Proceedings of the ASME 2026 21st International Manufacturing Science and Engineering Conference
MSEC2026
June 14–18, 2026, State College, Pennsylvania

MSEC2026-182780

UNCERTAINTY-AWARE MULTI-VIEW STEREO FOR GEOMETRY RECONSTRUCTION IN MANUFACTURING

Zihan Chen[1,†], Guanzhong Hu[1,†], Ping Guo[1,*]

[1]Department of Mechanical Engineering, Northwestern University, Evanston, IL, 60208

ABSTRACT

Multi-View Stereo (MVS) provides a flexible and low-cost solution for full-field three-dimensional (3D) geometry measurement in engineering and manufacturing applications. Compared with laser scanning or fringe projection profilometry, MVS offers greater scalability and measurement efficiency through direct imaging setups. As manufacturing system evolves toward intelligent and data-driven paradigms, the ability to quantify measurement uncertainty become essential for reliable process monitoring and control. However, despite extensive research on MVS reconstruction algorithms, rigorous uncertainty modeling and evalution remain largely underexplored, limiting the credibility of MVS as a metrological tool. This study presents an uncertainty-aware framework for estimating the geometric uncertainty of 3D points reconstructed by MVS. The approach is based on the law of uncertainty propagation, incorporating both camera calibration uncertainty and image matching errors. A multi-camera imaging system is deployed to monitor the English Wheel metal forming process, which serves as a representative case involving non-Lambertian and low-texture surfaces. Camera parameter covariances obtained from experimental calibration and empirically characterized matching uncertainties are propagated through the multi-view triangulation relationships to estimate the uncertainty associated with each reconstructed 3D point. Building on this uncertainty estimation, an uncertainty-guided view selection strategy is proposed to identify stable camera combinations for reconstructing different surface regions. By excluding unreliable observations and favoring views with lower expected uncertainty, the method improves both computational efficiency and reconstruction reliability. Experimental results demonstrate that the proposed approach enhances reconstruction accuracy and robustness compared with conventional MVS methods that utilize all available views indiscriminately. The framework therefore enables more reliable real-time deformation monitoring and provides a foundation for uncertainty-aware 3D metrology in advanced manufacturing environments.

[†]Joint first authors
*Corresponding author: ping.guo@northwestern.edu

Keywords: Multi-View Stereo, Uncertainty Propagation, Metal Forming

1. INTRODUCTION

The transition toward digital and data-driven production in advanced manufacturing has elevated in-situ and real-time three-dimensional (3D) measurement from a desirable capability to a fundamental requirement for modern process monitoring and control [1]. Within the Industry 4.0 paradigm, quality assurance is increasingly shifting from traditional offline inspection toward continuous, in-process monitoring integrated directly into the manufacturing workflow [2]. Such capabilities are essential for constructing accurate digital twins and enabling closed-loop corrective actions during production [3]. In-situ metrology provides this critical geometric feedback during fabrication, allowing deviations in processes such as additive manufacturing, machining, or sheet metal forming to be detected and corrected before defects propagate [4, 5]. As manufacturing systems adopt digital threads and cyber-physical production frameworks, reliable geometric sensing becomes a key enabler for high-confidence process monitoring and quality assurance. However, deploying metrology systems that are simultaneously accurate, dense, and robust within the complex and dynamic environment of a live production line remains a significant challenge.

Optical measurement technologies are widely employed to monitor the geometric evolution of large metal components during forming processes. Nevertheless, each category exhibits inherent trade-offs between accuracy, measurement speed, and spatial coverage. For instance, while laser line profilers can capture high-speed cross-sections measurements, but obtaining full-field 3D geometry requires mechanical scanning, which increase acquisition time and introduces vibration-induced noise [6]. Structured-light scanners are capable of producing high-accuracy and dense reconstructions [7], yet their sequential projection schemes require multiple exposures, which limits their applicability in monitoring rapidly evolving surfaces undergoing dynamic deformation [8]. Time-of-Flight cameras provide real-time depth sensing and convenient system integration [9], but their

Copyright ©2026 by ASME

millimeter-level depth accuracy and relatively low spatial resolution are often insufficient for the sub-millimeter deformation monitoring required in precision forming applications. Beyond these individual limitations, these methods share a fundamental, system-level vulnerability to occlusion. They are predominantly constrained by single-view imaging configurations, and in complex forming environments where tooling and the part's own geometry create significant and rapidly changing occlusions, this results in massive data loss. In addition, many of these approaches rely on single-view sensing configurations and are therefore highly susceptible to occlusion caused by tooling, machine components, or complex part geometries. Although multiple sensors can be deployed to mitigate occlusion, the commonly adopted "reconstruction-then-fusion" paradigm, in which individual point clouds are reconstructed separately and later registered together, inevitably accumulates registration errors, thereby degrading the global geometric consistency of the reconstructed model.

