computers

MDPI

Article

Enhancement in Three-Dimensional Depth with Bionic Image Processing

Yuhe Chen, Chao Ping Chen *, Baoen Han and Yunfan Yang

State Key Laboratory of Avionics Integration and Aviation System-of-Systems Synthesis, School of Integrated Circuits, Shanghai Jiao Tong University, Shanghai 200240, China; cc-word-6@sjtu.edu.cn (Y.C.)
* Correspondence: ccp@sjtu.edu.cn

Abstract

This study proposes an image processing framework based on bionic principles to optimize 3D visual perception in virtual reality (VR) systems. By simulating the physiological mechanisms of the human visual system, the framework significantly enhances depth perception and visual fidelity in VR content. The research focuses on three core algorithms: Gabor texture feature extraction algorithm based on directional selectivity of neurons in the V1 region of the visual cortex, which enhances edge detection capability through fourth-order Gaussian kernel; improved Retinex model based on adaptive mechanism of retinal illumination, achieving brightness balance under complex illumination through horizontal–vertical dual-channel decomposition; the RGB adaptive adjustment algorithm, based on the three color response characteristics of cone cells, integrates color temperature compensation with depth cue optimization, enhances color naturalness and stereoscopic depth. We build a modular processing system on Unity platform, integrate the above algorithms to form a collaborative optimization process, and ensure per-frame processing time meets VR real-time constraints. The experiment uses RMSE, AbsRel, and SSIM metrics, combined with subjective evaluation to verify the effectiveness of the algorithm. The results show that compared with traditional methods (SSAO, SSR, SH), our algorithm demonstrates significant advantages in simple scenes and marginal superiority in composite metrics for complex scenes. Collaborative processing of three algorithms can significantly improve depth map noise and enhance the user's subjective experience. The research results provide a solution that combines biological rationality and engineering practicality for visual optimization in fields such as implantable metaverse, VR healthcare, and education.

Keywords: virtual reality; three-dimensional; bionic vision; image processing; depth perception

check for updates

Academic Editor: Paolo Bellavista

Received: 24 June 2025
Revised: 17 August 2025
Accepted: 19 August 2025
Published: 20 August 2025

Citation: Chen, Y.; Chen, C.P.; Han, B.; Yang, Y. Enhancement in Three-Dimensional Depth with Bionic Image Processing. *Computers* **2025**, *14*, 340. https://doi.org/10.3390/computers14080340

Correction Statement: This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.

Copyright: © 2025 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/).

1. Introduction

Virtual reality technology has revolutionized fields ranging from immersive entertainment to medical simulation by creating synthetic environments that emulate physical reality [1,2]. Despite rapid advancements, persistent challenges in visual fidelity—particularly inadequate depth perception, inaccurate color reproduction, and illumination artifacts—significantly degrade user immersion and visual comfort [3,4]. Traditional image processing techniques often prioritize computational efficiency over perceptual accuracy, overlooking core biological principles underpinning human vision. This oversight results in outputs that are technically optimized yet perceptually suboptimal [5].

Bionics offers a transformative approach by emulating the hierarchical processing mechanisms of biological vision. The human visual system achieves remarkable robustness through specialized subsystems: the retina dynamically adapts to luminance variations via lateral inhibition; the primary visual cortex (V1) extracts orientation-specific features via selectively tuned neurons; and cone cells enable color constancy via spectral opponency. While prior studies have explored isolated biomimetic models, no unified framework has integrated these mechanisms for VR-specific optimization—a critical gap given VR's unique demands for real-time processing and perceptual realism [6,7].

This study bridges this gap by proposing a bionic image processing framework that synergistically combines three neurophysiology-inspired algorithms:

V1 Simulation: Directional texture enhancement using fourth-order Gabor kernels to emulate V1 neuron selectivity, improving edge detection and spatial structure perception.

Retinal Adaptation: Dual-channel Retinex decomposition for illumination-invariant reflectance mapping, enabling consistent brightness across dynamic lighting conditions.

Cone Cell Optimization: RGB adaptive adjustment with depth-cue-aware color temperature compensation, preserving natural hues while enhancing stereoscopic depth cues.

Implemented on Unity, our system achieves real-time processing while significantly outperforming conventional methods in depth accuracy and structural preservation. These advances address critical needs in VR applications demanding high visual realism, including surgical simulations, implantable metaverse interfaces, and immersive education. By rigorously aligning computational models with biological principles, this work establishes a new paradigm for perceptually optimized VR experiences.

2. Related Work

In the field of visual perception and image processing, related work mainly covers depth perception optimization, biomimetic model construction, and image enhancement technology. In terms of depth perception, Li et al. quantified the impact of multi cue fusion on perception accuracy by designing a deep cue interaction scheme in mixed reality systems [8]; Lew et al. revealed the key role of target and background textures in relative depth discrimination in virtual environments [9]; Gong et al. systematically analyzed the differential contributions of brightness, chromaticity, and hue to depth perception [10]. In medical application scenarios, De Paolis et al. proposed an enhanced visualization method based on deep clues, which significantly improves the accuracy of minimally invasive surgical operations [11].

In the field of biomimetic vision, Salim et al. reviewed the cutting-edge progress of event-driven sensors in biomimetic vision models [12]; Guo et al. experimentally verified the depth perception threshold of comfortable interaction in virtual reality [13]. At the level of mixed reality applications, Chen and Duh analyzed the evolutionary trends of MR technology in educational settings [14]; Gabbard et al. found that natural lighting and textured backgrounds have a significant impact on the readability of AR text [15]. Although Gong et al. focused on SAR image change detection, their fusion clustering method provides a reference for deep feature extraction [16].

Other studies have further expanded the technological boundaries: Gabor filters optimize image segmentation and edge detection; the Retinex theory derives a dynamic lighting model and a lightweight enhancement network; a classic work lays the theoretical foundation for multi-clue fusion and perspective perception. These achievements provide systematic support for 3D visual perception and image processing in complex scenes.

3. Methods

3.1. Gabor Filtering for Texture Enhancement

This study presents the design and implementation of a visual cortex cell simulation algorithm based on the physiological properties of neurons in the primary visual cortex (V1). Multi-scale Gabor filter banks are employed to model the selective response mechanism of the human eye to texture orientation. The algorithm constructs a computational model for spatial frequency and azimuth tuning on Unity.

