



특집논문 (Special Paper)

방송공학회논문지 제30권 제6호, 2025년 11월 (JBE Vol.30, No.6, November 2025)

<https://doi.org/10.5909/JBE.2025.30.6.952>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

AI기반 영상 향상 및 광학 보정 기술을 적용한 저저항 단일 화소 센서의 3D 집적 영상 이미징 및 디스플레이

조 수 연^{a)b)}, 윤 선 규^{c)†}

AI-based Image Enhancement and Optical Distortion Correction for 3D Integral Imaging and Display using Fermi-Level Pinning-Free and Low-Resistance Single-Pixel Imager

Suyeon Jo^{a)b)} and Seon Kyu Yoon^{c)†}

요 약

본 연구는 페르미 준위 고정(Fermi-level pinning, FLP)이 억제된 저저항 단일 화소 센서를 기반으로, 3차원 집적 영상에서 발생하는 해상도 저하와 기하학적 왜곡 문제를 개선하기 위한 영상 향상 프레임워크를 제안한다. 해당 센서는 안정적인 동작 특성을 지니고 있으나, 저해상도, 스캐닝 노이즈, 기하학적 왜곡 등으로 3D 집적 영상에 응용하기에는 제약이 존재한다. 이를 개선하기 위해 본 연구에서는 노이즈 억제와 딥러닝 기반 초해상도 복원을 결합하고, 추가적인 광학적 보정 단계를 도입하여 영상 품질과 재구성 정확도를 향상시켰다. 본 프레임워크는 데이터가 제한된 조건에서도 안정적으로 동작하며, 기존 단일 화소 기반 영상에서 발생하는 품질 저하 문제를 효과적으로 감소시켰다. 실험 결과, 영상의 해상도와 구조적 세부 표현이 뚜렷하게 향상되었으며, 시각적 비교에서도 재구성된 3차원 집적 영상이 공간적 일관성과 깊이감 측면에서 개선된 모습을 확인할 수 있었다.

Abstract

This study proposes an image enhancement framework for three-dimensional (3D) integral imaging using low-resistance single-pixel sensors with suppressed Fermi-level pinning (FLP). Although these sensors provide stable charge transport and reliable operation, their application to 3D imaging is constrained by low spatial resolution, scanning-induced noise, and geometric distortion. To address these limitations, the framework integrates noise suppression, deep learning-based super-resolution (SR), and an optical correction stage. Noise suppression stabilizes image acquisition under limited signal conditions, while SR restores fine structural details. In addition, optical correction reduces distortion and improves inter-view alignment, enabling more accurate 3D reconstructions. The framework operates robustly even in the absence of ground-truth reference data, mitigating the degradation typically observed in conventional single-pixel imaging. Experimental results confirm significant improvements in perceptual quality, and the reconstructed 3D scenes exhibit enhanced spatial consistency and realistic depth. Overall, the results demonstrate the feasibility of compact single-pixel-based 3D imaging systems for high-fidelity integral visualization.

Keyword : Fermi-level pinning (FLP)-free, Single-pixel imaging, Integral imaging, Super-resolution enhancement, Optical distortion correction

Copyright © 2025 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

I. Introduction

The Fermi level defines the energy distribution of electrons in solids and underpins charge transport phenomena. At metal-semiconductor interfaces, however, Fermi-level pinning (FLP) often arises and prevents the formation of an ideal Schottky barrier height (SBH). In principle, the SBH can be tuned by selecting metals with different work functions. Yet interface-induced states—such as metal- and defect-induced gap states (MIGS and DIGS)—fix the Fermi level at specific energies. As a result, large SBHs form regardless of the metal’s work function, reducing charge collection and fundamentally limiting the performance of 2D semiconductor optoelectronic devices.

In our previous work^[1], we addressed this challenge by implementing a photodiode structure designed to suppress FLP, thereby improving charge transport and signal stability. To validate the device, we employed a single-pixel imaging setup in which viewpoint images were sequentially acquired through mechanical scanning^[2,3]. These elemental images were reconstructed into 3D scenes using a microlens array-based integral imaging technique, which synthesizes depth from multiple perspectives.

While this configuration effectively demonstrated the sensor’s signal stability and feasibility for 3D imaging, several practical challenges emerged during implementation. As a single-pixel imaging approach, the method offers ad-

vantages of low cost, simple architecture, and spectral flexibility. Yet the inherently low photocurrent necessitates sequential scanning, during which signal fluctuations accumulate and manifest as acquisition noise. In addition, optical distortions—particularly pincushion effects on the display side—compromised reconstruction fidelity. Collectively, these issues motivated a system-level enhancement strategy to improve image quality and spatial consistency in single-pixel integral imaging.

