Proceedings of CONF-SEML 2026 Symposium: Computational Analysis and Modeling in Complex Intelligent Systems
DOI: 10.54254/2755-2721/2026.GU33657

Visual Perception and Weld Seam Tracking in Intelligent Robotic Welding

Hangrui Liu

Software College, Northeastern University, Shenyang, China
Liuhenry946@gmail.com

Abstract. Non-standard and high-mix manufacturing are increasingly using intelligent robotic welding, but seam tracking is challenging due to high reflectivity, arc glare, spatter and distortion. This review presents the latest advances in visual perception and weld seam tracking for intelligent robotic welding. It starts with multisource sensing and image preprocessing techniques to enhance 2D seam feature detection. It then considers deep learning-based feature detection and structured light 3D reconstruction as intermediate steps from image recognition to 3D weld seam modeling. It also reviews path generation, hand-eye transformation and closed-loop error correction to illustrate the transformation of perceived seam information into robot motions. Finally, multimodal quality monitoring and digital twin-based welding systems are reviewed as trends for online quality evaluation and process improvement and suggests that robust visual perception algorithms alone are insufficient for future intelligent welding systems; they also need lightweight 3D semantic models, safety-constrained motion planning, and plug-and-play data interfaces for commercial applications.

Keywords: Intelligent robotic welding, Visual perception, Weld seam tracking, Deep learning, Digital twins

1. Introduction

Smart manufacturing has increased the application of industrial robots in non-standardized working environments, such as shipbuilding, aerospace, and heavy steel fabrication. Recent industry reports forecast the global robotic welding industry will grow to $27.9 billion with a compound annual growth rate of 15.20% by 2034 [1]. In this context, ensuring welding robots can sense joint geometry, adjust welding trajectories, and handle welding process disturbances has emerged as a critical challenge in industrial automation [2].

Despite the rapid advancement of industrial automation, many welding robots still rely on teach-playback operation. This method is very effective for mass production, but less effective for non-standard products because misalignment, deformation and trajectory drift often occur during welding [2-4]. Furthermore, reflective materials such as stainless steel can introduce local saturation of images and arc glare, spatter, and electromagnetic interference degrade the extraction of image features [4, 5]. This makes it hard to accurately locate the weld seam in industrial environments [3, 4].

© 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: 10.54254/2755-2721/2026.GU33657

To overcome these issues, this paper reviews the recent research on multi-source vision, image enhancement, structured light reconstruction, and adaptive control for tracking weld seams. Section 2 presents sensor devices and multi-source perception. Section 3 presents interference-robust image pre-processing and deep learning-based feature extraction. Section 4 introduces structured light-based 3D reconstruction as a bridge from 2D perception to 3D modeling. Section 5 reviews trajectory planning and adaptive closed-loop control. Section 6 also covers multimodal monitoring and quality evaluation with digital twins.

2. The state of sensing systems in the intelligent robotic welding

2.1. Development trends in perception hardware

To break the restraints of teach-playback welding, recent research tends to integrate 3D vision systems with algorithm-based decision-making systems. According to Chen et al., this translates to a transition from virtual collaborative robotic welding to predict-optimize-control [2]. Thus, multi-source-sensing-based flexible and precise perception is a focal point of hardware development in intelligent welding.

2.2. Classification and principles of multi-source sensing

Sensing systems acquire diverse information about the welding environment, seam geometry and welding process. The information is used to support seam localization, trajectory generation, and closed-loop correction. Sensing systems in intelligent robotic welding can be classified into passive vision sensing, active vision sensing and process-signal sensing based on the method of information acquisition.

Passive vision sensing mainly uses general cameras or industrial cameras to acquire images of the welding zone. Weld seam information is then detected via edge detection, semantic segmentation, or keypoint detection. This approach is easy to implement with a simple and low-cost hardware system, but its performance is sensitive to low contrast, high reflection and severe arc interference.