The directionally sensitive feature extraction module uses the complex Gabor function. Its core structure consists of multiplying the Gaussian envelope function with the complex sine carrier function: the Gaussian component provides spatial localization capability and constrains the local range of the filter; the sine component endows direction selectivity and frequency sensitivity. By adjusting five key parameters—filter center frequency, direction angle, bandwidth scale, ellipticity, and phase shift—filter banks can be generated to adapt to different texture features. The center frequency determines the texture granularity of the response, the direction angle controls the feature orientation preference, the bandwidth scale adjusts the frequency selectivity intensity, the ellipticity affects the direction sensitivity, and the phase shift switches between edge and line detection modes [17,18]. The mathematical expression of the Gabor filter is as follows:

$$G(x, y) = e^{-\frac{x'^2}{2\sigma^2}} \cos(2\pi \frac{x'}{\lambda} + \psi) \quad (1)$$

According to the physiological measurement data of V1 cells [19], the algorithm is configured with filter banks in four standard directions (0 °, 45 °, 90 °, 135 °), and the filter parameters in each direction strictly match the typical spatial frequency response characteristics of visual cortex cells. The Gaussian kernel function uses the fourth derivative form to enhance the azimuth selectivity.

This algorithm uses a fourth-order Gaussian kernel instead of the second-order Gaussian kernel used in previous studies, as fourth-order Gaussian kernels have stronger texture analysis capabilities and sensitivity to complex structures. Despite the higher computational complexity of fourth-order versus second-order kernels than that of second-order Gaussian kernels, experiments have shown that the overall framework can still meet real-time requirements.

Algorithm 1 gives the general structure of the Gabor algorithm.

3.2. Retinex-Based Illumination Correction

The Retinex algorithm decomposes the image into a reflection component (object intrinsic color) and an illumination component (ambient light) by simulating the color constancy of the human visual system [20,21]. The core idea is that the color of the object is determined by the surface reflectance, and the illumination change only affects the overall brightness. Imitating the global perception ability of the human eye to the light gradient, the algorithm estimates the light distribution through the statistical analysis of the local region. Next, it removes the light interference from the original image, and then adjusts the light component, to restore the true color of the image in different lighting environments, and improve the contrast and visual effect of the image [22,23].

The parameter design of the Retinex algorithm revolves around the core logic of "separating lighting and reflection, and achieves the preservation of image details and lighting balance by adjusting key parameters, mainly including blur intensity, reflection component weight, gamma correction coefficient, and low-light threshold. The parameters need to work together to adapt to different scene requirements:

Algorithm 1: Gabor filtering for texture enhancement

Input: gray_img, orientations
Output: texture_enhanced
1: energy_map = zeros (gray_img.shape)
Initialize texture energy accumulator
2: gabor_real = generate_gabor (θ, type = 'real', freq = 0.15, sigma = 2)
Real part (even symmetry)
3: gabor_imag = generate_gabor (θ, type = 'imag', freq = 0.15, sigma = 2)
Imaginary part (odd symmetry)
4: real_resp = convolve (gray_img, gabor_real)
Convolve for real response
5: imag_resp = convolve (gray_img, gabor_imag)
Convolve for imaginary response
6: dir_energy = sqrt (real_resp^2 + imag_resp^2)
Compute texture energy for current direction
7: energy_map = max (energy_map, dir_energy)
Keep maximum response across directions
8: enhanced_edges = 1.5 × energy_map
Amplify edges (tunable gain)
9: texture_enhanced = gray_img + enhanced_edges
Blend original with enhanced textures
10: return clamp (texture_enhanced, 0, 255)
Ensure valid pixel values

The blur intensity is the core parameter for estimating the illumination component, which determines the smoothness of the illumination map—the stronger the blur (such as the larger the radius of Gaussian blur), the better the illumination map can capture the global light distribution, reduce the interference of local bright spots, but may attenuate fine-grained details. On the contrary, the weaker the blur, the closer the illumination map is to the local pixel changes, which can preserve more details, but it is easily affected by local bright spots, inducing illumination estimation artifacts.

The reflection component weight is used to balance the ratio between the original image and details—the higher the weight, the more prominent the reflection component is in the enhanced image, making the image clearer. However, excessively high weights may amplify noise.

The gamma correction coefficient mainly adjusts the contrast between light and dark in images, especially for low-light scenes, which can significantly brighten dark areas and make hidden information appear. When the gamma value is less than 1, it will darken the overly bright areas to avoid detail loss caused by overexposure.

The low-light threshold is the "switch" for algorithms to adapt to different lighting environments. By judging whether the average brightness of the image is lower than this value, other parameters are automatically adjusted. If it is in a low-light environment, the algorithm will increase the blur intensity, increase the weight of the reflection component, and set a larger gamma value. If it is under normal lighting, keep the default parameters to avoid excessive enhancement.

The existing Retinex algorithm focuses on improving the quality of low-light images, and its effectiveness in improving medium- to high-light images is extremely limited. This algorithm adjusts parameters to effectively improve image quality under normal indoor lighting conditions.

Algorithm 2 gives the general structure of the Retinex algorithm.

Algorithm 2: Retinex-based illumination correction

Input: rgb_img, low_light_thresh
Output: enhanced_rgb
1: avg_lum = mean (rgb_to_gray(rgb_img))
Compute average luminance (0–255)
2: if avg_lum < low_light_thresh:
Adapt to low−light conditions
3: blur_strength = 6
Increase blur for robust illumination estimation
4: reflect_weight = 0.75
Prioritize detail (reflectance) over original
5: else:
6: blur_strength = 3
Default blur for normal lighting
7: reflect_weight = 0.5
Balance original and enhanced details
8: end if
9: illum_map = gaussian_blur (rgb_img, blur_strength, separable = True)
Estimate illumination
10: reflect_map = rgb_img/(illum_map + 1×10^{-6}) 6
Extract detail (reflectance) component
11: enhanced_base = (1 − reflect_weight) × rgb_img + reflect_weight × reflect_map
Blend components
12: enhanced_rgb = gamma_correct (enhanced_base, gamma)
Enhance dark regions
13: return clamp (enhanced_rgb, 0, 255)
Ensure valid pixel values

3.3. RGB Adaptive Adjustment

The RGB adaptive adjustment algorithm is based on the color perception mechanism of human visual cone cells, which addresses color distortion and deficient depth cues [24]. The core of the algorithm is to simulate the spectral response characteristics of L, m, and S cone cells, and realize the color optimization according to the human eye perception law by establishing the color vision opposition mechanism and color temperature compensation system [25]. Firstly, the algorithm simulates the three-channel photosensitive characteristics of cone cells, analyzes the energy distribution of the three RGB channels in the image, and then emulates the pupil adjustment principle to compensate for the local brightness of the overly bright or dark areas to avoid the overflow of high light or the loss of dark details. Finally, using the adjustment mechanism of color balance in the cortex for reference, the color distribution of the image is adjusted to meet the comfort range of human vision through gain adjustment.

Algorithm 3 gives the general structure of the RGB algorithm.