Active Multi-View Stereo (MVS) provides a promising alternative that balances measurement speed, geometric fidelity, and spatial completeness. By deploying multiple cameras around the target object, MVS systems capture the surface from different viewpoints, allowing regions occluded in one view to be observed in others. When combined with projected texture patterns, these systems can acquire dense 3D data at video rates while maintaining sub-millimeter precision, making them well suited for monitoring high-rate forming processes involving complex geometries [10]. Compared with scanning-based methods, multi-view imaging avoids mechanical motion and enables full-field measurements in dynamic environments, which is particularly advantageous for in-situ monitoring scenarios.

Despite these advantages, applying MVS in manufacturing environments presents several practical challenges. In sheet metal forming processes, the workpiece surface often exhibits low natural texture, strong specular reflections, and rapidly evolving geometry due to plastic deformation. In addition, tooling and machine components frequently introduce dynamic occlusions that complicate reliable image correspondence across views. These factors lead to spatially varying reconstruction reliability across the surface, where some regions can be reconstructed accurately while others may contain significant errors due to ambiguous matching or unfavorable viewing geometry. For applications such as deformation tracking, process monitoring, or digital twin updating, simply generating dense point clouds is therefore insufficient. It is equally important to quantify the reliability of reconstructed geometry so that measurements with large uncertainty can be identified and excluded from downstream analysis.

Uncertainty quantification therefore plays a critical role in enabling MVS to function as a metrological tool in manufacturing. Previous studies have investigated uncertainty estimation in stereo vision systems. For example, Chen *et al.* [11] and Leo *et al.* [12], has established theoretical frameworks for uncertainty estimation in stereo vision. Concurrently, experimental approaches, like those from Sankowski *et al*. [13], have focused on evaluating stereo system uncertainty using objects of known geometry. However, most existing studies focus on conventional stereo configurations and have not been widely extended to multi-view reconstruction systems operating under complex industrial conditions. As a result, uncertainty-aware reconstruction methods suitable for manufacturing applications remain relatively underdeveloped. This limitation poses challenges for digital-twin frameworks and closed-loop process monitoring, where reliable geometric information and its associated confidence are essential for trustworthy decision-making.

To address this gap, this paper presents an uncertainty-aware multi-view stereo framework for in-situ geometry reconstruction in manufacturing environments. The proposed approach analytically propagates uncertainties originating from camera calibration and image correspondence into reconstructed 3D geometry and utilizes the resulting uncertainty estimates to guide the reconstruction process. By incorporating uncertainty information into view selection and point filtering, the framework improves reconstruction robustness under challenging surface conditions commonly encountered in forming applications. The main contributions of this work are summarized as follows:

1. **A systematic uncertainty evaluation framework for multi-view stereo reconstruction:** an analytical formulation is developed to quantify the uncertainty of reconstructed 3D points by propagating errors originating from camera calibration and image correspondence. This framework provides physically interpretable covariance estimates for reconstructed geometry and establishes a foundation for uncertainty-aware metrology in MVS systems.

2. **An uncertainty-guided reconstruction strategy for improving measurement reliability:** the estimated uncertainty is incorporated directly into the reconstruction pipeline to guide the selection of reliable camera views prior to triangulation. By excluding unstable view combinations and filtering reconstructed points with large predicted uncertainty, the proposed method improves robustness under challenging imaging conditions such as specular reflections and weak texture.

3. **Experimental validation in a forming measurement scenario:** the proposed framework is experimentally validated on industrial metal forming components. The results demonstrate that uncertainty-guided reconstruction significantly improves measurement stability and produces accuracy comparable to that of a high-precision commercial laser scanner, highlighting the potential of uncertainty-aware MVS for reliable in-process monitoring in advanced manufacturing.

Overall, this work demonstrates that uncertainty-aware multi-view stereo reconstruction provides a credible and traceable measurement approach for in-situ geometric monitoring, supporting reliable data acquisition for digital-twin-enabled manufacturing and closed-loop process control.

