Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD30294

Light Field Image Processing: A Comprehensive Overview

Zeheng Li

Beijing Normal-Hong Kong Baptist University, Zhuhai, China
s230026088@mail.uic.edu.cn

Abstract. Light field imaging repress ents a paradigm shift in computational photography, capturing not only the intensity of light but also its direction within a scene. This capability enables a host of post-capture manipulations, such as digital refocusing, depth estimation, and perspective shifting, which are impossible with traditional photography. However, the high-dimensional nature of light field data poses significant challenges in acquisition, processing, compression, and display. This paper provides a comprehensive overview of the field of light field image processing, and through introducing the fundamental theory behind light fields, including the plenoptic function and its representations. This paper then systematically review and compare the primary methods for light field acquisition. Subsequent sections explore in depth core processing tasks: super-resolution across spatial, angular, temporal, and spectral domains; depth estimation using various algorithmic approaches; and quality assessment paired with compression techniques tailored for light field data. This also highlight key application areas, from computational photography to scientific imaging like light field microscopy. Finally, this paper discuss the persistent challenges and future research directions, emphasizing the trade-offs between resolution dimensions and the need for efficient algorithms. This review aims to serve as a valuable resource for researchers and engineers entering this rapidly evolving field.

Keywords: Light Field, Plenoptic Function, Image Processing, Super-Resolution, Depth Estimation, Compression, Computational Photography

1. Introduction

Traditional photography, despite its advancements, captures only a two-dimensional projection of a scene, integrating the radiance along all directions onto a sensor plane. This process inherently discards the angular distribution of light rays, a critical piece of information that defines how we perceive depth, parallax, and focus. This limitation restricts post-capture editing and scene understanding. Light field imaging, also known as plenoptic imaging, overcomes this fundamental constraint. By capturing the light field—the function that describes the amount of light traveling in every direction through every point in space—it records a richer representation of the visual information within a scene. The concept was formally introduced to computer graphics by Levoy and Hanrahan and by Gortler et al. in 1996, though its roots trace back to the plenoptic function of Adelson and Bergen in 1991 [1-3]. A light field is typically modeled as a 4D function, L(u, v, s, t), where (u, v) and (s, t) represent the coordinates on two parallel planes, effectively parameterizing

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD30294

the position and direction of each light ray. The advantages of this approach are profound. It enables computational techniques for digital refocusing, where the focal plane can be chosen after the photo is taken; aperture adjustment; and the generation of novel viewpoints. Furthermore, the dense multi-view nature of light fields facilitates highly accurate depth estimation and the reconstruction of occluded objects. Despite these powerful advantages, the transition from theory to practical application faces several hurdles. The central challenge is the "data explosion" problem: a 4D light field contains orders of magnitude more data than a conventional 2D image. This impacts every stage of the pipeline: acquisition requires specialized and often expensive hardware; processing demands high computational power; storage and transmission necessitate efficient compression schemes. Furthermore, the inherent trade-off between spatial and angular resolution means that increasing the number of viewpoints often comes with the cost of reducing the resolution of each individual view. This paper provides a structured overview of the entire light field image processing pipeline. We will explore the theory, how light fields are captured, the algorithms that enhance and extract information from them, how their quality is measured and data reduced, and where they are applied, concluding with an outlook on future challenges.

2. Light field theory

The theoretical foundation of light field imaging is the plenoptic function. This high-dimensional function aims to describe everything that is visible from any point in space, at any time, for any wavelength. The full plenoptic function is a 7D scalar function: $P = P(\theta, \varphi, \lambda, t, Vx, Vy, Vz)$, representing intensity as a function of direction (θ, φ), wavelength (λ), time (t), and viewer position (Vx, Vy, Vz). For practical static imaging scenarios, this is often reduced to a 4D function—the light field—assuming monochromatic or RGB capture and a fixed time instant. The most common parameterization is the two-plane parameterization, where a light ray is defined by its intersections with two parallel planes: the (u, v) plane (aperture plane) and the (s, t) plane (sensor or image plane). Thus, the light field is represented as L(u, v, s, t). A powerful tool for analyzing and visualizing 4D light fields is the Epipolar Plane Image (EPI). An EPI is a 2D slice of the 4D light field, typically generated by fixing one spatial coordinate (e.g., v) and one angular coordinate (e.g., u), and varying the other two (s and t). In an EPI, a point in the scene manifests not as a point but as a straight line. Crucially, the slope of this line is directly proportional to the depth of the corresponding point in the scene. This property makes EPIs a cornerstone for many geometry-based light field processing algorithms, particularly for depth estimation [5]. Visualizing a 4D dataset is non-trivial. Common methods include displaying the light field as a grid of subaperture images (each is a 2D image from a specific viewpoint, obtained by fixing (u, v) and varying (s, t)) or as a set of EPIs. These representations serve as the basis for both human interpretation and algorithmic processing of light field data.