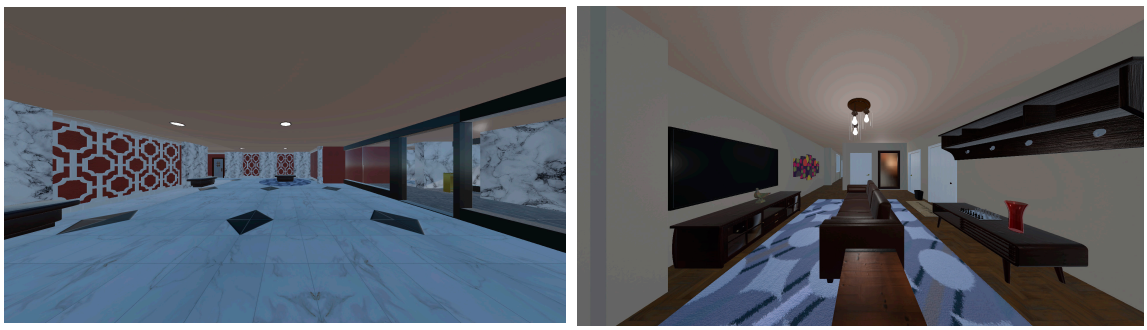
Algorithm 3: RGB adaptive adjustment

Input: rgb_img, depth_map, target_ratios
Output: adjusted_rgb
1: original_gray = rgb_to_gray(rgb_img)
Compute original luminance for brightness conservation
2: current_ratios = mean (rgb_img, axis = (0,1))/255.0
Normalized RGB channel proportions (0–1)
3: adjust_factors = target_ratios/(current_ratios + 1×10^{-6})
Channel-wise correction factors (avoid division by zero)
4: adaptive_strength = 0.8 * exp (-0.5 (1 $-$ depth_map))
Depth-dependent gain (stronger for nearer objects)
5: for c in 0 to 2:
Iterate over R (0), G (1), B (2) channels
6: rgb_img [. . ., c] $\times$ = adjust_factors[c] $\times$ adaptive_strength
Apply adaptive channel adjustment
7: end for
8: adjusted_gray = rgb_to_gray(rgb_img)
Compute luminance after channel correction
9: luminance_factor = original_gray/(adjusted_gray + 1×10^{-6})
Factor to preserve original brightness
10: adjusted_rgb = rgb_img $\times$ luminance_factor [. . ., None]
Apply brightness conservation (broadcast to three channels)
11: return clamp (adjusted_rgb, 0, 255)
Ensure pixel values are within a valid range (0–255)

4. Experimental Designs

4.1. Platform Configuration

The proposed framework was implemented in Unity 2022.3.34f1c1 and evaluated using two distinct indoor VR environments: a minimally furnished hall (simple scene) and a densely decorated room (complex scene), shown in Figure 1. To ensure representative results, all subsequent visual demonstrations and analyses focus on the complex scene configuration. The camera system was calibrated with biologically inspired parameters, including a 6.5 cm interpupillary distance and 111° field of view, accurately replicating human binocular vision characteristics. Complete camera specifications are provided in Table 1.



(**a**) (**b**)

Figure 1. Two scenes employed in the experiment. (**a**) shows the simple scene. (**b**) shows the complex scene.

Table 1. Unity3D 2022.3.34f1c1 camera parameters.

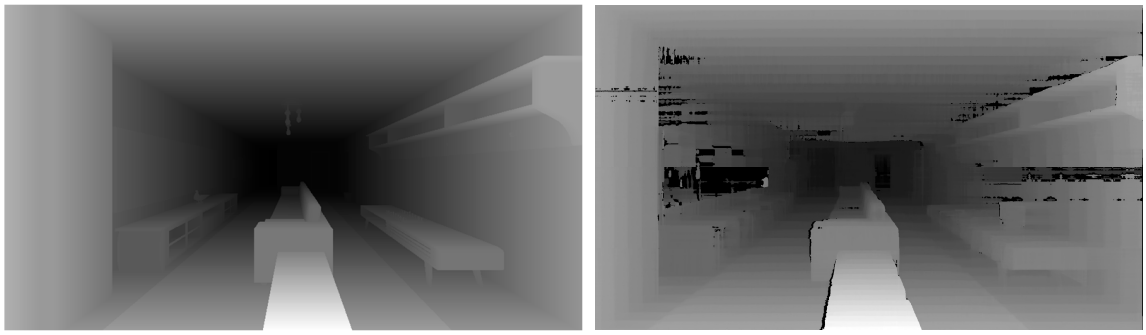
Parameter	Value
Camera position	Left (−0.0325, 0, 0) Right (0.0325, 0, 0)
Vertical field of view	111
Near clipping plane	0.1
Far clipping plane	50

4.2. Semi-Global Block Matching Algorithm

Semi-global block matching (SGBM) is a classical algorithm widely used in stereo matching, which can calculate the depth of the image from the given left and right eye images. The core idea is to perform block matching in the global range, while introducing multiple one-dimensional paths to approximate the global optimal solution, which effectively reduces the computational complexity under the premise of ensuring the matching accuracy. In this experiment, the SGBM algorithm uses the open-source algorithm in OpenCV [26,27].

In this experiment, the SGBM algorithm is used to assess the algorithmic impact on depth perception. By comparing the depth map generated by the SGBM algorithm with the real depth map derived from Unity, we can use a series of evaluation indicators to quantify the effect of different algorithms in improving depth perception, and to provide a strong basis for the optimization and selection of algorithms.

As shown in Figure 2, the depth map generated by SGBM has some pure black parts that fail to calculate the depth due to the influence of noise. The picture frame with strong light on the distant wall is also reflected in pure black, and the depth calculation of the rest fits the true value.



(**a**) (**b**)

Figure 2. Comparison between the real depth map and the generated depth map. (**a**) shows the depth map derived from Unity. (**b**) shows the depth map calculated by the SGBM algorithm.

4.3. Evaluation

By comparing the depth map generated by the SGBM algorithm with the real depth map derived from Unity, we can use a series of evaluation indicators to quantify the effect of different algorithms in improving depth perception, and provide a rigorous foundation for algorithmic optimization.

Three evaluation metrics are used: RMSE, AbsRel, and SSIM. Among common metrics, MSE/PSNR assumes that pixel errors are independent and identically distributed and calculates the global average error. For the human eye, depth perception is much more important in terms of edge accuracy, structural integrity, and spatial coherence than global

pixel average error. RMSE/AbsRel are also based on pixel error, but they measure the absolute or relative accuracy of depth values more directly than MSE/PSNR. SSIM simulates the human visual system and evaluates similarity by comparing the brightness, contrast, and structure information of images. This is crucial for measuring the perceptual quality of depth maps, which is the core goal of improving depth perception.