Active vision sensing augments weld geometry by projecting additional patterns. Two examples include laser vision and structured-light vision. Laser vision typically projects a laser stripe on the workpiece surface and records the deformation of the laser stripe along the groove or seam. The stripe is then used to extract the seam centerline, groove width or height. In recent applications of online seam tracking, laser vision has been combined with industrial robots to calculate seam centerlines and modify robot trajectories [3]. Structured-light vision is considered a more advanced version of active optical sensing and its process of 3D reconstruction is described in Section 4. Process-signal sensors consist of arc voltage, welding current, acoustic emission, and thermal imaging. These do not yield full seam geometry information directly, but they are indicative of process stability, heat-input variation and defect creation. Visual features are often enhanced by using narrow band filters to reduce arc interference and enhance signal-to-noise ratios. For highly reflective stainless-steel welding, active laser vision can be more stable than passive vision when combined with arc signals [2-4].

Therefore, the key to multi-source sensing is not to simply add more sensors, but to use complementary information. Imaging sensors provide geometric information, laser or structured-light sensors improve morphological features, and arc or thermal sensors represent process states. By synchronizing, calibrating and fusing these different types of data, the system can get more robust weld seam recognition under spatter, arc glare and local overexposure. But multi-source

DOI: 10.54254/2755-2721/2026.GU33657

acquisition does not completely eliminate image disturbances. The visual inputs of welding may still suffer from arc glare, spatter, local overexposure, and reflection. These disturbances need to be removed to ensure robust seam feature extraction. Thus, Section 3 introduces image enhancement and anti-interference pre-processing techniques to transform visual data into valid weld seam features.

3. Imaging technology and enhancement in noisy environments

3.1. Image preprocessing with anti-interference

In image enhancement, image enhancement is moving toward feature segregation and frequency domain adaptive enhancement.

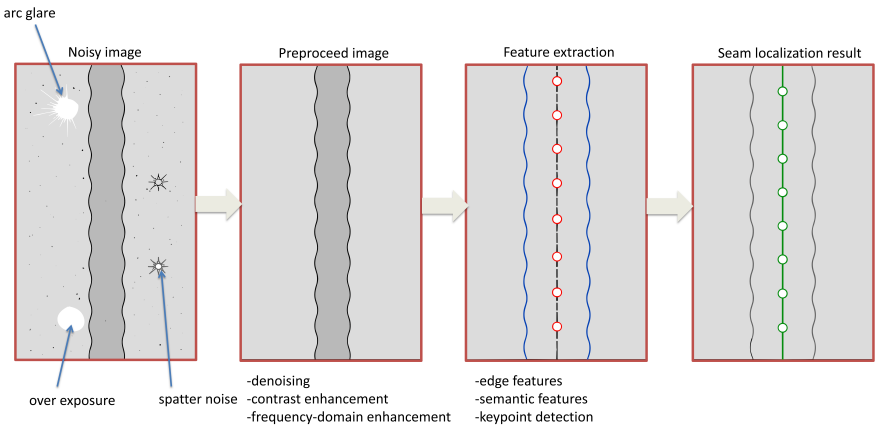


Figure 1. Schematic pipeline of weld seam image preprocessing and feature extraction

As illustrated in Figure 1, the typical weld seam image processing workflow goes from noisy visual perception to feature enhancement to seam extraction. In the acquisition stage, reflection, welding spatter and arc glare make it difficult to obtain clear and stable seam contours. Song and Wang introduced the Y-shaped Fast Fourier Convolution Network (Y-FFC Net), modelling in the frequency domain for better reconstruction of highly reflective materials [5]. While this approach is not welding-specific, it offers a technical starting point for addressing issues of reflection-caused visual degradation in weld seam recognition.

3.2. Deep learning weld seam feature extraction

In addition to image enhancement, weld seam detection has transitioned from hard-coded features to learned semantics. Traditional edge and threshold-based approaches heavily rely on expert knowledge, while deep learning algorithms can automatically learn seam-related features from pixels. Kang et al. introduced a weld image feature point detection algorithm with an improved Deeplabv3+ semantic segmentation network, demonstrating the potential of semantic extraction in

DOI: 10.54254/2755-2721/2026.GU33657

identifying weld features [6]. Mobaraki et al. also proposed a model for keypoint detection of GMAW fillet weld seam tracking, which shows the suitability of deep learning to detect sensitive seam features [7].

However, deep learning-based visual feature recognition should be considered as a baseline approach. Future trends include vision transformers, explainable AI and edge computing [6, 7]. These usually work well with carefully prepared samples, but their reliance on pixels can make generalization in welding challenging [5-7]. Thus, for variable workpieces and spatial trajectory correction, weld seam recognition should be extended from 2D planar recognition to 3D morphological recognition, as presented in the structured light reconstruction in Section 4.