To address these challenges, this study extends our previous work by proposing a system-level enhancement framework based on a low-resistance, FLP-mitigated single-pixel imager. The proposed approach integrates computational and optical strategies to improve the quality of 3D reconstruction in integral imaging systems. A two-stage computational pipeline combines non-local filtering with deep learning-based SR to suppress noise and restore spatial detail. In parallel, geometric distortion is reduced through optical design, using a fixed-focal-length Fresnel lens and optimized lens-display spacing. This integrated strategy enables accurate 3D reconstruction with consistent inter-view alignment. An overview of the system architecture is provided in Fig 1.

II. Methods

1. 3D Integral Imaging Setup

A single-pixel photodiode previously developed with a conductive-bridge interlayer contact (CBIC) structure was employed as the imaging sensor^[1]. This architecture suppresses Fermi-level pinning by mitigating metal- and defect-induced gap states, thereby enabling stable charge transport suitable for reliable 2D/3D image acquisition. For elemental image acquisition, the photodiode was mechanically translated across a uniform grid using a motorized XY stage, while a stationary object was illuminated by a

a) 고려대학교 전기전자공학부(School of Electrical Engineering, Korea University)

b) 한국과학기술연구원(Korea Institute of Science and Technology)

c) 한국광기술원(Korea Photonics Technology Institute)

‡ Corresponding Author : 윤선규(Seon Kyu Yoon)

E-mail: skyo2n@kopti.re.kr

Tel: +82-62-605-9255

ORCID: <https://orcid.org/0000-0002-5050-4675>

※ 이 논문의 결과 중 일부는 한국방송·미디어공학회 2025년 하계학술대회에서 발표한 바 있음

※ This work was supported by the Technology Innovation Program (02316245) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea)

· Manuscript September 15, 2025; Revised October 31, 2025; Accepted November 3, 2025.

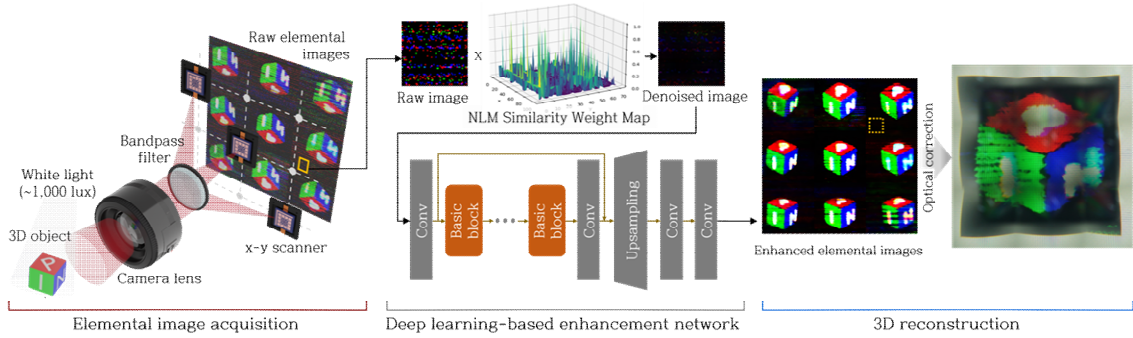


그림 1. 제안된 영상 향상 프레임워크 개요. 시점 영상을 획득한 뒤, 노이즈 억제와 디러닝 기반 초해상도를 통해 영상 품질을 향상시키고, 광학 보정 단계를 거쳐 3차원으로 재구성.

Fig 1. Schematic overview of the proposed framework. Elemental images are acquired, enhanced by denoising and SR, and reconstructed into 3D with optical correction.

white LED (~1000 lux). RGB elemental images were sequentially captured at 450, 550, and 650 nm through narrow bandpass filters, and multiple frames were averaged at each wavelength to improve the signal-to-noise ratio. The resulting dataset was then used to validate the proposed noise suppression, SR enhancement, and optical distortion correction framework.

2. Noise Suppression and Resolution Enhancement

In integral imaging, the quality of elemental images is essential for maintaining spatial consistency and accurate inter-view alignment^[4]. Image degradation—such as noise and reduced resolution—can arise in scanning-based single-pixel acquisition as well as in microlens array capture, especially under data-limited or compact configurations^[5]. To mitigate these issues, a two-stage enhancement pipeline was implemented, combining non-local means (NLM) filtering^[6] with deep learning-based SR.