2. THEORY

2.1. Camera Model

The pin-hole model is widely used to describe the imaging process of a camera [14]. Let $\tilde{\mathbf{P}}^{\mathrm{w}} = [x^{\mathrm{w}}, y^{\mathrm{w}}, z^{\mathrm{w}}, 1]^{\top}$ represent the homogeneous coordinates of a 3D point in the world coordinate

Copyright ©2026 by ASME

system. Its corresponding coordinates in the camera coordinate system, $\tilde{\mathbf{P}}^{\mathrm{c}} = [x^{\mathrm{c}}, y^{\mathrm{c}}, z^{\mathrm{c}}, 1]^\top$, are obtained via a homogeneous transformation:

$$\tilde{\mathbf{P}}^{\mathrm{c}} = \mathbf{T}_{\mathrm{w}}^{\mathrm{c}} \tilde{\mathbf{P}}^{\mathrm{w}}, \tag{1}$$

where $\mathbf{T}_{\mathrm{w}}^{\mathrm{c}} \in \mathrm{SE}(3)$ is the homogeneous transformation matrix (extrinsic matrix) and can be written as:

$$\mathbf{T}_{\mathrm{w}}^{\mathrm{c}} = \begin{bmatrix} \mathbf{R}_{\mathrm{w}}^{\mathrm{c}} & \mathbf{t}_{\mathrm{w}}^{\mathrm{c}} \\ \mathbf{0} & 1 \end{bmatrix}. \tag{2}$$

Here, $\mathbf{R}_{\mathrm{w}}^{\mathrm{c}} \in \mathrm{SO}(3)$ is the rotation matrix which can be parameterized by three Euler angles $\boldsymbol{\phi}_{\mathrm{w}}^{\mathrm{c}} = [\varphi_{\mathrm{w}}^{\mathrm{c}}, \theta_{\mathrm{w}}^{\mathrm{c}}, \psi_{\mathrm{w}}^{\mathrm{c}}]^\top$ [15], and $\mathbf{t}_{\mathrm{w}}^{\mathrm{c}} = [t_{\mathrm{w,x}}^{\mathrm{c}}, t_{\mathrm{w,y}}^{\mathrm{c}}, t_{\mathrm{w,z}}^{\mathrm{c}}]^\top$ is the translation vector. $\mathbf{T}_{\mathrm{w}}^{\mathrm{c}}$ can be represented by a vector $\boldsymbol{\eta}_{\mathrm{w}}^{\mathrm{c}} = [(\boldsymbol{\phi}_{\mathrm{w}}^{\mathrm{c}})^\top, (\mathbf{t}_{\mathrm{w}}^{\mathrm{c}})^\top]^\top$.

The projection of $\tilde{\mathbf{P}}^{c}$ onto the image plane, denoted by homogeneous coordinates $\tilde{\mathbf{p}}^{\mathrm{c}} = [u^{\mathrm{c}}, v^{\mathrm{c}}, 1]^\top$, is defined as:

$$\tilde{\mathbf{p}}^{\mathrm{c}} = \frac{1}{z^{\mathrm{c}}} \mathbf{K} \tilde{\mathbf{P}}^{c}, \tag{3}$$

where $\mathbf{K}$ is the camera intrinsic matrix:

$$\mathbf{K} = \begin{bmatrix} f_x & s & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{4}$$

Here, f_x and f_y represent the focal lengths along the x- and y-axes, s is the skew coefficient, and $[u_0, v_0]^\top$ is the principal point. $\mathbf{K}$ can be vectorized by $\boldsymbol{\kappa} = [f_x, f_y, s, u_0, v_0]^\top$.

Real-world lenses introduce optical aberrations that deviate from the ideal pinhole geometry. These are modeled using radial and tangential distortion terms:

$$\tilde{\mathbf{p}}^{\mathrm{c}} = \mathbf{K} \left[\frac{1}{z^{\mathrm{c}}} \tilde{\mathbf{P}}^{\mathrm{c}} + \boldsymbol{\delta}_{\mathrm{r}} + \boldsymbol{\delta}_{\mathrm{t}} \right], \tag{5}$$

where the radial distortion $\boldsymbol{\delta}$r and tangential distortion $\boldsymbol{\delta}$t are given by [16, 17]:

$$\boldsymbol{\delta}_{\mathrm{r}} = \begin{bmatrix} x_{\mathrm{n}}^{\mathrm{c}} \\ y_{\mathrm{n}}^{\mathrm{c}} \\ 0 \\ 0 \end{bmatrix} [k_1 (r_{\mathrm{n}}^{\mathrm{c}})^2 + k_2 (r_{\mathrm{n}}^{\mathrm{c}})^4 + k_3 (r_{\mathrm{n}}^{\mathrm{c}})^6], \tag{6}$$

