceedings of the International Conference on Distributed Smart Cameras ICDSC, Palm Springs, CA, USA, 29 October–1 November 2013.
14. Cheng, Y.; Cao, J.; Zhang, Y.; Hao, Q. Review of state-of-the-art artificial compound eye imaging systems. *Bioinspir. Biomim.* **2019**, *14*, 031002. [[CrossRef](#)]
15. Schwartz, E.L. A quantitative model of the functional architecture of human striate cortex with application to visual illusion and cortical texture analysis. *Biol. Cybern.* **1980**, *37*, 63–76. [[CrossRef](#)]
16. Traver, V.J.; Bernardino, A. A review of log-polar imaging for visual perception in robotics. *Robot. Auton. Syst.* **2010**, *58*, 378–398. [[CrossRef](#)]
17. Li, S.; Cao, J.; Cheng, Y.; Meng, L.; Xia, W.; Hao, Q.; Fang, Y. Spatially adaptive retina-like sampling method for imaging LiDAR. *IEEE Photonics J.* **2019**, *11*, 1–16. [[CrossRef](#)]
18. Sungheetha, A.; Rajesh, S.R. GTIKF—Gabor-transform incorporated K-means and fuzzy C means clustering for edge detection in CT and MRI. *J. Soft Comput. Paradig.* **2020**, *2*, 111–119. [[CrossRef](#)]
19. Versaci, M.; Morabito, F.C. Image edge detection: A new approach based on fuzzy entropy and fuzzy divergence. *Int. J. Fuzzy Syst.* **2021**, *23*, 1–19. [[CrossRef](#)]
20. Benoit, A.; Caplier, A.; Durette, B.; Herault, J. Using Human Visual System modeling for bio-inspired low level image processing. *Comput. Vis. Image Underst.* **2010**, *114*, 758–773. [[CrossRef](#)]
21. Tistarelli, M.; Sandini, G. Estimation of depth from motion using an anthropomorphic visual sensor. *Image Vis. Comput.* **1990**, *8*, 271–278. [[CrossRef](#)]
22. Sandini, G.; Questa, P.; Scheffer, D.; Diericks, B.; Mannucci, A. A retina-like CMOS sensor and its applications. In Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop, Cambridge, MA, USA, 16–17 March 2000.
23. Jie, C.; Qun, H.; Yong, S.; Fan, F.; Liu, T.; Yang, Y.; Gao, H. Non-uniform lens array based on log-polar mapping. *Acta Photonica Sin.* **2014**, *4*, 91–96.
24. Carles, G.; Chen, S.; Bustin, N.; Downing, J.; McCall, D.; Wood, A.; Harvey, A.R. Multi-aperture foveated imaging. *Opt. Lett.* **2016**, *41*, 1869–1872. [[CrossRef](#)]
25. Xu, C.; Cheng, D.; Chen, J.; Wang, Y. Design of all-reflective dual-channel foveated imaging systems based on freeform optics. *Appl. Opt.* **2016**, *55*, 2353–2362. [[CrossRef](#)] [[PubMed](#)]
26. Carles, G.; Babington, J.; Wood, A.; Ralph, J.F.; Harvey, A.R. Superimposed multi-resolution imaging. *Opt. Express* **2017**, *25*. [[CrossRef](#)]
27. Thiele, S.; Arzenbacher, K.; Gissibl, T.; Giessen, H.; Herkommer, A.M. 3D-printed eagle eye: Compound microlens system for foveated imaging. *Sci. Adv.* **2017**, *3*, e1602655. [[CrossRef](#)] [[PubMed](#)]
28. Wang, X.; Chang, J.; Niu, Y.; Du, X.; Zhang, K.; Xie, G.; Zhang, B. Design and demonstration of a foveated imaging system with reflective spatial light modulator. *Front. Optoelectron.* **2017**, *10*, 89–94. [[CrossRef](#)]