3. Light field acquisition

Capturing the 4D light field requires innovative hardware designs that go beyond a single conventional sensor. Acquisition strategies can be broadly categorized into three main approaches, each with unique trade-offs.

DOI: 10.54254/2755-2721/2025.LD30294

3.1. Multi-sensor systems (camera arrays)

This approach uses an array of cameras, each capturing the scene from a slightly different viewpoint simultaneously. The advantages are significant: instantaneous capture of the entire light field, making it suitable for dynamic scenes. Furthermore, each camera can be a high-resolution sensor, potentially yielding a high-quality light field. The primary disadvantages are the large physical size of the array, the complexity of calibrating all cameras, and the very high cost associated with multiple high-end sensors and syncing systems.

3.2. Time-sequential acquisition (gantry-based systems)

This method uses a single camera that is moved to different positions over time to capture the various viewpoints sequentially. The main advantage is the very low system cost, as it requires only one camera. However, the capture process is time-consuming, making it completely unsuitable for any dynamic scene or live recording. It is effectively limited to static objects and controlled environments, such as studio-based object scanning.

3.3. Multiplexed acquisition (plenoptic cameras)

This is the most common consumer-grade approach, exemplified by Lytro and Raytrix cameras [4]. A single main lens images the scene onto a microlens array placed just in front of the sensor. Each microlens effectively acts as a tiny camera, sampling the angular distribution of rays arriving at that spatial point. The 4D light field is thus optically encoded into a single 2D sensor image (a light field raw image or lenslet image), which must be decoded computationally. The key advantage is the compact form factor and the ability to capture a light field in a single snapshot. The fundamental disadvantage is the inherent spatial-angular resolution trade-off. The total number of pixels is fixed; allocating more to angular resolution (more microlenses) reduces spatial resolution (pixels per microlens image), and the reverse is also true. This limits the resolution of each subaperture image compared to a traditional photo.

4. Light field super-resolution

The often-limited resolution of captured light fields, especially from plenoptic cameras, has spurred extensive research into super-resolution (SR) techniques. SR for light fields is more nuanced than for 2D images due to the availability of complementary information across different viewpoints.

4.1. Spatial super-resolution

This aims to enhance the resolution of each individual subaperture image. Techniques often leverage the high degree of redundancy across angular views. Algorithms can fuse information from multiple low-resolution views of the same scene to reconstruct a high-resolution image, using techniques derived from multi-frame image super-resolution [5].

4.2. Angular super-resolution

Also known as view synthesis, this task involves generating novel intermediate viewpoints that were not physically captured. This is crucial for creating smooth parallax effects and increasing the angular sampling rate. Methods range from depth-image-based rendering (DIBR), where a depth

DOI: 10.54254/2755-2721/2025.LD30294

map is used to warp pixels from captured views to novel ones, to learning-based approaches that directly synthesize new views from a sparse set of inputs [6].

4.3. Temporal super-resolution (frame interpolation)

Capturing high-frame-rate light field video is extremely challenging due to the massive data throughput required. Temporal SR addresses this by applying video frame interpolation techniques to a low-frame-rate light field video, generating intermediate frames to improve temporal smoothness.

4.4. Spectral reconstruction (hyperspectral light fields)

Directly capturing a hyperspectral light field (adding the wavelength dimension) would exacerbate the resolution trade-off. An alternative is to capture a standard RGB light field and a separate lower-dimensional hyperspectral image, and then use computational methods to reconstruct a full hyperspectral light field, effectively performing super-resolution in the spectral domain.

5. Light field depth estimation

The dense multi - view nature of light fields makes them exceptionally well - suited for accurate depth estimation [7]. The disparities between corresponding pixels in different subaperture images provide a strong cue for depth calculation.

5.1. Multi-view stereo matching

Traditional stereo matching can be extended to the multiple views of a light field. By comparing blocks of pixels across subaperture images, correspondences can be found, and depth can be triangulated. The abundance of views allows for robust matching and noise suppression.

5.2. Epipolar Plane Image (EPI) analysis

This is a distinctive and powerful approach unique to light fields. As mentioned in Section 2, a scene point corresponds to a line with a specific slope in an EPI. Depth estimation then becomes the task of estimating these slopes, often using structure tensor analysis or deep learning models trained on EPI patches [5,8].

5.3. Deep learning-based methods

Convolutional Neural Networks (CNNs) have become dominant in this area. Networks can be designed to take a stack of subaperture images or an EPI volume as input and directly regress to obtain a depth map. These data-driven methods often outperform traditional algorithms, especially in handling occlusions and textureless regions, by learning priors from large datasets [6].

