

Article

Not peer-reviewed version

Multi-Spectral Reflectance Encoding for Robotic Vision in Metallic Manufacturing Environments

Haruka Tanaka , Kenta Mori , [Shiori Fujimoto](#) *

Posted Date: 2 December 2025

doi: 10.20944/preprints202512.0277.v1

Keywords: reflectance encoding; spectral imaging; metallic perception; robotic vision; illumination robustness



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

Multi-Spectral Reflectance Encoding for Robotic Vision in Metallic Manufacturing Environments

Haruka Tanaka *, Kenta Mori and Shiori Fujimoto

Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan

* Correspondence: haruka.tanaka@u-tokyo.ac.jp

Abstract

Robotic perception in metallic factories is hindered by glare, dynamic reflections, and inconsistent lighting. This paper proposes a multi-spectral reflectance encoding method that captures stable reflectance signatures through narrow-band illumination and spectral normalization. A reflectance-stabilization module reconstructs geometry-consistent features under fluctuating brightness conditions. Evaluations on MetalBench-2025 and SteelLine-Vision datasets show improvements of 21.4% in recognition accuracy and a 33.7% reduction in reflection-induced errors. When deployed on a welding robot, the system improves joint alignment precision by 18.6% and reduces failure rates in reflective surface detection by 27.3%.

Keywords: reflectance encoding; spectral imaging; metallic perception; robotic vision; illumination robustness

1. Introduction

Robotic vision is increasingly deployed in modern manufacturing, yet metallic environments remain among the most challenging settings for reliable perception. Highly reflective surfaces, welding arcs, and polished coatings create strong glare, saturated pixels, and unstable appearance changes that interfere with detection, segmentation, and pose estimation. Studies on metal surface inspection report that even advanced CNN- and Transformer-based detectors suffer substantial performance degradation when illumination angles or backgrounds change, revealing the fragility of intensity-based features under specular conditions [1,2]. On real production lines, where cameras, tools, and fixtures move continuously, these appearance shifts occur frequently and lead to unstable predictions instead of repeatable measurements, which is problematic for closed-loop robotic control. These difficulties are especially evident in robotic welding, machining, and handling of metallic workpieces, where perception must operate near molten pools and highly reflective materials. Weld-tracking systems, for example, must localize seams close to high-temperature regions in which reflections distort projected patterns, laser stripes, or time-of-flight signals, causing sizeable errors in seam geometry estimation [3]. Investigations of multi-sensor welding and metallic bin-picking show that spatter, moving highlights, and mirror-like surfaces can corrupt stripe extraction, edge detection, and depth measurements, ultimately degrading pose estimation and grasp planning [4]. Both passive and active sensors are affected because metallic surfaces behave as dynamic mirrors rather than diffuse reflectors, so small changes in viewpoint or lighting can trigger large variations in the captured image [5]. Prior work on deep-learning-based vision recognition and positioning optimization for industrial robots further confirms that geometric accuracy and cycle-time reliability are highly sensitive to perception noise and instability in such environments [6]. Together, these results indicate that improving downstream algorithms alone is insufficient if the underlying reflectance signal remains unstable.

Lighting control is a common practical workaround in industry. Technical guidelines and case studies describe dome lighting, polarizers, structured light, and carefully tuned exposure as effective ways to suppress glare and stabilize appearance on reflective parts within dedicated workcells [7,8].

