Designing a Rendering Density-Adaptive Light Field Display System for Enhanced Depth Perception in Glasses-Free 3D Environments

Chunlei Cao, Yuting Zhang, Zhaoli Wang, Minghao Liu, Weijing Zhong*

Harbin Institute of Information Technology, Harbin 150431, China

*Corresponding author: 1280990489@qq.com

Abstract. Light field display technology can achieve glasses-free 3D visualization by simulating the direction of light rays. However, the balance between angular resolution and spatial resolution restricts the fidelity of depth perception. Previous light field display technologies uniformly distribute the rendering density over the entire light field volume, ignoring the fact that the depth complexity of scenes varies over spatial regions. This paper presents a rendering density adaptive light field display system with the addition of depth gradient analysis and the development of a perceptual depth salience model incorporating a binocular sensitivity function and a texture masking factor based on the characteristics of the human visual system. The system employs a three-tier quantization strategy with adaptive Otsu thresholding and a boundary blending mechanism to ensure seamless transitions between density tiers. Experimental results confirm the proposed system can achieve improved depth perception MOS by 15.1 percent and reduce rendering computation by 38.2 percent with similar accuracy as the uniform density rendering.

Keywords: Light field display, Adaptive rendering density, Depth perception, Glasses-free 3D, Perceptual optimization

1 Introduction

The glasses-free technology of three-dimensional display offers the facility of stereoscopic visualization in the domains of immersive visualization and medical imaging, with the light field display technology being the best choice by simulating the distribution of light rays in multiple directions [1]. The performance of such display technology depends on the accuracy of depth perception, with the maximization of depth perception quality with limited hardware bandwidth being the major challenge.

A trade-off between angular resolution and spatial resolution is seen in light field display technology due to the space bandwidth product constraint [2]. The existing schemes make uniform angular sampling density applicable to all regions irrespective of the severity of the depth variation, with no consideration of the rendering density in the region of high depth variation. The nonlinear distribution of depth perception by the

R. Zur et al. (eds.), *Proceedings of the 2026 5th International Conference on Art Design and Digital Technology (ADDT 2026)*, Advances in Computer Science Research 131,
https://doi.org/10.2991/978-94-6239-737-8_6

human eye is seen with inacreased sensitivity at near and intermediate distances and decreased sensitivity at greater distances [3], which suggests that the actual requirement of rendering density is different in different regions.

No framework exists that takes the complexity of scene depth and the sensitivity of depth perception together into consideration in the context of rendering density. This paper proposes a density adaptive light field display system, where the depth gradient analysis module measures the complexity of depth in a scene, and the depth salience model performs a perception weighting calculation and outputs an adaptive angular density map for non-uniform light field rendering. Scene depth analysis and human perception are integrated in the process of rendering density allocation, and a region-layered pipeline and boundary blending strategy are used in the transition between different tiers.

Previous studies on light field rendering have primarily employed uniform angular sampling without adapting to scene content. Recent advances include radiance field representations for efficient rendering [4], a systematic review of naked-eye light field display trade-offs between viewpoint density and display quality [5], and real-time rendering of dynamic light fields through spatial-temporal encodings [6]. On the perceptual side, depth perception in glasses-free 3D displays is influenced by binocular disparity, motion parallax, and accommodation-vergence conflicts [7], with the interaction of multiple depth cues in three-dimensional space receiving increasing attention. Experimental research has further examined depth perception under comfortable viewing conditions in virtual reality [8]. Although perceptual models have been applied during coding processes, their utilization as input for determining rendering density allocation has not been investigated. The present work bridges this gap by jointly considering scene depth complexity and human perceptual sensitivity.

2 Proposed Method

2.1 System Architecture Overview

The proposed system comprises six sequential stages. The input scene provides an RGB-D representation consisting of a color image and a depth map $D(x, y)$ with resolution $H \times W$. The depth map is processed by the Depth Gradient Analysis Module to produce the normalized depth complexity map $C(p) \in [0,1]$. Subsequently, $C(p)$ is combined with the depth values $D(p)$ and the local texture information in the Perceptual Depth Salience Model to yield the perceptual salience map $S(p)$ and the discrete adaptive density map $A(p) \in \{1, 2, 3\}$. The density map is then passed to the GPU batch scheduler, which groups contiguous same-tier regions into unified rendering batches. Finally, tiered light field rendering with boundary blending at inter-tier transitions produces the output light field. The system architecture is illustrated in Fig. 1.