5.4. Focal stack analysis

Some methods use the refocusing capability of light fields to generate a focal stack—a series of images focused at different depths. The depth for a pixel can then be inferred by finding the image in the stack where that pixel is most in focus (i.e., has the highest sharpness).

DOI: 10.54254/2755-2721/2025.LD30294

6. Light field assessment and compression

The massive volume of light field data makes it necessary to have robust quality assessment metrics and efficient compression schemes. The quality assessment of light field data involves several metrics, including spatial quality, angular consistency, and spatial-angular resolution. Standard 2D image quality metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are used to evaluate spatial fidelity. Angular consistency measures the consistency of views across the angular dimension, penalizing artifacts like ghosting or broken edges. Effective resolution is measured using charts like ISO12233, considering the interplay between the two dimensions. Compression methods include pseudo-sequence coding, transform-based coding, and lenslet image compression. Lossless compression is desirable for archives, while high-ratio lossy compression is necessary for practical storage and streaming due to high redundancy in light fields [9].

7. Light field applications

Light field technology has applications beyond research labs, including computational photography, 3D reconstruction, VR/AR, occlusions removal and see-through imaging, light field microscopy in biology, and industrial inspection and robotics. It enables post-capture refocusing, aperture adjustment, perspective shift in consumer cameras, accurate depth maps for 3D models, algorithmic reconstruction of occluded areas, real-time 3D tracking of dynamic processes, and precise object measurement, defect detection, and improved scene understanding for robotic navigation and manipulation [10].

8. Conclusion

Light field imaging has matured significantly, transitioning from a theoretical concept to a technology with real-world applications. This paper have reviewed its core principles, ranging from acquisition and super-resolution to depth estimation and compression. Despite this progress, several challenges remain open for future research. The most fundamental challenge is the inherent trade-off among spatial, angular, temporal, and spectral resolutions. Breaking this trade-off requires innovations at both the hardware level and the software level. Data efficiency and computational complexity are persistent hurdles. Processing 4D/5D data is inherently expensive. Future work must focus on developing highly efficient algorithms, perhaps by leveraging neural networks with specialized architectures for light field data and model compression techniques. On the hardware front, the development of compact, low-cost, high-resolution light field capture systems is crucial for widespread adoption. Research into micro-sized camera arrays and next-generation plenoptic sensor designs is ongoing. Finally, expanding the range of application fields is a key direction. Interdisciplinary research will unlock new uses, particularly in scientific imaging (like medical endoscopy and astronomy), robotics, and immersive telepresence. As the underlying technology becomes more powerful and accessible, the potential for light fields to revolutionize how we capture, interact with, and understand visual information is immense.

References

[1] Adelson, E. H., & Bergen, J. R. (1991). The plenoptic function and the elements of early vision. Computational models of visual processing, 1(2), 3-20.

[2] Levoy, M., & Hanrahan, P. (1996). Light field rendering. Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 31-42.

DOI: 10.54254/2755-2721/2025.LD30294

[3] Gortler, S. J., Grzeszczuk, R., Szeliski, R., & Cohen, M. F. (1996). The lumigraph. Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 43-54.
[4] Ng, R., Levoy, M., Brédif, M., Duval, G., Horowitz, M., & Hanrahan, P. (2005). Light field photography with a hand-held plenoptic camera. Computer Science Technical Report CSTR, 2(11), 1-11.
[5] Wanner, S., & Goldluecke, B. (2014). Variational light field analysis for disparity estimation and super-resolution. IEEE transactions on pattern analysis and machine intelligence, 36(3), 606-619.
[6] Wang, T. C., Zhu, J. Y., Hiroaki, E., Efros, A. A., & Ramamoorthi, R. (2017). A 4D light field dataset and CNN architectures for material recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(5), 1040-1050.
[7] Wu, G., Masia, B., Jarabo, A., Zhang, Y., Wang, L., Dai, Q., … & Liu, Y. (2017). Light field image processing: A survey. IEEE Journal of Selected Topics in Signal Processing, 11(7), 926-954.
[8] Zhang, S., Sheng, H., Li, C., Zhang, J., & Xiong, Z. (2019). Robust depth estimation for light field via spinning parallelogram operator. Computer Vision and Image Understanding, 187, 102771.
[9] Conti, C., Soares, L. D., & Nunes, P. (2020). Light field image coding survey: A taxonomy of methods and performance evaluation. IEEE Access, 8, 14199-14226.
[10] Jin, J., Hou, J., Chen, J., & Kwong, S. (2021). Light field super-resolution via attention-based fusion of hybrid lenses. IEEE Transactions on Circuits and Systems for Video Technology, 31(11), 4334-4346.