29. Wang, S.; Chen, X.; Yang, Y.; Ye, M. Foveated imaging using a liquid crystal lens. *Optik* **2019**, *193*, 163041. [[CrossRef](#)]
30. Cao, J.J.; Hou, Z.S.; Tian, Z.N.; Hua, J.G.; Zhang, Y.L.; Chen, Q.D. Bioinspired zoom compound eyes enable variable-focus imaging. *ACS Appl. Mater. Interfaces* **2020**, *12*, 10107–10117. [[CrossRef](#)]
31. Phillips, D.B.; Sun, M.-J.; Taylor, J.M.; Edgar, M.P.; Barnett, S.M.; Gibson, G.M.; Padgett, M.J. Adaptive foveated single-pixel imaging with dynamic supersampling. *Sci. Adv.* **2017**, *3*, e1601782. [[CrossRef](#)]
32. Cao, J.; Li, Y.; Zhou, D.; Zhang, F.H.; Zhang, K.; Tang, M.; Wang, X.; Hao, Q. Foveal scanning based on an optical-phases array. *Appl. Opt.* **2020**, *59*, 4165–4170. [[CrossRef](#)]
33. Gamba, P.; Lombardi, L.; Porta, M. Log-map analysis. *Parallel Comput.* **2008**, *34*, 757–764. [[CrossRef](#)]
34. Wong, W.K.; Choo, C.W.; Loo, C.K.; Teh, J.P. FPGA implementation of log-polar mapping. In Proceedings of the International Conference on Mechatronics & Machine Vision in Practice, Auckland, New Zealand, 2–4 December 2008.

35. Cheung, B.; Weiss, E.; Olshausen, B. Emergence of foveal image sampling from learning to attend in visual scenes. *arXiv* **2016**, arXiv:1611.09430. Available online: <https://arxiv.org/abs/1611.09430> (accessed on 10 June 2021).
36. Ortiz-Coder, P.; Sánchez-Ríos, A. A self-assembly portable mobile mapping system for archeological reconstruction based on VSLAM-photogrammetric algorithm. *Sensors* **2019**, *19*, 3952. [[CrossRef](#)]
37. Nguyen, S.V.; Tran, H.M.; Maleszka, M. Geometric modeling: Background for processing the 3D objects. *Appl. Intell.* **2021**, *51*, 1–20.
38. Xue, F.; Lu, W.; Chen, K.; Webster, C., J. BIM reconstruction from 3D point clouds: A semantic registration approach based on multimodal optimization and architectural design knowledge. *Adv. Eng. Inform.* **2019**, *42*, 100965.1–100965.12. [[CrossRef](#)]
39. Ning, J.; McCLean, S.; Cranley, K. Using simulated annealing for 3D reconstruction of orthopedic fracture. *Med. Phys.* **2004**, *31*. [[CrossRef](#)]
40. Huh, J.; Sang, J.P.; Lee, J.K. Measurement of proptosis using computed tomography based three-dimensional reconstruction software in patients with Graves' orbitopathy. *Sci. Rep.* **2020**, *10*, 14554. [[CrossRef](#)] [[PubMed](#)]
41. Lee, I.H.; Mahmood, M.T.; Choi, T.S. Robust focus measure operator using adaptive log-polar mapping for three-dimensional shape recovery. *Microsc. Microanal.* **2015**, *21*, 442–458. [[CrossRef](#)] [[PubMed](#)]
42. Akter, M.; Lambert, A.J.; Pickering, M.R.; Scarvell, J.M.; Smith, P.N. A non-invasive method for kinematic analysis of knee joints. In Proceedings of the IEEE International Symposium on Signal Processing & Information Technology, Athens, Greece, 12–15 December 2014.