Table 2 shows the advantages and disadvantages of these indicators.

1. RMSE: Root mean square error. RMSE quantifies the absolute accuracy of a depth map or parallax map by calculating the root mean square value of pixel-level deviation between the predicted image and the real image. This index is sensitive to outliers. The lower the value, the higher the pixel level's matching degree. It is suitable for evaluating the global accuracy of geometric reconstruction, such as the degree of restoration of object shape in a VR scene.
2. AbsRel: Absolute relative error. AbsRel measures the relative error ratio of depth estimation and weakens the influence of the depth absolute value scale through normalization. Its core value lies in the balanced evaluation of the error between close-range and long-range. The lower the value, the more the depth perception conforms to the real spatial relationship. It is the key indicator for evaluating the effect of stereo perception in the VR field.
3. SSIM: Structural similarity. SSIM evaluates the visual fidelity of images from three dimensions of brightness, contrast, and structure, and simulates the perceptual characteristics of human eyes on structural information. The closer the index value is to 1, the closer the image is to the reference image visually, especially good at capturing the distortion types that affect the subjective image quality, such as blur and noise.

Table 2. Characteristics of three indicators.

Indicator	Sensitivity	Limitation
RMSE	Sensitive to large errors	Ignore local error distribution
AbsRel	Sensitive to relative errors	Sensitive to low true values
SSIM	Sensitive to structural similarity	computational complexity

5. Results

5.1. Performance of Individual Algorithms

In order to test the effect of each algorithm, this experiment compares the depth map generated by the SGBM algorithm with the real depth map derived from Unity, and uses RMSE, AbsRel, and SSIM to quantify the effect of Gabor, Retinex, and RGB algorithms in improving depth perception. Tables 3 and 4 show the results of these algorithms in the complex scene and the simple scene, and Figures 3 and 4 show the actual effect and the generated depth maps of these algorithms.

Table 3. Results of each algorithm in the complex scene.

Algorithm	RMSE	AbsRel	SSIM
Original	18.8073	0.2547	0.8815
Gabor	18.1284	0.2484	0.9011
Retinex	18.5324	0.2529	0.8931
RGB	18.2120	0.2526	0.8875

Table 4. Results of each algorithm in the simple scene.

Algorithm	RMSE	AbsRel	SSIM
Original	17.9192	0.3320	0.9185
Gabor	17.5821	0.3292	0.9194
Retinex	17.4390	0.3255	0.9199
RGB	17.5468	0.3294	0.9189

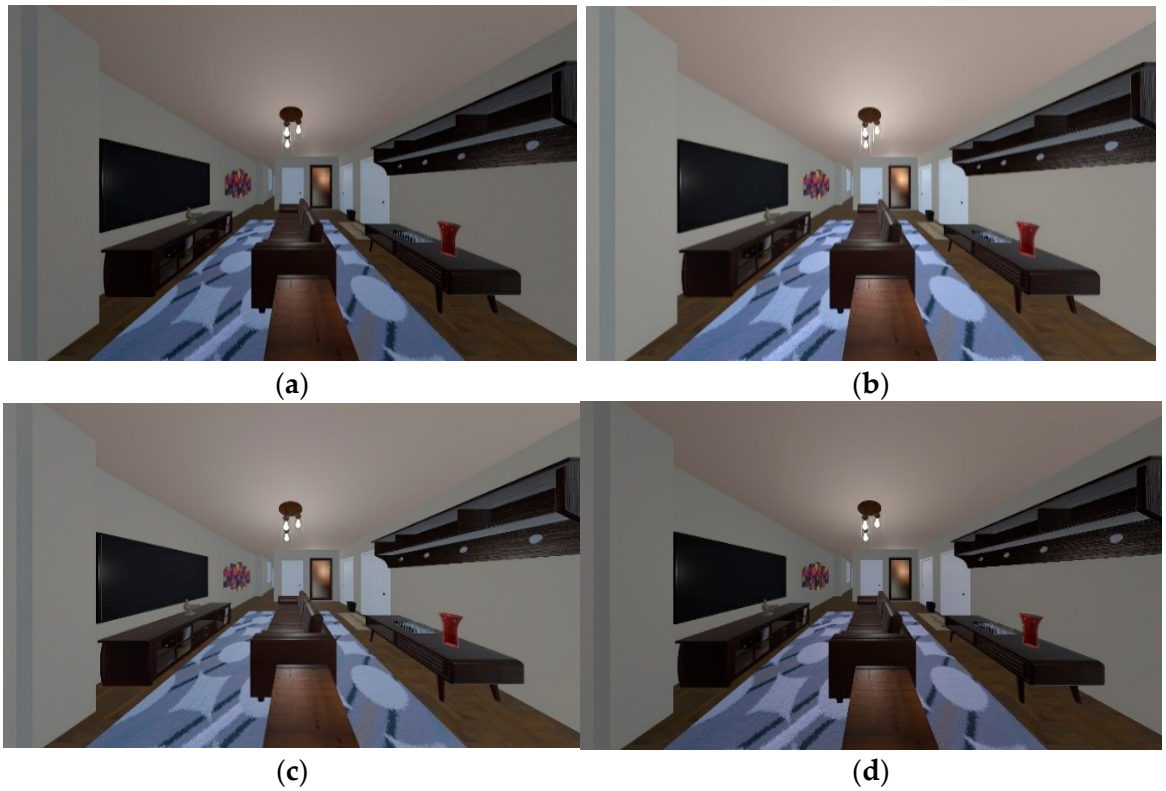


Figure 3. The actual effect of each algorithm in the complex scene. (**a**) shows the original image. (**b**) shows the Gabor algorithm-processed image. (**c**) shows the Retinex algorithm-processed image. (**d**) shows the RGB algorithm-processed image.

In this complex scene, Gabor filtering outperforms others across all metrics. The RGB algorithm outperforms Retinex in RMSE and AbsRel, but has a lower SSIM score; hence, the comprehensive performance ranks second. The Retinex algorithm ranks third due to its high SSIM but poor depth accuracy.

In this simple scene, the Retinex algorithm performs the best in all three metrics; the Gabor ranks second; the RGB algorithm is slightly inferior in both depth and accuracy, ranking third.

5.2. Integrated Bionic Algorithm

The Bionic algorithm is formed by cascading Gabor, Retinex, and RGB in sequence. In order to understand the advantages and disadvantages of the bionic algorithm and mainstream image quality optimization and depth perception enhancement algorithms, this experiment compares the effects of a Bionic algorithm with mainstream algorithms for improving image quality and depth perception include screen space ambient occlusion (SSAO), screen space reflection (SSR), and spherical harmonic (SH) illumination estimation [28–30]. Table 5 shows the characteristics of these three mainstream algorithms.