However, these solutions typically assume fixed sensor and light positions, rigid fixturing, and limited variation in part pose. Reports from industrial deployments show that recognition performance degrades rapidly once illumination deviates from the calibrated setup, for example when robots operate over larger workspaces or when multiple tools share the same cell [9]. Illumination-invariant sensing has also been explored in outdoor and mobile robotics using strobed light, adaptive exposure control, or active shading, yet these techniques often rely on strict geometric constraints that are difficult to maintain in flexible metallic factories where both sensors and workpieces move [10,11]. As a result, robust perception in metallic environments cannot depend solely on external lighting constraints. Multi-spectral and hyperspectral imaging offer a promising alternative because they capture reflectance at multiple wavelengths and can reveal material-dependent signatures that are less sensitive to intensity fluctuations. Surveys on non-destructive testing show that combining visible, near-infrared, and infrared bands can provide more stable cues for metal inspection under varying lighting [12]. Recent developments in compact spectral cameras and on-chip filters have produced sensors that are increasingly suitable for embedded or robotic use, with reduced size, weight, and power consumption [13]. Multi-spectral fusion has also been examined for 3D reconstruction and material analysis, where spectral cues are combined with stereo, structured light, or shape-from-shading to improve both geometry and surface characterization [14,15]. Nevertheless, most of these systems target static inspection, cultural-heritage imaging, or outdoor scenes with relatively slow motion. They rarely consider the combined challenges of strong specular reflections, moving robots, and continuously changing illumination that characterize metallic production lines. These observations reveal several persistent gaps. Many existing approaches treat reflections purely as noise to be suppressed, rather than as structured signals that could be encoded to obtain stable reflectance descriptors across lighting and viewpoint changes. Despite the advantages of spectral imaging, there is still limited work on leveraging narrow-band or multi-spectral signals to build robust appearance representations tailored to metallic environments. In addition, evaluations are often performed on static or laboratory datasets and pay limited attention to robot-level performance indicators such as joint alignment accuracy, seam-tracking precision, or failure rates in detecting reflective surfaces and edges. These gaps highlight the need for methods that encode reflectance information directly and maintain consistency even when brightness, viewing direction, and workpiece configuration change during operation.

In this study, we introduce a multi-spectral reflectance encoding method designed specifically for metallic manufacturing environments. The proposed system employs narrow-band illumination and spectral normalization to extract stable reflectance signatures that remain consistent under varying brightness and viewing geometry. A reflectance-stabilization module reconstructs wavelength-conditioned features that suppress specular artifacts while preserving task-relevant structure, and a fusion-based encoder integrates these multi-band cues into reflection-aware representations for detection and pose estimation. We evaluate the approach on the MetalBench-2025 and SteelLine-Vision datasets and observe higher recognition accuracy and substantially fewer reflection-induced errors than RGB and broadband baselines. The method is further validated on a welding robot, where it improves joint alignment precision and reduces failures in detecting reflective seams and workpiece boundaries. These results suggest that encoding reflectance at selected wavelengths provides a practical and scalable foundation for more reliable robotic vision in metallic factories, and they point toward a broader class of reflectance-aware perception systems that can support robust, high-precision automation in challenging industrial settings.

2. Materials and Methods

2.1. Dataset Description and Sampling Conditions

This study uses two multi-spectral datasets collected in metallic manufacturing settings: MetalBench-2025 and SteelLine-Vision. MetalBench-2025 contains 32,400 image sets of aluminum, stainless steel, and coated metal parts. SteelLine-Vision includes 27,100 image sets taken along an

operating production line with moving tools and changing lighting. Each set has six narrow-band channels between 450 nm and 950 nm. Recordings were made under three lighting conditions: fixed intensity, varying intensity, and mixed light sources. Surfaces with flat, curved, and welded geometries were included so that different reflective behaviors were represented. Image sets with corrupted channels or missing exposure information were removed during preprocessing.

2.2. Experimental Setup and Control Conditions

Two groups were used for comparison. The multi-spectral group used the proposed reflectance encoding method. The control group used standard RGB images with broadband light. Both groups shared the same camera positions, motion speed, and exposure rules. For robot tests, a welding robot repeated the same joint-tracking task with fixed speed and path. This setup ensures that differences in results come from the sensing method rather than task variations. The control group provides a reference for judging whether multi-spectral reflectance encoding offers measurable improvement over common RGB-based inspection.

2.3. Measurement Methods and Quality Control

All images were recorded with synchronized narrow-band LEDs and a calibrated multi-channel camera. Before capturing data, white-reference and dark-reference frames were collected at each wavelength to correct variations in channel intensity. During robot experiments, joint positions were recorded at 200 Hz, and deviations were computed with respect to a reference path. Three indicators were used: recognition accuracy, reflection-related error rate, and alignment deviation measured in millimeters. To maintain data quality, 5% of image sets were manually checked for channel misalignment, exposure issues, and motion blur. Any sequence showing sensor saturation or blur was removed.