NLM filtering was adopted for its data-independent nature and ability to exploit repetitive patterns. It is defined as:

$$NL[v](i) = \sum_{j \in I} w(i, j) \cdot v(j) \quad (1)$$

$$w(i, j) = \frac{1}{z(i)} \exp\left(-\frac{|v(N_i) - v(N_j)|_2^2}{h^2}\right). \quad (2)$$

This approach exploits non-local similarity to suppress noise while preserving structural features, making it suitable for data-scarce imaging. Although it effectively reduces noise, it can also attenuate fine edges and diminish inter-view distinctiveness. To complement these limitations, a deep learning-based SR method was employed, as it can restore high-frequency details by utilizing learned structural priors beyond the capacity of classical filtering. To restore high-frequency details, SR was therefore applied as a second stage.

SR models are abundant, but not all are suitable for single-pixel imaging. Because the reconstructed data directly reflect the device response, post-processing must avoid introducing artificial distortions. This requires methods that preserve structural fidelity, maintain geometric consistency, and remain robust to noise. Real-ESRGAN^[7] is particularly suited to this context, as its blind SR framework delivers stability under noisy and distorted conditions while preserving spatial coherence. The generator, based on residual-in-residual dense blocks (RRDBs), is trained with a blind degradation model that incorporates a range of noise levels, blur operators, and downsampling kernels, provid-

ing robustness to measurement artifacts. Perceptual and adversarial losses further enable the recovery of fine textures without compromising structural fidelity. In this study, elemental images denoised by NLM filtering were processed with a pretrained Real-ESRGAN model, and the outputs were reassembled to yield super-resolved elemental images suitable for reliable 3D reconstruction.

3. Geometric Distortion Minimization by Optical Design

Geometric distortion is a persistent challenge in integral imaging, particularly pincushion distortion at wide viewing angles. Lenses with a wide field of view (FOV) tend to exhibit stronger nonlinear distortions, such as barrel and pincushion effects, since the FOV is determined by the magnification and focal length of the optical system^[8]. Conventional convex lenses are cost-effective and widely used, but they are prone to distortion. Aspheric lenses mitigate such distortion by correcting for nonlinear aberrations, but are bulky and costly. Fresnel lenses, by contrast, are thin, lightweight, and easily integrated, making them suitable for compact configurations.

A fixed-focal-length Fresnel lens was adopted, and the distance between the lens and the display was adjusted to optimize system magnification. Because magnification directly affects both the viewing angle and the degree of geometric distortion, small deviations in spacing can lead to

misalignment across viewpoints. We therefore tuned the spacing experimentally to minimize pincushion distortion while maintaining consistent inter-view registration. Proper adjustment of this spacing reduced magnification-induced distortion and stabilized inter-view alignment. This optical strategy enhanced spatial registration and reconstruction fidelity without additional computational processing, contributing to clearer and more reliable 3D reconstructions.

III. Results

Perceptual quality was evaluated using the Perceptual Index (PI)^[9], a no-reference metric well suited to single-pixel imaging, where reconstructed images directly reflect device signals and ground-truth references are unavailable. PI has been widely acknowledged as a practical perceptual metric^[10-12]. PI linearly combines two no-reference image quality measures, NIQE (Naturalness Image Quality Evaluator) and the Ma score. The Ma score is a learned perceptual quality metric that estimates the human-perceived naturalness and aesthetic preference of a single image; higher Ma values indicate images that are visually more pleasing and less degraded^[13]. NIQE quantifies the deviation from natural scene statistics; lower NIQE values indicate fewer distortions and a more natural appearance^[14].

PI is defined as:

표 1. 다양한 노이즈 억제 방법 및 NLM으로 노이즈 억제 후 적용한 초해상도 모델 간 시각적 품질 비교. PI 값이 낮을수록 우수한 화질을 의미함. Right/Middle/Left는 대표 시점이며, All-viewpoints Avg는 전체 획득 시점의 평균 PI를 뜻하며 시점 전반에서 일관된 개선이 확인됨.

Table 1. Quantitative comparison of perceptual quality across denoising methods and SR methods applied after NLM denoising. Lower PI indicates better perceptual quality. Right/Middle/Left are representative viewpoints. All-viewpoints Avg is the average PI over all acquired viewpoints, showing that the improvement is consistent across viewpoints.