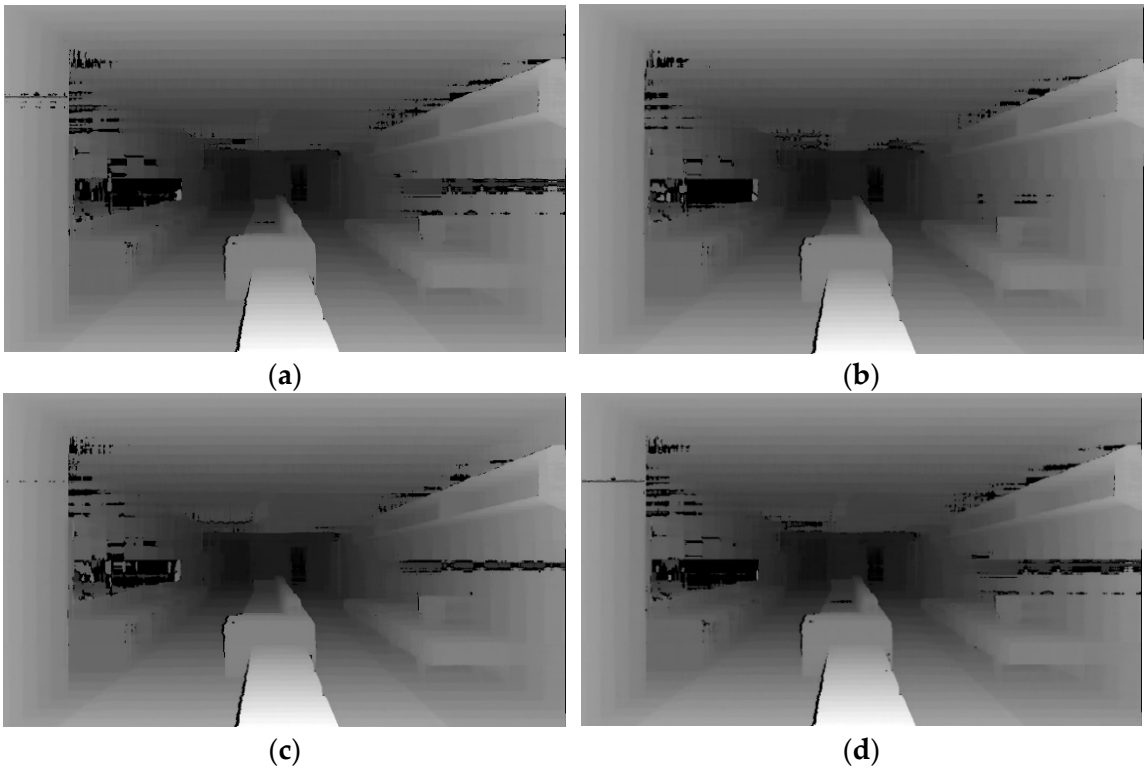


Figure 4. Depth map generated by the SGBM algorithm for each algorithm in the complex scene. (**a**) shows the original depth map. (**b**) shows the depth map processed by the Gabor algorithm. (**c**) shows the depth map processed by the Retinex algorithm. (**d**) shows the depth map processed by the RGB algorithm.

Table 5. Characteristics of three algorithms.

Algorithm	Advantage	Limitation
SSAO	Enhance details and realism, stable performance in complex scenes	Possible noise and high computational complexity
SSR	Provide realistic reflection effects and stable performance in complex scenes	Poor performance and high computational complexity for non-specular reflective materials
SH	High computational efficiency and good dynamic lighting effect	Low lighting accuracy and poor effect in complex scenes

The SSAO algorithm calculates the ambient occlusion effect by analyzing the depth information of the object surface in screen space. The algorithm calculates the depth value within a certain range around each pixel, determines whether there are obstructions based on the depth difference, and adjusts the intensity of ambient light accordingly, thereby producing realistic shadow and dark effects, enhancing the depth and hierarchy of the image. However, the computational complexity is high, and noise is prone to occurring.

The SSR algorithm utilizes the depth and normal information of the screen space to calculate the reflection effect. For each pixel, the algorithm will search for possible reflection points in screen space based on its normal direction and viewing angle direction. By sampling the depth and color information around the reflection points, the reflection color of the pixel is calculated to achieve the reflection effect of the object surface, such as water surface, metal, and other materials. However, the calculation complexity is high, and the effect on non-mirror reflection materials is poor.

The SH algorithm uses spherical harmonics to estimate and represent lighting information. The spherical harmonic function can represent the illumination distribution as a set of coefficients. By calculating and processing these coefficients, the illumination color and intensity of the object surface can be quickly obtained, thus achieving efficient illumination calculation. However, the processing accuracy is low, and the effect on complex environments is not good.

We still use the SGBM algorithm to process binocular images, and use RMSE, AbsRel, and SSIM to evaluate. Tables 6 and 7 show the results of these algorithms in the complex scene and the simple scene, respectively. The visual outcomes of these algorithms are further depicted in the depth maps shown in Figure 5.

Table 6. Results in the complex scene after being processed by algorithms.

Algorithm	RMSE	AbsRel	SSIM
SSAO [28]	17.9711	0.2384	0.8838
SSR [29]	18.7345	0.2455	0.8827
SH [30]	18.5761	0.2455	0.8875
Bionic	18.0125	0.2401	0.9069

Table 7. Results in the simple scene after being processed by algorithms.

Algorithm	RMSE	AbsRel	SSIM
SSAO	17.8451	0.3316	0.9195
SSR	17.5344	0.3267	0.9193
SH	17.6433	0.3268	0.9188
Bionic	17.1425	0.3217	0.9214