2.4. Data Processing and Model Formulation

Each multi-spectral set is represented as $I=[I_{\lambda_1}, I_{\lambda_2}, \dots, I_{\lambda_L}]$. A reflectance-normalized value is computed as:

$$R_{\lambda_i} = \frac{I_{\lambda_i} - D_{\lambda_i}}{W_{\lambda_i} - D_{\lambda_i}},$$

where D_{λ_i} is the dark-reference value and W_{λ_i} is the white-reference value at wavelength λ_i . Recognition accuracy is calculated as:

$$Accuracy = \frac{N_{correct}}{N_{total}},$$

and the reflection-related error rate is:

$$Error_{reflect} = \frac{N_{fail,reflect}}{N_{total}}.$$

All channels were resized and normalized before entering the reflectance-stabilization module, which produces descriptors that remain consistent under lighting or viewpoint changes.

2.5. Computational Environment and Reproducibility

Experiments were conducted on a workstation with two 24-GB GPUs and 128 GB RAM. Calibration settings, wavelength choices, and exposure parameters were kept constant for all datasets. Each experiment was repeated with fixed random seeds. All preprocessing steps, normalization functions, and evaluation scripts were stored under version control to allow reproduction under identical conditions.

3. Results and Discussion

3.1. Performance on Metallic Vision Benchmarks

Across both metallic datasets, the multi-spectral reflectance encoder gives clear gains over the RGB baselines. On MetalBench-2025, average recognition accuracy increases by 21.4%, while reflection-related errors drop by 33.7%. These improvements appear on polished, brushed, and coated surfaces, which suggests that the method is not tied to a single type of finish. Figure 1 summarizes accuracy and error rates for all tested methods. The plot shows that the multi-spectral encoder maintains stable accuracy even when glare is strong, whereas RGB-only models show abrupt performance loss. On SteelLine-Vision, the model reduces missed detections around reflective edges and improves contour stability near weld joints. In many scenes with mixed direct and indirect reflections, the RGB baseline saturates at bright spots and loses seam boundaries, while the encoded multi-spectral signatures allow the classifier to follow the actual object outline more closely [15].

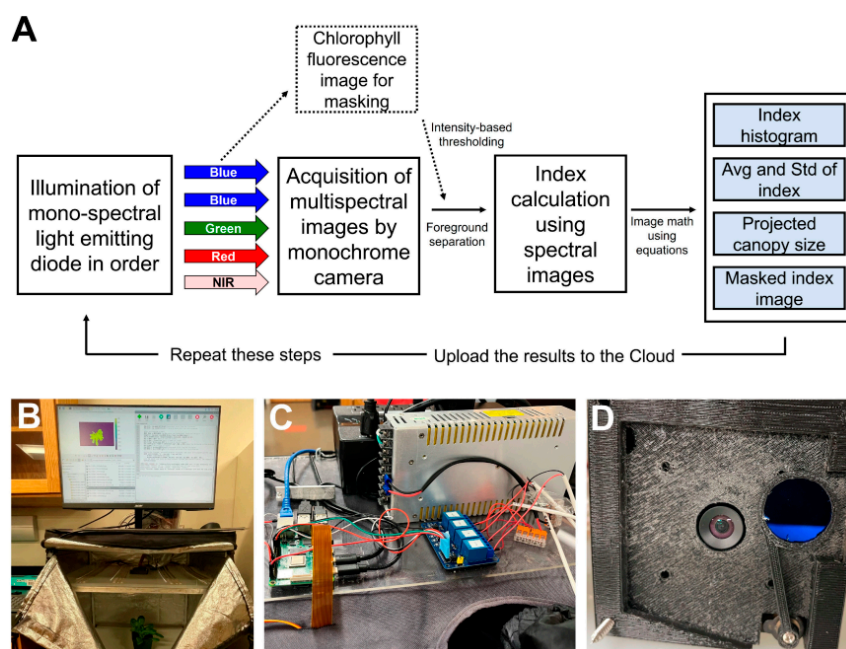


Figure 1. Accuracy and reflection-related error of all tested methods on the metallic datasets.