PI (↓)	raw	Denoising			SR on Denoising(NLM)			
		Median Filter	BM3D	NLM	SRCNN	EDSR	BSRGAN	Real ESRGAN
Right	16.06	13.43	10.99	9.94	5.51	7.22	5.05	5.03
Middle	13.12	14.24	11.95	13.39	8.57	7.67	6.29	5.93
Left	12.12	12.96	6.96	6.95	6.28	5.80	4.79	4.64
All-viewpoints Avg	13.44	12.89	10.08	7.88	6.43	5.96	5.08	4.80

$$PI = \frac{1}{2}((10 - Ma) + NIQE), \quad (3)$$

such that lower PI values correspond to higher perceptual quality. Consequently, a reduction in PI indicates simultaneous suppression of noise and artifacts and an improvement in perceived visual quality, even in the absence of ground truth.

As summarized in Table 1, results are reported for three representative viewpoints (Right/Middle/Left) and for the average PI across all acquired viewpoints of the integral imaging array. The raw elemental images exhibit the greatest perceptual degradation, with an all-viewpoints mean PI of 13.44, indicating severe noise and distortion. The consistent reduction in PI at both the individual-viewpoint level and the all-viewpoints level indicates that the proposed denoising and SR pipeline improves perceptual quality across the full multi-view set.

Among denoising methods, NLM attains the lowest PI in the Right/Left viewpoints (9.94/6.95), reflecting a favorable balance between noise suppression and structural fidelity. In the Middle viewpoint, BM3D yields a lower PI than NLM, contrary to the pattern observed elsewhere. Nevertheless, the all-viewpoints average confirms that

NLM achieves the lowest PI on average, indicating that the dataset-level behavior is consistent with NLM being the most robust denoiser.

When SR was applied to NLM-denoised images, Real-ESRGAN achieved the greatest improvement, reducing PI to 4.6 - 5.9 across viewpoints. This performance reflects its robustness to residual noise and distortions under a blind SR formulation. For comparison, representative SR models known for minimizing distortion were also evaluated. SRCNN^[15] provided only marginal improvements, whereas EDSR^[16] and BSRGAN^[17] preserved fidelity under clean conditions but degraded in noisy environments. In contrast, Real-ESRGAN restored edges and textures while maintaining cross-view alignment, enabling reliable 3D reconstructions in single-pixel integral imaging.

Qualitative results further support these findings. As shown in Fig 2, raw images exhibit pronounced noise and low resolution, which in turn degrade spatial fidelity and disrupt inter-view alignment. After applying the proposed pipeline, elemental images display sharper edges and richer textures, thereby enabling more consistent 3D reconstruction and enhancing angular coherence across viewpoints.

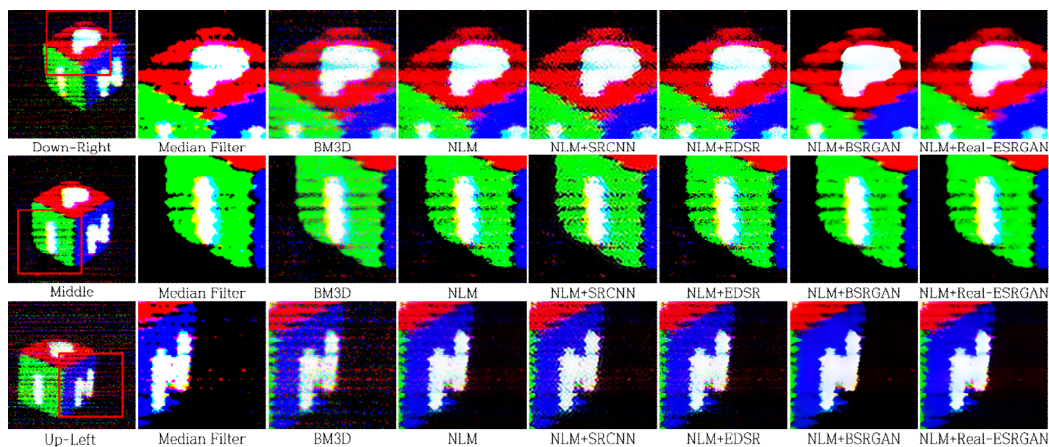


그림 2. 시점별 영상 비교 결과. NLM과 Real-ESRGAN 적용 시, 원본 및 기존 기법 대비 개선된 경계 선명도와 미세구조 복원. Fig 2. Comparison of elemental images across viewpoints. NLM and Real-ESRGAN yield clearer edges and textures compared with raw and baseline methods.