4. Structured light 3D welding guidance

4.1. 3D point cloud acquisition and spatial modeling

Structured-light 3D reconstruction provides spatial geometry of non-standard workpieces and overcomes the depth deficiency of 2D image analysis. In a typical welding guidance system, structured fringe patterns are projected onto the surface of a workpiece, and the fringe patterns are distorted due to the height variation. With phase decoding, calibration, and coordinate transform, the system translates the image information into dense 3D point clouds for extracting weld geometry. Once the point cloud is obtained, spatial modeling is carried out by fitting the groove profile of the weld, extracting geometric descriptors including the seam centerline, groove width, root opening and bevel angle, and converting these descriptors into the robot coordinate system. In this way, structured-light 3D reconstruction not only provides visual depth data, but also transforms the weld structure into geometric descriptors for trajectory planning and correction. This reconstruction process is illustrated in Figure 2.

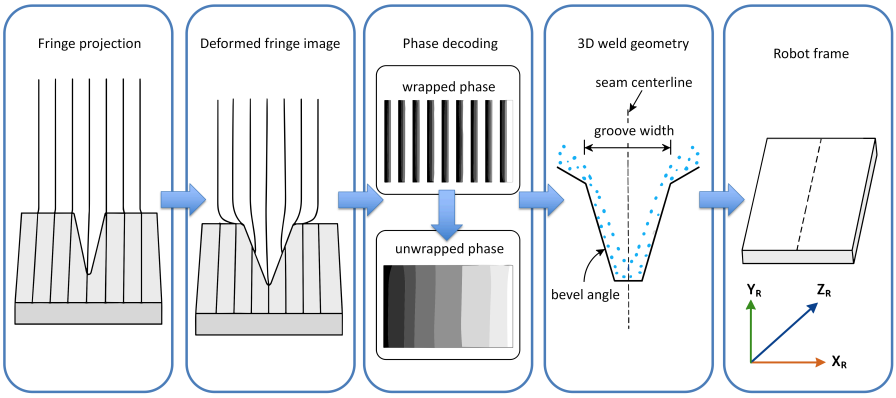


Figure 2. Workflow of structured-light 3D reconstruction for weld seam spatial modeling

DOI: 10.54254/2755-2721/2026.GU33657

4.2. Trade-offs between real-time performance and high dynamic range

Highly reflective surfaces create trade-offs among measurement accuracy, dynamic range, and real-time performance. Specular reflections and multi-path interference can violate the sinusoidal intensity model of conventional fringe profilometry, causing phase errors or depth loss. Micro-electro-mechanical system (MEMS) projection may be a hardware solution because micro-mirror devices can facilitate compact and fast projection. At the software level, deep learning-based phase extraction can reduce the need for multiple temporal projections, which are prone to motion blur and time delay [8, 9].

However, fast projection and phase recovery do not directly produce robot control commands. Even when seam-related geometric features are extracted from the point cloud, they need to be further filtered, simplified and structured into control inputs. Hence, future 3D welding guidance must include efficient point-cloud processing and task-oriented geometric modeling, not just hardware improvement of the projector. These geometric features must then be converted into smooth and coordinated welding paths, which leads to the path planning and adaptive control techniques discussed in Section 5.

5. Path planning technologies and adaptive correction control

5.1. Visual feedback and closed-loop control

Using spatial features of the weld reconstructed from 3D vision, the first step of dynamic tracking is to convert visual features into the robot motion space. This involves precise hand-eye calibration, using which 3D points in the camera coordinate system are converted into the robot base frame through homogeneous transformation matrices. Once the mapping is done, the position of the detected seam can be compared with the welding trajectory, and the deviation can be used to update the torch position.

In a closed-loop welding system, visual perception does not implement autonomous welding. It is instead used to provide measurement feedback for trajectory correction. The seam centerline, height offset, or groove profile need to be transformed into control offsets by coordinate conversion, path filtering, and motion compensation. Research on online seam tracking demonstrates that vision feedback, such as laser-vision sensing, can be used in conjunction with industrial robots to correct the weld trajectory online [3]. Thus, the control problem is to ensure a constant torch-workpiece relationship under assembly errors, thermal distortion and sensor delay.