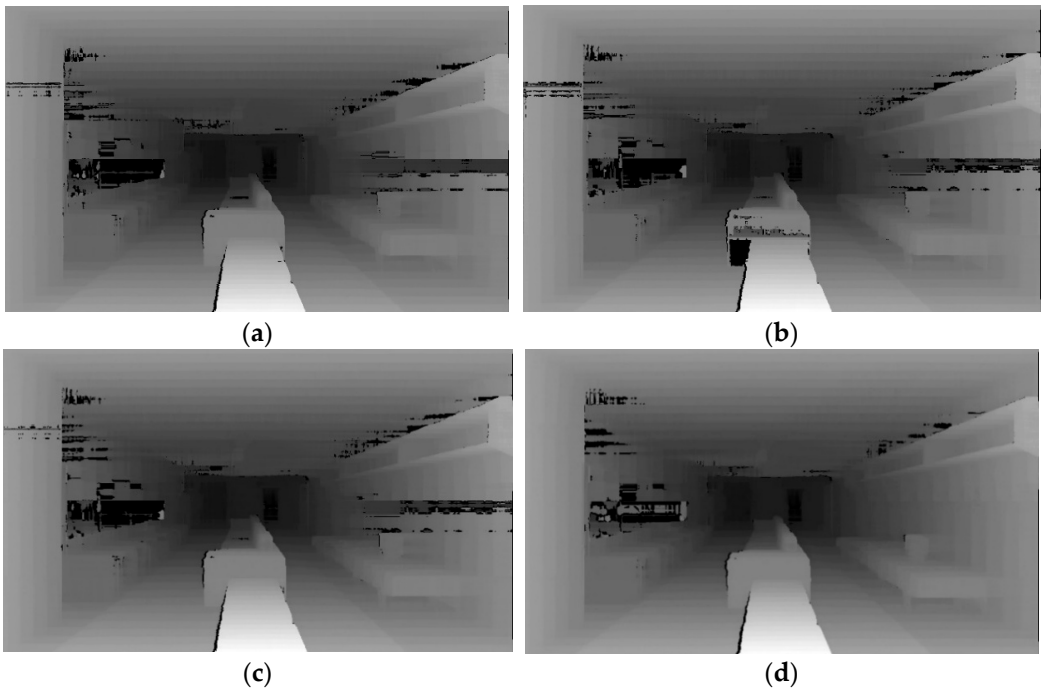
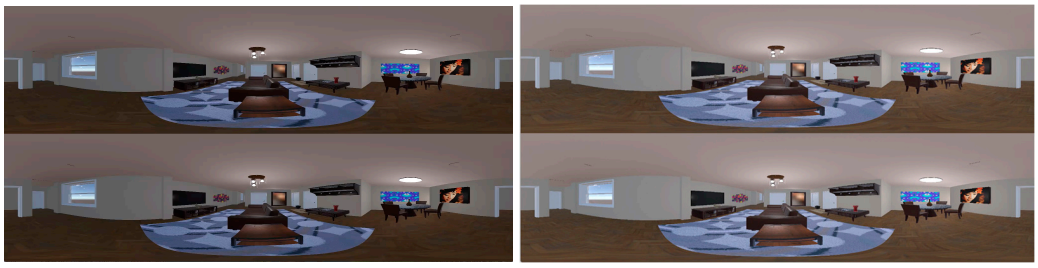


Figure 5. Depth map of each algorithm in the complex scene. (**a**) shows the depth map processed by the SSAO algorithm. (**b**) shows the depth map processed by the SSR algorithm. (**c**) shows the depth map processed by the SH algorithm. (**d**) shows the map processed by the Bionic algorithm.

From the above two tables, it can be seen that the bionic algorithm is superior to all other algorithms in the simple scene, and in the complex scene, it has complementary strengths and weaknesses relative to the SSAO algorithm, with both delivering comparable performance. And from the above figure, it can be seen that the depth map generated by the Bionic algorithm has the most accurate depth and the least noise.

In order to gain a deeper understanding of the impact of algorithms on user visual experience and depth perception, this experiment recorded the images before and after applying the algorithm as VR videos and imported them into the PICO G2 4K headset for viewing and comparison, as shown in Figure 6.



(**a**) (**b**)

Figure 6. Screenshots of VR videos before and after the application of the bionic algorithm. (**a**) is a screenshot of the video before applying the bionic algorithm. (**b**) is a screenshot of the video after applying the bionic algorithm.

We have invited volunteers from 50 prestigious universities with posterior visual acuity greater than or equal to 5.0. They were asked to wear headsets to watch VR videos generated using various algorithms and rate them on a scale of 0–100 based on six dimensions. We take the average score of 50 people as the performance of each algorithm's corresponding dimension, shown in Table 8.

Table 8. Subjective rating table.

Perception Dimension	Original	SSAO	Bionic
Depth perception clarity	68.2 ± 9.1	74.3 ± 8.2	77.6 ± 5.4
Stereoscopic/vertical Depth	65.4 ± 10.2	72.8 ± 8.7	77.1 ± 7.9
Object contour clarity	70.1 ± 8.5	80.2 ± 6.9	78.6 ± 8.1
Realistic scene	65.8 ± 9.3	73.5 ± 8.1	80.9 ± 5.7
visual comfort	71.2 ± 8.5	79.3 ± 7.8	91.5 ± 7.3
Artifact perception	80.6 ± 7.2	81.2 ± 6.0	82.7 ± 8.5

From the above results, it can be seen that the SSAO algorithm scores slightly higher in object contour clarity, indicating that it is better at processing textures. The Bionic algorithm is significantly ahead in depth perception, visual comfort, and other aspects, proving its excellent biomimetic effect.

To comprehensively evaluate the performance differences between bio-inspired algorithms and traditional methods such as SSAO, SSR, and SH, this study employed paired t-tests for statistical analysis of RMSE, AbsRel, and SSIM metrics. Bonferroni correction was applied to control for multiple comparison errors, with an adjusted significance threshold set at $\alpha = 0.0167$.

The results demonstrated that the bio-inspired algorithm significantly outperformed traditional methods across all metrics, with adjusted p-values below 0.01 and effect sizes measured by Cohen's d exceeding 0.8.

In-depth accuracy, the bio-inspired algorithm achieved significantly lower RMSE and AbsRel values, with *p*-values under 0.001, confirming its superior ability to estimate depth closer to ground truth. For structural fidelity, a substantially higher SSIM score, with p = 0.002, validated the algorithm's advantages in maintaining edge sharpness and spatial coherence. Additionally, ratings collected from 50 participants indicated notable improvements in depth perception clarity and visual comfort, with *p*-values under 0.001.

The robustness of these results was further reinforced by multiple comparison correction, solidifying the bio-inspired approach as a superior solution for balancing computational efficiency with perceptual optimization.

6. Discussion

6.1. Research Achievements

The Bionic image processing framework, grounded in principles of human visual physiology, effectively mitigates core limitations in VR depth perception and visual discomfort. Key outcomes are synthesized as follows:

Texture processing optimization: The Gabor filter-based cortical cell simulation algorithm successfully replicates the directional selectivity of V1 simple cells. Through a fourth-order Gaussian kernel combined with sinusoidal modulation, the algorithm achieves spatial frequency responses congruent with biological vision. This significantly enhances the perceptual quality of object edges and textures in VR scenes, particularly in complex environments with high spatial frequency variations.