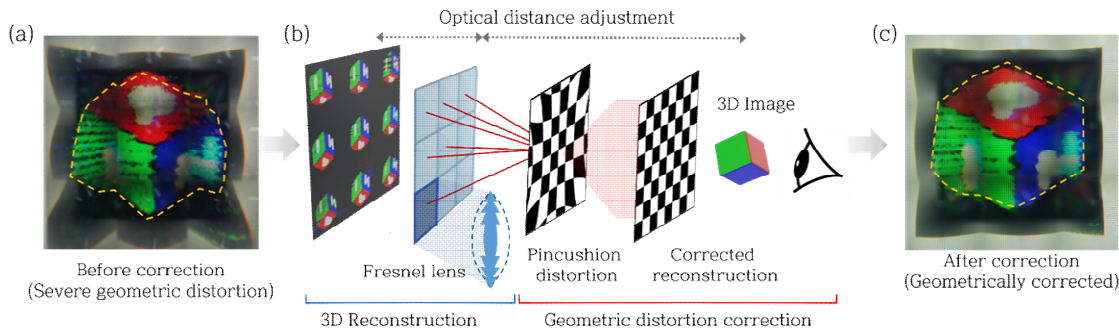


그림 3. 3차원 집적 영상에서의 광학 보정. (a) 보정 전 3차원 재구성 시 심한 핀쿠션 왜곡 관찰. (b) 프레넬 렌즈 기반 광학 설계 및 렌즈와 디스플레이 간격 조정을 통한 기하학적 왜곡 보정. (c) 시점 간 정렬 및 형상 정확도가 향상된 광학 보정 후의 3차원 재구성 영상.
 Fig. 3. Optical correction in 3D integral imaging. (a) 3D reconstruction before correction with severe pincushion distortion. (b) Fresnel-lens-based optical design with geometric correction of elemental images. (c) Reconstruction after correction yielding improved fidelity and alignment.

Fig 3 illustrates the effect of optical correction in 3D integral imaging. In the uncorrected case, pincushion distortion and inter-view misalignment are clearly visible near the image periphery. These artifacts compromise reconstruction fidelity. Optimizing the Fresnel lens-spacing substantially reduced distortion, which in turn improved spatial registration and reconstruction accuracy.

IV. Discussion & Conclusions

In this study, we demonstrated an integrated framework to enhance image quality and correct geometric distortion in 3D integral imaging systems using FLP-suppressed, low-resistance photodiodes. Building on our previous work with CBIC-based single-pixel imagers^[1], we addressed the limitations of low resolution, noise, and distortion through a system-level strategy integrating computational and optical refinements. Noise suppression and deep learning-based super-resolution restored structural details under data-limited conditions, while Fresnel-lens-based optical correction minimized distortion and stabilized inter-view alignment. As a result, the reconstructed 3D images exhibited sharper edges, enhanced spatial coherence, and im-

proved perceptual fidelity.

These findings demonstrate the value of device-level innovations. When combined with computational and optical refinement, they enable accurate and practical autostereoscopic 3D visualization. The proposed framework provides a scalable route for compact and high-quality imaging. It also points to potential applications in wearable and portable 3D visualization platforms.

참고 문헌 (References)

- [1] J. Jang et al., "Conductive-bridge interlayer contacts for two-dimensional optoelectronic devices," *Nature Electronics*, Vol. 8, No. 4, pp. 298 - 308, Feb 2025.
doi: <https://doi.org/10.1038/s41928-025-01339-9>
- [2] T. H. Kim et al., "Self-powering sensory device with multi-spectrum image realization for smart indoor environments," *Advanced Materials*, vol. 36, no. 2, p. 2307523, Nov 2024.
doi : <https://doi.org/10.1002/adma.202307523>
- [3] S. Oh et al., "Robust Imaging through Light-Scattering Barriers via Energetically Modulated Multispectral Organic Photodetectors," *Advanced Materials*, p. 2503868, Apr 2025.
doi: <https://doi.org/10.1002/adma.202503868>
- [4] Y. Huang et al., "Performance enhanced elemental array generation for integral image display using pixel fusion," *Frontiers in Physics*, vol. 9, p. 639117, Apr 2021.
doi: <https://doi.org/10.3389/fphy.2021.639117>
- [5] M.-J. Sun, M. P. Edgar, D. B. Phillips, G. M. Gibson, and M. J. Padgett,