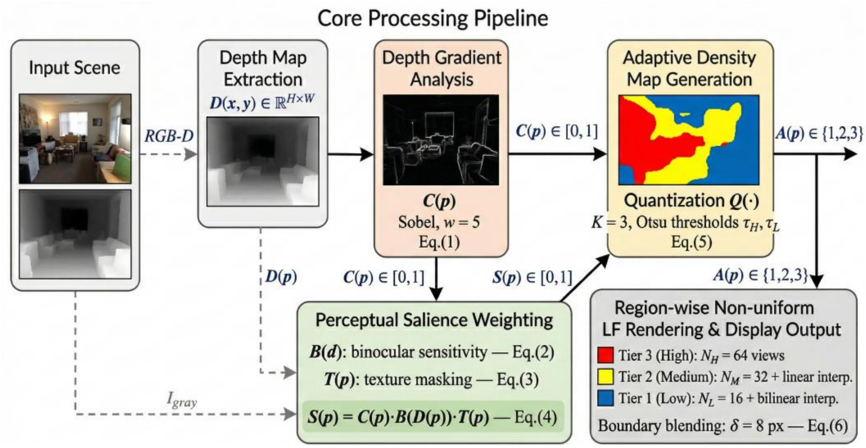

Fig. 1. System architecture of the proposed rendering density-adaptive light field display system.

2.2 Depth Gradient Analysis Module

Angular sampling density dictates continuity of disparities between adjacent viewpoints, and disruptions occur where there is abrupt change in depth. The actual cue for angular sampling density is the magnitude of the depth gradient and not the actual value of the depth. The module employs a Sobel operator on the scene's depth map $D(x, y)$ to compute spatial gradients, chosen for its effective balance between edge localization accuracy and noise suppression. A local depth complexity index $C(p)$ is defined as the mean aggregation of gradient magnitudes over a local $w \times w$ neighborhood $\Omega(p)$, formulated as

$$DCI(x, y) = \frac{1}{w^2} \sum_{(i,j) \in \Omega_w(x,y)} \|\nabla D(i, j)\| \tag{1}$$

where $\nabla D(p)$ denotes the spatial gradient vector and $\| \nabla D(p)\|$ its Euclidean norm. The neighborhood size is set to $w = 5$ based on empirical evaluation. The resulting $C(p)$) values are mapped to [0, 1] through scene-level min-max normalization.

2.3 Perceptual Depth Salience Model

Not all depth gradients are of equal perceptual importance. Depth boundaries in the near range produce larger binocular disparity variations than those in the far range, and high-texture regions partially mask the perception of depth discontinuities. The perceptual depth salience $S(p)$ is obtained by modulating $C(p)$ with two factors.

The first is a depth-dependent binocular sensitivity function:

$$B(d) = \frac{}{1+\left(\frac{d}{d_0}\right)^{\alpha}} \tag{2}$$

where d denotes the normalized depth value in [0, 1], $d_0 = 0.4$ is the half-attenuation depth, and $\alpha = 2.0$ controls the steepness of attenuation. This monotonically decreasing form reflects the psychophysical finding that stereoscopic acuity decreases nonlinearly with viewing distance [9]. At $d = 0$ (nearest depth), $B = 1$; at $d = 1$ (farthest depth), $B \approx 0.14$, representing a reduction by a factor of approximately seven across the full depth range.

The second is a texture masking factor:

$$T(p) = \frac{1}{1+\gamma \cdot \hat{\sigma}^2_{\text{tex}}(p)} \tag{3}$$

where $\hat{\sigma}^2_{\text{tex}}(p)$ is the local grayscale variance computed within the same $w \times w$ window, normalized to [0, 1], and $\gamma = 5.0$ controls the masking strength. In uniform regions, $T \approx 1$, preserving the full salience of depth edges. In highly textured regions, T decreases toward 0.17, reflecting the reduced perceptibility of depth transitions under texture masking [10].