3.2. Effect of Spectral Encoding and Normalization

Ablation tests show that narrow-band illumination and spectral normalization each play an important role. Removing the narrow-band channels and using only RGB images causes accuracy to fall, especially on curved metal objects that generate strong highlights. When normalization is disabled, variations between lighting conditions increase, and class clusters in the feature space become less compact. In both cases, predictions near edges and weld roots become less stable. With the full encoder, feature distributions from different lighting conditions align more closely, making class boundaries easier to maintain. This pattern is consistent with observations in multispectral defect-inspection studies, where simple band selection and per-band correction improve feature separation compared with intensity-only inputs [16,17]. The present results extend these insights to metallic factory environments, where saturation and blooming are frequent and can distort the appearance of reflective surfaces.

3.3. Robotic Deployment and Alignment Behaviour

When deployed on a welding robot, the encoder leads to better tracking of joint positions. Alignment precision improves by 18.6% compared with the RGB-only setup, and the rate of missed

detections on reflective surfaces decreases by 27.3%. The improvement is most clear at weld start and end points. At these locations, standard edge detectors often follow bright reflections instead of the seam centre. The multi-spectral encoder reduces this problem, resulting in smoother tracking and fewer abrupt path deviations. Figure 2 shows an example of the sensing unit and how the spectral channels are used during robot motion. Unlike seam-tracking systems based mainly on a single laser stripe or RGB segmentation, the proposed method keeps the robot program unchanged and improves performance through more stable visual input [18]. This makes integration into existing production cells easier because no additional mechanical modules or complex calibration procedures are required.

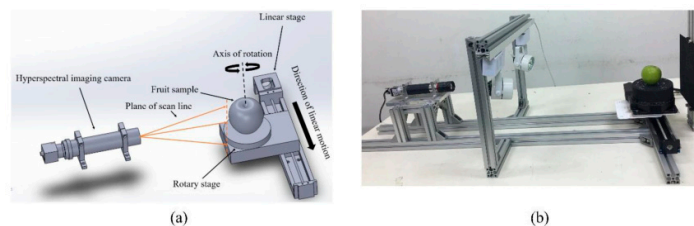


Figure 2. Multi-spectral imaging setup used during robotic joint tracking.

3.4. Comparison with Existing Work and Remaining Limitations

Compared with traditional weld-inspection pipelines that rely on single-band cameras and fixed thresholds to reduce glare, the proposed encoder offers a more direct way to handle dynamic reflections in metallic plants. Deep segmentation models such as improved YOLOv8s-seg have shown strong weld recognition when contrast is adequate [19]. Multispectral imaging studies in agriculture and environmental sensing also report better tolerance to lighting variations when several narrow bands are combined [20,21]. This study brings these observations to metallic manufacturing and shows that a small number of narrow bands, with simple normalization, can improve both recognition and robot-alignment performance. There are, however, limits. The experiments use only two datasets and one type of industrial robot. Tests on more alloys, coatings, and factory layouts are needed. The spectral bands are fixed and may not cover all material types or lighting conditions. More flexible band selection or adaptive exposure may help in factories with stronger illumination variation. Finally, the current system does not recover full 3D geometry, so depth cues remain implicit. Combining the reflectance encoder with established 3D sensing—for example, structured-light systems used in weld reconstruction—may further improve seam tracking under difficult lighting.

4. Conclusion

This study introduces a multi-spectral reflectance method for vision tasks in metallic manufacturing settings. By applying narrow-band lighting and simple spectral normalization, the system produces reflectance signals that remain more consistent under glare and fluctuating brightness. Tests on the MetalBench-2025 and SteelLine-Vision datasets show higher recognition accuracy and fewer reflection-related errors than RGB-based baselines. When used on a welding robot, the method also reduces alignment deviation and lowers detection failures on reflective surfaces. These results show that wavelength-selective reflectance cues can support more stable perception in metallic factories, where changing light often disrupts standard imaging. The work still has limits, including the use of fixed bands, two datasets, and one robot type. Future studies should examine adaptive selection of spectral bands, combine reflectance cues with depth sensing, and evaluate performance in different factory layouts and with a broader range of metal surfaces.