- “Improving the signal-to-noise ratio of single-pixel imaging using digital microscanning,” *Optics express*, vol. 24, no. 10, pp. 10476-10485, May 2016.
doi: <https://doi.org/10.1364/OE.24.010476>
- [6] A. Buades, B. Coll, and J.-M. Morel, “Non-local means denoising,” *Image Processing On Line*, vol. 1, pp. 208-212, Sep 2011.
doi: https://doi.org/10.5201/ipol.2011.bcm_nlm
- [7] X. Wang, L. Xie, C. Dong, and Y. Shan, “Real-esrgan: Training real-world blind super-resolution with pure synthetic data,” in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 1905-1914, 2021.
doi: <https://doi.org/10.1109/iccvw54120.2021.00217>
- [8] R. E. Fischer, B. Tadic-Galeb, and P. R. Yoder Jr, *Optical system design*, Mcgraw-hill, pp.85-89, 2008.
- [9] Y. Blau, R. Mechrez, R. Timofte, T. Michaeli, and L. Zelnik-Manor, “The 2018 PIRM challenge on perceptual image super-resolution,” in *Proceedings of the European conference on computer vision (ECCV) workshops*, pp. 0-0, 2018.
doi: https://doi.org/10.1007/978-3-030-11021-5_21
- [10] X. Luo, R. Chen, Y. Xie, Y. Qu, and C. Li, “Bi-GANs-ST for perceptual image super-resolution,” in *Proceedings of the European conference on computer vision (ECCV) workshops*, pp. 0-0, 2018.
doi: https://doi.org/10.1007/978-3-030-11021-5_2
- [11] K. Zhang, S. Gu, and R. Timofte, “Ntire 2020 challenge on perceptual extreme super-resolution: Methods and results,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp. 492-493, 2020.
doi: <https://doi.org/10.1109/cvprw63382.2024.00624>
- [12] F. Zhang, S. B. Rangrej, T. Aumentado-Armstrong, A. Fazly, and A. Levinshtein, “Augmenting Perceptual Super-Resolution via Image Quality Predictors,” in *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 2311-2322, June 2025.
doi: <https://doi.org/10.1109/cvpr52734.2025.00221>
- [13] C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang, “Learning a no-reference quality metric for single-image super-resolution,” *Computer Vision and Image Understanding*, vol. 158, pp. 1-16, May 2017.
doi: <https://doi.org/10.1016/j.cviu.2016.12.009>
- [14] A. Mittal, R. Soundararajan, and A. C. Bovik, “Making a “completely blind” image quality analyzer,” *IEEE Signal processing letters*, vol. 20, no. 3, pp. 209-212, March 2013.
doi: <https://doi.org/10.1109/lsp.2012.2227726>
- [15] C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295-307, Jun 2015.
doi: <https://doi.org/10.1109/TPAMI.2015.2439281>
- [16] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, “Enhanced deep residual networks for single image super-resolution,” in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 136-144, 2017.
doi: <https://doi.org/10.1109/cvprw.2017.151>
- [17] K. Zhang, J. Liang, L. Van Gool, and R. Timofte, “Designing a practical degradation model for deep blind image super-resolution,” in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 4791-4800, October 2021.
doi: <https://doi.org/10.1109/iccv48922.2021.00475>

저 자 소 개

조 수 연



- 2023년 2월 : 경희대학교 전자공학과 공학사
- 2025년 2월 : 고려대학교 전기전자공학부 공학석사
- 2023년 3월 ~ 현재 : 한국과학기술연구원 차세대반도체연구소 학생연구원
- 2025년 3월 ~ 현재 : 고려대학교 전기전자공학부 박사과정
- ORCID : <https://orcid.org/0009-0000-8138-594X>
- 주관심분야 : 인공지능 기반 영상처리, 광학 기반 스마트 센싱, 홀로그래프 및 AI 기반 3D 디스플레이

윤 선 규



- 2016년 2월 : 고려대학교 물리학과 이학박사
- 2016년 2월 ~ 2017년 4월 : 한국과학기술연구원 영상미디어연구단 박사 후 연구원
- 2017년 5월 ~ 현재 : 한국과학기술원 광응용연구본부 공간광정보연구센터 선임연구원
- ORCID : <https://orcid.org/0000-0002-5050-4675>
- 주관심분야 : 3차원 디스플레이 시스템, 디지털 홀로그래피, 3차원 디스플레이 시스템 특성 측정/분석