The combined perceptual depth salience is then:

$$S(p) = C(p) \cdot B(D(p)) \cdot T(p) \tag{4}$$

The continuous salience values are quantized into $K = 3$ discrete angular sampling tiers:

$$A(p) = Q(S(p)) = \begin{cases} 3 \text{ (High)}, & \text{if } S(p) \geq \tau_H \\ 2 \text{ (Medium)}, & \text{if } \tau_L \leq S(p) < \tau_H \\ 1 \text{ (Low)}, & \text{if } S(p) < \tau_L \end{cases} \tag{5}$$

The thresholds τ_H and τ_L are determined by applying two-level Otsu thresholding to the histogram of $S(p)$ for each scene. A minimum tier occupancy constraint of 10% is enforced to prevent degenerate cases where the salience histogram is heavily skewed.

2.4 Adaptive Light Field Rendering Pipeline

The adaptive density map $A(p)$ guides a three-tier rendering pipeline. Tier 3 (high-density) regions are rendered with $N_H = 64$ viewpoints at full angular resolution. Tier 2 (medium-density) regions are rendered with $N_M = 32$ viewpoints, with intermediate views reconstructed via linear interpolation. Tier 1 (low-density) regions are rendered with $N_L = 16$ viewpoints, with missing views generated through bilinear interpolation. The halving scheme ($64 \rightarrow 32 \rightarrow 16$) ensures that interpolated angular positions align exactly with the midpoints between rendered viewpoints.

A GPU batch scheduler groups spatially contiguous regions sharing the same tier label into unified draw call batches. At tier boundaries, an overlap band of width $\delta = 8$ pixels is established to ensure seamless visual transitions. Within this band, the rendered viewpoints from adjacent tiers i and j are blended as:

$$L_{\text{blend}}(p) = \lambda(p) \cdot L_i(p) + (1 - \lambda(p)) \cdot L_j(p) \quad (6)$$

where $\lambda(p)$ varies linearly from 1 to 0 as pixel p transitions from the interior of tier i to the interior of tier j across the overlap band.

3 Experiments

3.1 Experimental Setup

Four synthetic scenes with graduated depth complexity (single plane, dual-layer, multiple objects, complex indoor) and two real scenes are used in the experiment. All the experiments are done on a light field display simulation platform with the capability of rendering 64 viewpoints at 1920 × 1080 resolution with a 60° horizontal field of view, where the multi-view angular sampling is designed to minimize accommodation-vergence conflicts. The baseline methods are Uniform Full-Density rendering, Fixed Three-Level rendering, and the content-adaptive rendering algorithm [11]. The evaluation criteria are the mean opinion score (MOS) with a five-level scale and 20 participants, rendering time in milliseconds, PSNR, and SSIM.

3.2 Results and Analysis

The quantitative results across all test scenes are summarized in Table 1.

Table 1. Quantitative comparison of rendering methods.

Method	MOS↑	Time(ms)↓	PSNR(dB)↑	SSIM↑
Uniform Full-Density	3.51	187.4	34.82	0.961
Fixed Three-Level	3.24	131.6	32.47	0.938
Content-adaptive	3.72	126.3	33.85	0.953
Proposed	4.04	115.8	34.26	0.957

The proposed method achieves the highest MOS value of 4.04, which is an improvement of 15.1% over the uniform scheme, while simultaneously reducing the rendering time by 38.2%, from 187.4 ms to 115.8 ms. The PSNR and SSIM of the proposed method are 34.26 dB and 0.957, slightly lower than those of the uniform full-density scheme (34.82 dB and 0.961). This marginal decrease is attributable to the reduced angular sampling in Tier 1 regions, where bilinear interpolation introduces minor reconstruction errors. These low-density regions correspond to perceptually smooth areas where the human visual system exhibits low sensitivity to angular interpolation artifacts. By reallocating rendering density from perceptually insignificant smooth regions to perceptually critical depth boundaries, the proposed method concentrates angular sampling where observers are most sensitive, resulting in higher MOS despite slightly lower pixel-level metrics.

The adaptive density allocation maps overlaid on scene depth maps are depicted in Fig. 2.

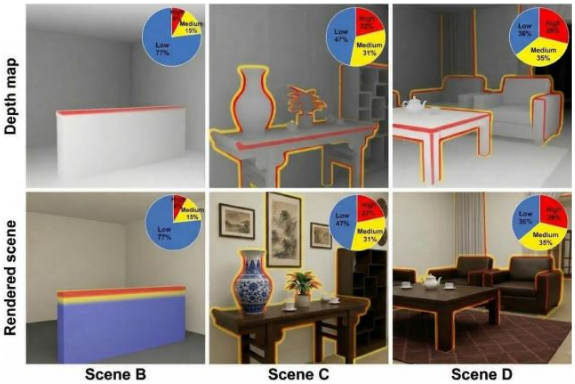

Fig. 2. Adaptive density allocation maps overlaid on depth maps (top) and rendered scenes (bottom) across three test scenes.