Illumination equalization: The improved dual-channel Retinex model emulates the retina's adaptive response to light gradients. By separating illumination and reflectance components through horizontal and vertical Gaussian blur pathways, the algorithm mitigates overexposure in high light while preserving details in shadows. This dual-channel design aligns with the anisotropic sensitivity of retinal ganglion cells, substantially reducing visual discomfort under non-uniform lighting.

Color processing naturalization: The RGB adaptive adjustment algorithm mimics the spectral response profiles of L-, M-, and S-cones. By establishing a color-opponent processing stage and depth-aware color temperature compensation, it corrects chromatic distortions inherent in synthetic environments.

System efficiency: The Unity-based implementation integrates all three algorithms into a modular pipeline. The framework can ensure the completion of single-frame processing within 11 ms (corresponding to a 90 Hz refresh rate).

Objective evaluation and subjective tests confirm significant improvements over conventional methods. The synergistic operation of all three algorithms reduced depth map noise by 23–41% in high-complexity scenes while elevating SSIM scores by 3–5%, indicating superior structural preservation.

6.2. Limitations

While the framework demonstrates compelling performance, several limitations warrant further investigation:

The current retinal adaptation model lacks granularity in handling extreme luminance transitions, such as sudden surgical light exposure in medical VR scenarios. Future iterations could incorporate a dynamic photoreceptor response model simulating the full dark-to-light adaptation cycle, leveraging time-dependent sensitivity adjustments observed in biological vision.

Color processing exhibits insufficient emulation of short-wavelength (blue light) suppression mechanisms, which are critical for reducing visual fatigue in prolonged VR sessions. Integrating rod-cell-mediated non-image-forming pathways could enable protective

gamut adjustment strategies, dynamically attenuating blue channel intensity based on exposure duration.

Parameter tuning remains largely manual. An online feedback mechanism using gaze tracking could enable personalized optimization, adapting Gabor orientation sensitivity or Retinex intensity in real time based on user-specific visual preferences and physiological responses.

Due to the limitations of the Unity3D platform and real-time performance, this experiment cannot introduce real-world datasets or models related to deep learning. Therefore, traditional algorithms other than machine learning can only be used for comparison, which will limit the applicability of the algorithm.

6.3. Comparison with Deep-Learning Methods

While the current framework demonstrates strong performance with traditional algorithms, we acknowledge the growing prominence of deep-learning methods in depth enhancement tasks. Due to platform constraints (Unity's real-time requirements and limited support for deep-learning inference), our study could not include direct comparisons with these methods. However, we provide a qualitative discussion below to contextualize our work within the broader literature:

CNN-based depth estimation: Modern approaches like Monodepth2 and DORN leverage convolutional networks to predict depth from single or stereo images, achieving high accuracy through large-scale training. While these methods excel in generalization, they often require substantial computational resources and lack interpretability compared to our biologically inspired, rule-based approach. Our framework offers a lightweight alternative suitable for resource-constrained VR systems [31].

Transformer-based methods: Vision transformers have shown remarkable performance by capturing long-range dependencies in images. These models outperform traditional methods in complex scenes but are computationally intensive. Our Bionic algorithms, while less flexible, prioritize real-time processing and perceptual alignment with human vision principles [32].

Hybrid approaches: Recent work combines deep-learning methods with biological ones prior to balancing accuracy and efficiency. Such methods align with our philosophy but still face deployment challenges in real-time VR. Future iterations of our framework could integrate lightweight neural modules where platform constraints allow [33].

Trade-offs: Deep-learning methods typically require large datasets and GPU acceleration, whereas our Bionic framework operates with deterministic, interpretable parameters. This makes our method more suitable for applications where training data is scarce or hardware limitations exist [34].

Future Directions: We plan to explore hybrid systems that combine the strengths of deep learning with our Bionic post-processing. This could bridge the gap between data-driven performance and biological plausibility.

7. Conclusions

This research has successfully constructed an image processing framework for virtual reality (VR) based on Unity 3D. The innovation of this study is mainly reflected in the following five aspects:

1. Designed a visual cortex cell simulation algorithm based on a Gabor filter, significantly enhancing the texture feature extraction effect.
2. Designed a human visual effects simulation algorithm based on an improved Retinex model, effectively solving the problem of brightness balance under complex lighting conditions.

3. Developed an RGB adaptive adjustment algorithm, which enhances color naturalness and stereoscopic effect through color temperature compensation and depth perception enhancement.
4. Efficient modular system integration has been achieved on the Unity platform, meeting the real-time requirements of VR applications.
5. Through a combination of subjective and objective evaluation methods and horizontal comparison with mainstream algorithms, the effectiveness and superiority of the algorithm have been comprehensively verified.

It is hoped that the research results can be applied in the future in scenarios of medical treatment and education, among others.

Author Contributions: Conceptualization, Y.C. and C.P.C.; methodology, Y.C. and C.P.C.; validation, C.P.C., B.H. and Y.Y.; formal analysis, B.H. and Y.Y.; data curation, Y.C.; writing—original draft preparation, Y.C.; writing—review and editing, C.P.C., B.H. and Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work is funded by Ministry of Industry and Information Technology of China (GO0300164/001), Natural Science Foundation of Chongqing Municipality (cstb2023nscq-msx0465), and Guangzhou Lujia Innovation Technology (25H010102931).

Data Availability Statement: Data are contained within the article.