References

1. Yan, R., Dang, D., Peng, K., Li, Y., Tao, Y., Hou, L., ... & Tang, J. (2025). Document-level Relation Extraction with Low Entity Redundancy Feature Map. *IEEE Transactions on Knowledge and Data Engineering*.
2. Sinoara, R. A., Antunes, J., & Rezende, S. O. (2017). Text mining and semantics: a systematic mapping study. *Journal of the Brazilian Computer Society*, 23(1), 9.
3. Genest, P. Y. (2024). Unsupervised open-world information extraction from unstructured and domain-specific document collections (Doctoral dissertation, INSA de Lyon).
4. Alvarez, J. E., & Bast, H. (2017). A review of word embedding and document similarity algorithms applied to academic text. Bachelor thesis, 1.
5. Murty, S., Verga, P., Vilnis, L., Radovanovic, I., & McCallum, A. (2018, July). Hierarchical losses and new resources for fine-grained entity typing and linking. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 97-109).
6. Jin, J., Su, Y., & Zhu, X. (2025). SmartMLOps Studio: Design of an LLM-Integrated IDE with Automated MLOps Pipelines for Model Development and Monitoring. *arXiv preprint arXiv:2511.01850*.
7. Zini, J. E., & Awad, M. (2022). On the explainability of natural language processing deep models. *ACM Computing Surveys*, 55(5), 1-31.
8. Toldo, M., Maracani, A., Michieli, U., & Zanuttigh, P. (2020). Unsupervised domain adaptation in semantic segmentation: a review. *Technologies*, 8(2), 35.
9. Wu, S., Cao, J., Su, X., & Tian, Q. (2025, March). Zero-Shot Knowledge Extraction with Hierarchical Attention and an Entity-Relationship Transformer. In *2025 5th International Conference on Sensors and Information Technology* (pp. 356-360). IEEE.
10. Chai, Y., Zhang, H., Yin, Q., & Zhang, J. (2023, June). Neural text classification by jointly learning to cluster and align. In *2023 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.
11. Liang, R., Ye, Z., Liang, Y., & Li, S. (2025). Deep Learning-Based Player Behavior Modeling and Game Interaction System Optimization Research.
12. Yin, Z., Chen, X., & Zhang, X. (2025). AI-Integrated Decision Support System for Real-Time Market Growth Forecasting and Multi-Source Content Diffusion Analytics. *arXiv preprint arXiv:2511.09962*.
13. Lopes Junior, A. G. (2025). How to classify domain entities into top-level ontology concepts using language models: a study across multiple labels, resources, domains, and languages.
14. Wu, C., Zhang, F., Chen, H., & Zhu, J. (2025). Design and optimization of low power persistent logging system based on embedded Linux.
15. Yuan, M., Qin, W., Huang, J., & Han, Z. (2025). A Robotic Digital Construction Workflow for Puzzle-Assembled Freeform Architectural Components Using Castable Sustainable Materials. Available at SSRN 5452174.
16. Grewal, D., & Compeau, L. D. (2017). Consumer responses to price and its contextual information cues: A synthesis of past research, a conceptual framework, and avenues for further research. In *Review of marketing research* (pp. 109-131). Routledge.
17. Chen, F., Liang, H., Yue, L., Xu, P., & Li, S. (2025). Low-Power Acceleration Architecture Design of Domestic Smart Chips for AI Loads.
18. Wu, C., & Chen, H. (2025). Research on system service convergence architecture for AR/VR system.
19. Tashakori, E., Sobhanifard, Y., Aazami, A., & Khanizad, R. (2025). Uncovering Semantic Patterns in Sustainability Research: A Systematic NLP Review. *Sustainable Development*.

20. Tan, L., Liu, D., Liu, X., Wu, W., & Jiang, H. (2025). Efficient Grey Wolf Optimization: A High-Performance Optimizer with Reduced Memory Usage and Accelerated Convergence.
21. Xu, K., Wu, Q., Lu, Y., Zheng, Y., Li, W., Tang, X., ... & Sun, X. (2025, April). Meatr: Multimodal anomalous tissue region detection enhanced with spatial transcriptomics. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 39, No. 12, pp. 12918-12926).

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.