The closed-loop feedback also defines the role of perception and control. The image recognizer and 3D reconstruction deliver spatial measurements and the robot controller translates these to motion commands. If they are not coordinated, accurate seam detection does not necessarily lead to stable welding. As such, seam tracking systems require coordinated integration of visual sensing, coordinate transformation, trajectory compensation and robot control.

5.2. Teach-free operations and global planning

Teach-free welding is concerned with generating weld paths directly from sensor data rather than from human teaching. Here, global planning refers to the initial planning of the weld path and welding sequence from CAD models or reconstructed point clouds while considering robot accessibility, torch orientation and collision avoidance. This level mainly focuses on deciding where to weld and the sequence to complete welding tasks.

DOI: 10.54254/2755-2721/2026.GU33657

Spline-based path planning can be used for this task as it can model complex weld paths with smooth splines, and reduce robot motion discontinuities [10]. For real-world robotic welding, real-time registration of reconstructed point clouds and CAD models can also be used to facilitate welding sequence planning and collision-free path generation. The system can detect potential collisions between the welding torch and complex parts in a simulation to minimize collision risk.

But a path is only a guide. In real welding, assembly error, welding distortion, and sensing lag may result in a shift in the seam position from the planned path. Therefore, geometric planning should not be an isolated module, but integrated with online correction. In practice, a hybrid approach is required that combines geometric planning for safe and smooth motion, and visual feedback and adaptive correction for error compensation.

While this combination may enhance geometric seam tracking, it still primarily controls the torch position relative to the seam. It cannot comprehensively assess process stability and weld quality, including melt-pool evolution, heat-input variation, and defect formation. This prompts the transition from geometric seam tracking to multimodal monitoring and digital twin-based quality assessment in Section 6.

6. Multimodal fusion and digital twins for future factory architectures

6.1. Multimodal data monitoring and online quality assessment

The above perception and tracking data provide a geometric basis for intelligent welding, but weld quality evaluation also needs process information. Multimodal monitoring in practice can be used to combine visual images, 3D point clouds, welding current and arc voltage, acoustic emissions and melt-pool thermal images. The signals above correspond to all sorts of issues in the welding process, such as seam shape, heat-input fluctuations, process instability, melt-pool characteristics, and defect risks. Optical point clouds show the structure and position of seams, and electrical, acoustic and thermal signals are emitted from reflecting hidden process states. The above data will be used to calibrate and combine them in the system to determine not only the position of the seam but also the quality of weld formation. Recent research on deep learning-based weld defect detection has achieved various results in classification, detection and segmentation, but its application in industry is still limited by the quality of data, real-time performance and generalization ability of models [11]. Therefore, many ways can be used to drive the ongoing improvement of quality control in the field of offline inspection and online prediction-correction.

6.2. Digital twins and cloud collaboration

Digital twins can be used to link welding operations with virtual models and data. For welding, the digital twin can align robot kinematics, vision, welding process parameters, and quality metrics for monitoring, simulation, and process parameter tuning [12, 13]. The cloud-edge paradigm can support this approach by distributing real-time sensing and control to the edge, while enabling cloud platforms to update models, store and analyse historical data, and optimize welding processes. But digital twins are only as good as their data synchronization and model-calibration capabilities, not their visual representation.

To implement this, the digital-twin functions can be divided into a data flow that connects sensors, robot controllers, quality databases, and virtual models. Data from sensors and controllers is first gathered at the edge, then stored, processed and analyzed in cloud or local servers, and then fed back as parameter setting updates or quality alarms. In this capacity, the digital twin serves as an

DOI: 10.54254/2755-2721/2026.GU33657

information coordination layer of a monitoring and process optimization system, rather than an independent visualization model. However, the industrial implementation of digital-twin-based welding systems remains hampered by interoperability, rather than just computing power. Multimodal sensors, robot controllers, welding power sources and quality databases can use different data formats and communication protocols, making it challenging to integrate process information across systems. Without uniform interfaces and verified mechanisms for updating models, digital twins may only serve as stand-alone monitoring systems instead of plug-and-play industrial decision-support systems. Thus, future smart welding factories require integrated data streams to link multimodal perception, robot control, quality monitoring and virtual simulation into a reliable closed-loop system.