43. Persad, R.A.; Armenakis, C. Automatic co-registration of 3D multi-sensor point clouds. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 162–186. [[CrossRef](#)]
44. Masuda, T. Log-polar height maps for multiple range image registration. *Comput. Vis. Image Underst.* **2009**, *113*, 1158–1169. [[CrossRef](#)]
45. Liu, Z.; Peng, M.; Di, K. A continuative variable resolution digital elevation model for ground-based photogrammetry. *Comput. Geosci.* **2014**, *62*, 71–79. [[CrossRef](#)]
46. Deng, L.; Wang, Y.; Liu, B.; Liu, W.; Qi, Y. Biological modeling of human visual system for object recognition using GLoP filters and sparse coding on multi-manifolds. *Mach. Vis. Appl.* **2018**, *29*, 965–977. [[CrossRef](#)]
47. Wang, H.; Hao, Q.; Cao, J.; Wang, C.; Zhang, H.; Zhou, Z.; Li, S. Target recognition method on retina-like laser detection and ranging images. *Appl. Opt.* **2018**, *57*, B135. [[CrossRef](#)] [[PubMed](#)]
48. Li, D.; Wen, G.; Kuai, Y.; Zhang, X. Log-polar mapping-based scale space tracking with adaptive target response. *J. Electron. Imaging* **2017**, *26*, 033003. [[CrossRef](#)]
49. Yamaguchi, T.; Hashimoto, S.; Berton, F.; Sandini, G. Edge-based extraction of a grasped object with retina-like sensor. In Proceedings of the International Workshop on Systems, Maribor, Slovenia, 27–30 June 2007.
50. Sahare, P.; Dhok, S.B. Script pattern identification of word images using multi-directional and multi-scalable textures. *J. Ambient Intell. Humaniz. Comput.* **2021**. [[CrossRef](#)]
51. Li, Y.; Zhu, J.; Hoi, S.; Song, W.; Wang, Z.; Liu, H. Robust estimation of similarity transformation for visual object tracking. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
52. Sharif, M.; Khan, S.; Saba, T.; Raza, M.; Rehman, A. Improved video stabilization using SIFT-log polar technique for unmanned aerial vehicles. In Proceedings of the International Conference on Computer and Information Sciences (ICCIS), Aljouf, Saudi Arabia, 10–11 April 2019.
53. Ellahyani, A.; Ansari, M.E. Mean shift and log-polar transform for road sign detection. *Multimed. Tools Appl.* **2016**, *76*, 24495–24513. [[CrossRef](#)]
54. Gudigar, A.; Chokkadi, S.; Raghavendra, U.; Acharya, U., R. Multiple thresholding and subspace based approach for detection and recognition of traffic sign. *Multimed. Tools Appl.* **2017**, *76*, 6973–6991. [[CrossRef](#)]
55. Zhao, J.; Xiong, R.; Zhao, R.; Wang, J.; Ma, S.; Huang, T. Motion estimation for spike camera data sequence via spike interval analysis. In Proceedings of the IEEE International Conference on Visual Communications and Image Processing (VCIP), Macau, China, 1–4 December 2020; pp. 371–374.
56. Zhu, L.; Dong, S.; Li, J.; Huang, T.; Tian, Y. Retina-like visual image reconstruction via spiking neural model. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 1438–1446.
57. Itti, L.; Koch, C. Computational modelling of visual attention. *Nat. Rev. Neurosci.* **2001**, *2*, 194–203. [[CrossRef](#)] [[PubMed](#)]
58. Dai, X.; Kong, X.; Guo, T.; Lee, J.B.; Liu, X.; Moore, C. Recurrent networks for guided multi-attention classification. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Virtual Event, CA, USA, 6–10 July 2020; pp. 412–420.
59. Xia, Y.; Kim, J.; Canny, J.; Zipser, K.; Canas-Bajo, T.; Whitney, D. Periphery-fovea multi-resolution driving model guided by human attention. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Snowmass Village, CO, USA, 1–5 March 2020; pp. 1767–1775.
60. Hayashi, S.; Raytchev, B.; Tamaki, T.; Kaneda, K. Biomedical image segmentation by retina-like sequential attention mechanism using only a few training images. In Proceedings of the International Workshop on Machine Learning in Medical Imaging, Shenzhen, China, 13 October 2019; pp. 283–291.