To validate the independent contribution of each module, three ablation configurations are tested as indicated in Table 2.

Table 2. Ablation study results.

Configuration	MOS↑	Time(ms)↓	SSIM↑
Full model	4.04	115.8	0.957
w/o Depth gradient analysis	3.89	162.3	0.954
w/o Perceptual weighting	3.58	118.5	0.949
w/o Boundary blending	3.96	114.2	0.931

Removing depth gradient analysis produces the greatest increase in rendering time (115.8 → 162.3 ms), validating its role in efficiency optimization. Removing perceptual weighting causes the largest MOS drop (4.04 → 3.58), confirming the perceptual model as the primary driver of depth enhancement. Removing boundary blending yields the largest SSIM decrease (0.957 → 0.931) with minimal rendering time change, validating its importance in visual continuity between tiers.

4 Conclusion

This paper presents an adaptive rendering density system driven by depth gradient analysis and a depth perception salience model for light field displays. The system allocates angular sampling density according to scene depth complexity modulated by human perceptual sensitivity. Experimental results indicate that the proposed system improves depth perception MOS by 15.1% and reduces rendering computation by 38.2%, with viewpoint reconstruction quality comparable to the full density strategy. A limitation of the current system is the validation on static scenes only. Future work will develop inter-frame temporal consistency strategies for dynamic light field video content and integrate eye-tracking technology to enable viewer-adaptive rendering.

References

1. Huang, F., Chen, K., Wetzstein, G.: The light field stereoscope: immersive computer graphics via factored near-eye light field displays with focus cues. ACM Trans. Graph. 34(4), 60:1–60:12 (2015)https://doi.org/10.1145/2766922
2. Wetzstein, G., Lanman, D., Hirsch, M., Raskar, R.: Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. ACM Trans. Graph. 31(4), 80:1–80:11 (2012)https://doi.org/10.1145/2185520.2185576
3. Held, R.T., Cooper, E.A., Banks, M.S.: Blur and disparity are complementary cues to depth. Curr. Biol. 22(5), 426–431 (2012)https://doi.org/10.1016/j.cub.2012.01.033
4. Yang, Z., Liu, B., Song, Y., Yi, L., Xiong, Y., Zhang, Z., Yu, X.: DirectL: efficient radiance fields rendering for 3D light field displays. ACM Trans. Graph. 43(6), 1–19 (2024) https://doi.org/10.1145/3687897
5. Wang, T., Yang, C., Chen, J., Zhao, Y., Zong, J.: Naked-eye light field display technology based on mini/micro light emitting diode panels: a systematic review and meta-analysis. Sci. Rep. 14, 24381 (2024)https://doi.org/10.1038/s41598-024-75172-z
6. Coomans, A., Dominici, E.A., Döring, C., Mueller, J.H., Hladky, J., Steinberger, M.: Real-time neural rendering of dynamic light fields. Comput. Graph. Forum 43(2), e15014 (2024)https://doi.org/10.1111/cgf.15014
7. Hoffman, D.M., Girshick, A.R., Akeley, K., Banks, M.S.: Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. J. Vis. 8(3), 33:1–33:30 (2008) https://doi.org/10.1167/8.3.33
8. 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. 33(4), 263–273 (2025)https://doi.org/10.1002/jsid.2030
9. Li, S., He, S., Dong, Y., Dai, C., Liu, J., Wang, Y., Shigemasu, H.: Depth perception based on the interaction of binocular disparity and motion parallax cues in three-dimensional space. Sensors 25(10), 3171 (2025)https://doi.org/10.3390/s25103171
10. Mahmoudpour, S., Pagliari, C., Schelkens, P.: Learning-based light field imaging: an overview. EURASIP J. Image Video Process. 2024, 15 (2024) https://doi.org/10.1186/s13640-024-00628-1
11. Wu, G., Masia, B., Jarabo, A., Zhang, Y., Wang, L., Dai, Q., Chai, T., Liu, Y.: Light field image processing: an overview. IEEE J. Sel. Top. Signal Process. 11(7), 926–954 (2017)https://doi.org/10.1109/JSTSP.2017.2747126