7. Conclusion

In summary, intelligent welding robots are emerging as key elements in high-mix low-volume manufacturing. This review has outlined the technologies from perception to closed-loop weld seam tracking. At the level of perception, multi-source sensing and image enhancement enhance the stability of image-based seam feature extraction in complex welding environments. At the modeling level, structured-light 3D reconstruction provides a feasible way to convert image-level features into spatial geometric representations. At the execution level, visual servo control, spline-based trajectory generation and adaptive control convert seam geometry perception into robot motion. At the monitoring level, multimodal fusion and digital twin technologies bring weld seam tracking to real-time quality monitoring and process optimization.

However, the technologies reviewed here still suffer from a number of challenges to industrial applications. Visual perception at the image level is still susceptible to arc flare, spatter, smoke, reflection, and local overexposure, particularly in unstructured environments. Structured-light reconstruction needs to compromise between measurement accuracy and real-time point-cloud processing and control cycles. Hence, low-complexity and task-specific models must be developed to identify control-relevant features such as seam centerlines, groove boundaries and height errors. Furthermore, path planning and automatic correction should meet industrial safety standards. Online planning should be integrated with geometric constraints, calibration knowledge, and deterministic safety boundaries, calibration data, and deterministic safe zones for industrial applications.

Lastly, robotic welding is advancing towards digital and networked manufacturing, where the challenge is moving beyond individual algorithmic performance to system interoperability. Digital-twin systems for welding rely on seamless data exchange, model calibration and unified interfaces between sensors, robot controllers, quality databases and simulation models. Without standardized data formats and communication protocols, multimodal data may be isolated in different manufacturing platforms. Thus, future intelligent welding systems should consider anti-interference perception, lightweight 3D reconstruction, safe trajectory correction, and interoperable digital-twin monitoring to implement a verifiable closed-loop architecture for industrial use.

References

[1] Fortune Business Insights, Robotic welding market size, share & growth report 2024-2034, 2024.
[2] H. Chen et al., Progress and trends towards visual perception technologies in intelligent robotic welding, Transactions of the China Welding Institution, vol. 45, no. 11, pp. 1-9, 2024.
[3] Q.-C. Nguyen, H. Q. B. Hua, and P.-T. Pham, Development of a vision system integrated with industrial robots for online weld seam tracking, Journal of Manufacturing Processes, vol. 119, pp. 414-424, 2024.

DOI: 10.54254/2755-2721/2026.GU33657

[4] J. Wang, L. Li, and P. Xu, Visual sensing and depth perception for welding robots and their industrial applications, Sensors, vol. 23, no. 24, 9700, 2023.
[5] X. Song and L. Wang, Y-FFC Net for 3D reconstruction of highly reflective surfaces, IEEE Transactions on Industrial Informatics, vol. 20, no. 12, pp. 13966-13974, 2024.
[6] S. Kang, H. Qiang, J. Yang, K. Liu, W. Qian, W. Li, and Y. Pan, Research on a feature point detection algorithm for weld images based on deep learning, Electronics, vol. 13, no. 20, 4117, 2024.
[7] M. Mobaraki, S. Ahani, R. Gonzalez, K. M. Yi, K. Van Heusden, and G. A. Dumont, Vision-based seam tracking for GMAW fillet welding based on keypoint detection deep learning model, Journal of Manufacturing Processes, vol. 117, pp. 315-328, 2024.
[8] Z. Zhang et al., Fringe-based structured-light 3D reconstruction: Principles, projection technologies, and deep learning integration, Sensors, vol. 25, no. 20, 6296, 2025.
[9] M. Han, X. Wang, and X. Li, Fast and accurate fringe projection based on a MEMS micro-vibration mirror, in Optical Metrology and Inspection for Industrial Applications XI, SPIE, vol. 13241, 2024.
[10] G. He, H. H. Teo, and L. K. Moey, Review of path planning of welding robot based on spline curve, Journal of Mechanical Engineering and Sciences, vol. 18, no. 1, pp. 9928-9948, 2024.
[11] B. Eren, Deep learning approaches for weld defect detection: A comprehensive review of models, applications, and future directions, Computers & Industrial Engineering, vol. 212, 111725, 2026.
[12] Y. Kang and R. Chen, Welding robot automation technology based on digital twin, Frontiers in Mechanical Engineering, vol. 10, 1367690, 2024.
[13] S. Wang et al., Research on the digital twin system of welding robots, Sensors, vol. 25, no. 13, 3889, 2025.