61. Daucé, E.; Albiges, P.; Perrinet, L.U. A dual foveal-peripheral visual processing model implements efficient saccade selection. *J. Vis.* **2020**, *20*, 22. [[CrossRef](#)]

62. Azevedo, P.; Panceri, S.S.; Guidolini, R.; Cardoso, V.B.; Badue, C.; Oliveira-Santos, T.; De Souza, A.F. Bio-inspired foveated technique for augmented-range vehicle detection using deep neural networks. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.
63. Kim, S.; Park, S.; Na, B.; Yoon, S. Spiking-YOLO: Spiking neural network for energy-efficient object detection. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 11270–11277.
64. Cao, Y.; Chen, Y.; Khosla, D. Spiking deep convolutional neural networks for energy-efficient object recognition. *Int. J. Comput. Vis.* **2015**, *113*, 54–66. [[CrossRef](#)]
65. Esteves, C.; Allen-Blanchette, C.; Zhou, X.; Daniilidis, K. Polar transformer networks. *arXiv* **2017**, arXiv:1709.01889. Available online: <https://arxiv.org/abs/1709.01889> (accessed on 20 June 2021).
66. Bian, L.; Suo, J.; Dai, Q.; Chen, F. Experimental comparison of single-pixel imaging algorithms. *J. Opt. Soc. Am. A* **2017**, *35*, 78–87. [[CrossRef](#)] [[PubMed](#)]
67. Sun, M.J.; Zhang, J.M. Single-pixel imaging and its application in three-dimensional reconstruction: A brief review. *Sensors* **2019**, *19*, 732. [[CrossRef](#)]
68. Xu, Z.-H.; Chen, W.; Penuelas, J.; Padgett, M.; Sun, M.-J. 1000 fps computational ghost imaging using LED-based structured illumination. *Opt. Express* **2018**, *26*, 2427–2434. [[CrossRef](#)] [[PubMed](#)]
69. Sun, S.; Liu, W.T.; Lin, H.Z.; Zhang, E.F.; Liu, J.Y.; Li, Q.; Chen, P.X. Multi-scale adaptive computational ghost imaging. *Sci. Rep.* **2016**, *6*, 37013. [[CrossRef](#)] [[PubMed](#)]
70. Zhang, K.; Cao, J.; Hao, Q.; Zhang, F.; Feng, Y.; Cheng, Y. Modeling and simulations of retina-like three-dimensional computational ghost imaging. *IEEE Photonics J.* **2019**, *11*, 1–13. [[CrossRef](#)]
71. Zhai, X.; Cheng, Z.-D.; Hu, Y.-D.; Chen, Y.; Liang, Z.-Y.; Wei, Y. Foveated ghost imaging based on deep learning. *Opt. Commun.* **2019**, *448*, 69–75. [[CrossRef](#)]
72. Gaol, Z.; Cheng, X.; Zhang, L.; Hu, Y.; Hao, Q. Compressive ghost imaging in scattering media guided by region of interest. *J. Opt.* **2020**, *22*, 055704.
73. Cao, J.; Zhou, D.; Zhang, F.; Cui, H.; Zhang, Y.; Hao, Q. A novel approach of parallel retina-like computational ghost imaging. *Sensors* **2020**, *20*, 7093. [[CrossRef](#)] [[PubMed](#)]
74. Qiu, Z.; Zhang, Z.; Zhong, J. Efficient full-color single-pixel imaging based on the human vision property—“giving in to the blues”. *Opt. Lett.* **2020**, *45*, 3046–3049. [[CrossRef](#)]
75. Zhu, J.; Li, M.; Qiu, J.; Ye, H. Fabrication of high fill-factor aspheric microlens array by dose-modulated lithography and low temperature thermal reflow. *Microsyst. Technol.* **2019**, *25*, 1235–1241. [[CrossRef](#)]