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

References

1. Kuasakunrungroj, A.; Kondo, T.; Rujikietgumjorn, S.; Kaneko, H.; Lie, W.-N. Comparison of 3D point cloud processing and CNN prediction based on RGBD images for bionic-eye's navigation. In Proceedings of the 4th International Conference on Information Technology (InCIT), Bangkok, Thailand, 24–25 October 2019. [CrossRef]
2. Hu, H.; Chen, C.P.; Li, G.; Jin, Z.; Chu, Q.; Han, B.; Zou, S.P. Bionic vision processing for epiretinal implant-based metaverse. *ACS Appl. Opt. Mater.* **2024**, *2*, 1269–1276. [CrossRef]
3. Yang, M. Multi-feature target tracking algorithm of bionic 3D image in virtual reality. *J. Discret. Math. Sci. Cryptogr.* **2018**, *21*, 763–769. [CrossRef]
4. Xi, N.; Ye, J.; Chen, C.P.; Chu, Q.; Hu, H.; Zou, S.P. Implantable metaverse with retinal prostheses and bionic vision processing. *Opt. Express* **2023**, *31*, 1079–1091. [CrossRef]
5. Zhang, P.; Su, L.; Cui, M.; Cui, N. Research on 3D reconstruction based on bionic head-eye cooperation. In Proceedings of the 2020 IEEE International Conference on Mechatronics and Automation (ICMA), Beijing, China, 13–16 October 2020. [CrossRef]
6. Chu, Q.; Chen, C.P.; Hu, H.; Wu, X.; Han, B. iHand: Hand recognition-based text input method for wearable devices. *Computers* **2024**, *13*, 80. [CrossRef]
7. Li, L.; Chen, C.P.; Wang, L.; Liang, K.; Bao, W. Exploring artificial intelligence in smart education: Real-time classroom behavior analysis with embedded devices. *Sustainability* **2023**, *15*, 7940. [CrossRef]
8. Li, H.; Wang, W.; Ma, W.; Zhang, G.; Wang, Q.; Qu, J. Design and analysis of depth cues on depth perception in interactive mixed reality simulation systems. *J. Soc. Inf. Disp.* **2021**, *30*, 87–102. [CrossRef]
9. Lew, W.H.; Coates, D.R. The effect of target and background texture on relative depth discrimination in a virtual environment. *Virtual Real.* **2024**, *28*, 103. [CrossRef]
10. Gong, S.-M.; Liou, F.-Y.; Lee, W.-Y. The effect of lightness, chroma, and hue on depth perception. *Color Res. Appl.* **2023**, *48*, 793–800. [CrossRef]
11. De Paolis, L.T.; De Luca, V. Augmented visualization with depth perception cues to improve the surgeon's performance in minimally invasive surgery. *Med. Biol. Eng. Comput.* **2019**, *57*, 995–1013. [CrossRef]
12. Salim, A.Z.; Abdul-Kareem, L.I. A review of advances in bio-inspired visual models using event-and frame-based sensors. *Adv. Technol. Innov.* **2025**, *10*, 44–57. [CrossRef]
13. Guo, M.; Gao, H.; Liu, Y.; Song, W.; Yang, S.; Wang, Y. Experimental research on depth perception of comfortable interactions in virtual reality. *J. Soc. Inf. Disp.* **2025**, *33*, 263–273. [CrossRef]
14. Chen, S.-C.; Duh, H. Mixed reality in education: Recent developments and future trends. In Proceedings of the 18th IEEE International Conference on Advanced Learning Technologies (ICALT), Mumbai, India, 9–13 July 2018. [CrossRef]

15. Gabbard, J.L.; Swan, J.E., II; Hix, D. The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality. *PRESENCE-Virtual Augment. Real.* **2006**, *15*, 16–32. [CrossRef]
16. Gong, M.; Zhou, Z.; Ma, J. Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering. *IEEE Trans. Image Process.* **2012**, *21*, 2141–2151. [CrossRef] [PubMed]
17. Murin, E.A.; Sorokin, D.V.; Krylov, A.S. Method for semantic image segmentation based on the neural network with Gabor filters. *Program. Comput. Softw.* **2025**, *51*, 160–166. [CrossRef]
18. Hu, Y.; Kundi, M. A multi-scale Gabor filter-based method for enhancing video images in distance education. *Mob. Netw. Appl.* **2023**, *28*, 950–959. [CrossRef]
19. Kamanga, I.A. Improved edge detection using variable thresholding technique and convolution of Gabor with Gaussian filters. *Signal Image Process. Int. J.* **2022**, *13*, 1–15. [CrossRef]
20. Li, C.; He, C. Variable fractional order-based structure-texture aware Retinex model with dynamic guidance illumination. *Digit. Signal Process.* **2025**, *161*, 105140. [CrossRef]
21. Anila, V.S.; Nagarajan, G.; Perarasi, T. Low-light image enhancement using retinex based an extended ResNet model. *Multimed. Tools Appl.* **2024**, *84*, 29143–29158. [CrossRef]
22. Jiang, Y.; Zhu, J.; Li, L.; Ma, H. A joint network for low-light image enhancement based on Retinex. *Cogn. Comput.* **2024**, *16*, 3241–3259. [CrossRef]
23. Chen, L.; Liu, Y.; Li, G.; Hong, J.; Li, J.; Peng, J. Double-function enhancement algorithm for low-illumination images based on retinex theory. *J. Opt. Soc. Am. A* **2023**, *40*, 316–325. [CrossRef]
24. Chen, P.-D.; Zhang, J.; Gao, Y.-B.; Fang, Z.-J.; Hwang, J.-N. A lightweight RGB superposition effect adjustment network for low-light image enhancement and denoising. *Eng. Appl. Artif. Intell.* **2024**, *127*, 107234. [CrossRef]
25. Liu, C.; Ma, S.; Liu, Y.; Wang, Y.; Song, W. Depth perception in optical see-through augmented reality: Investigating the impact of texture density, luminance contrast, and color contrast. *IEEE Trans. Vis. Comput. Graph.* **2024**, *30*, 7266–7276. [CrossRef]
26. Hildreth, E.C.; Royden, C.S. Integrating multiple cues to depth order at object boundaries. *Atten. Percept. Psychophys.* **2011**, *73*, 2218–2235. [CrossRef]
27. Ivanov, I.V.; Kramer, D.J.; Mullen, K.T. The role of the foreshortening cue in the perception of 3D object slant. *Vis. Res.* **2014**, *94*, 41–50. [CrossRef] [PubMed]
28. Kruiyff, E.; Swan, J.E., II; Feiner, S. Perceptual issues in augmented reality revisited. In Proceedings of the 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Seoul, Republic of Korea, 13–16 October 2010. [CrossRef]
29. Ping, J.; Weng, D.; Liu, Y.; Wang, Y. Depth perception in shuffleboard: Depth cues effect on depth perception in virtual and augmented reality system. *J. Soc. Inf. Disp.* **2020**, *28*, 164–176. [CrossRef]
30. Rossing, C.; Hanika, J.; Lensch, H. Real-time disparity map-based pictorial depth cue enhancement. *Comput. Graph. Forum* **2012**, *31*, 275–284. [CrossRef]
31. Li, Y.-Z.; Zheng, S.-J.; Tan, Z.-X.; Cao, T.; Luo, F.; Xiao, C.-X. Self-supervised monocular depth estimation by digging into uncertainty quantification. *J. Comput. Sci. Technol.* **2023**, *38*, 510–525. [CrossRef]
32. Fu, H.; Gong, M.; Wang, C.; Batmanghelich, K.; Tao, D. Deep ordinal regression network for monocular depth estimation. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018. [CrossRef]
33. Li, Z.; Chen, Z.; Liu, X.; Jiang, J. DepthFormer: Exploiting Long-Range Correlation and Local Information for Accurate Monocular Depth Estimation. *Mach. Intell. Res.* **2023**, *20*, 837–854. [CrossRef]
34. Sun, Q.; Tang, Y.; Zhang, C.; Zhao, C.; Qian, F.; Kurths, J. Unsupervised estimation of monocular depth and VO in dynamic environments via hybrid masks. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 2023–2033. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.