$$\boldsymbol{\delta}_{\mathrm{t}} = \begin{bmatrix} 2k_4 x_{\mathrm{n}}^{\mathrm{c}} y_{\mathrm{n}}^{\mathrm{c}} + k_5 [(r_{\mathrm{n}}^{\mathrm{c}})^2 + 2(x_{\mathrm{n}}^{\mathrm{c}})^2] \\ k_4 [(r_{\mathrm{n}}^{\mathrm{c}})^2 + 2(y_{\mathrm{n}}^{\mathrm{c}})^2] + 2k_5 x_{\mathrm{n}}^{\mathrm{c}} y_{\mathrm{n}}^{\mathrm{c}} \\ 0 \\ 0 \end{bmatrix}, \tag{7}$$

where $x_{\mathrm{n}}^{\mathrm{c}} = x^{\mathrm{c}}/z^{\mathrm{c}}$, $y_{\mathrm{n}}^{\mathrm{c}} = y^{\mathrm{c}}/z^{\mathrm{c}}$ are the normalized image coordinates, and $r_{\mathrm{n}}^{\mathrm{c}} = \sqrt{(x_{\mathrm{n}}^{\mathrm{c}})^2 + (y_{\mathrm{n}}^{\mathrm{c}})^2}$. The total distortion term $\boldsymbol{\delta} = \boldsymbol{\delta}_{\mathrm{r}} + \boldsymbol{\delta}_{\mathrm{t}}$ can be parameterized by $\mathbf{k} = [k_1, k_2, k_3, k_4, k_5]^\top$. Once all the parameters for a single camera defined as $\boldsymbol{\xi} = [\boldsymbol{\kappa}^\top, (\boldsymbol{\eta}_{\mathrm{w}}^{\mathrm{c}})^\top, \mathbf{k}^\top]^\top$ are known, the forward projection model is fully determined.

2.2. Calibration and Triangulation

The pinhole camera model defines the projective mapping from 3D world coordinates to their corresponding 2D image coordinates. Camera calibration then estimates the parameters of this model based on the 2D projections. In practice, a planar target with well-designed features of high geometric accuracy (e.g., a checkerboard) is used for calibration. Consider N feature points on a target, where the n-th point in the board coordinate system be defined as $\tilde{\mathbf{P}}_n^{\mathrm{b}} = [x_n^{\mathrm{b}}, y_n^{\mathrm{b}}, z_n^{\mathrm{b}}, 1]^\top$. When the board is placed in the l-th pose $\mathbf{T}_l^{\mathrm{w}}$ (out of L poses), its world coordinates are $\tilde{\mathbf{P}}_{n,l}^{\mathrm{w}} = \mathbf{T}_l^{\mathrm{w}} \tilde{\mathbf{P}}_n^{\mathrm{b}}$. The projection of this point onto the m-th camera (out of M cameras) results in observed image coordinates $\tilde{\mathbf{p}}_{n,l}^{m} = [u_{n,l}^{m}, v_{n,l}^{m}, 1]^\top$ can be derived based on Eq. (5):

$$\tilde{\mathbf{p}}_{n,l}^{m} = \mathbf{K}^m \left(\frac{1}{z_{n,l}^{m}} \mathbf{T}_{\mathrm{w}}^{m} \mathbf{T}_l^{\mathrm{w}} \tilde{\mathbf{P}}_n^{\mathrm{b}} + \boldsymbol{\delta}_{n,l}^{m} \right), \tag{8}$$

where $\mathbf{K}^m$ and $\mathbf{T}_{\mathrm{w}}^{m}$ are the intrinsic and extrinsic matrices for m-th camera, and $\boldsymbol{\delta}_{n,l}^{m}$ is the distortion term. Actually, Eq. (8) is redundant, and only two independent equations can be derived from it. Let vector $\boldsymbol{\xi}^m$ represents the m-th camera's parameters and vector $\boldsymbol{\eta}_l^{\mathrm{w}}$ represents l-th target pose $\mathbf{T}_l^{\mathrm{w}}$, Eq. (8) can be rearrange as:

$$\begin{cases} u_{n,l}^{m} - f_u(\boldsymbol{\xi}^m, \boldsymbol{\eta}_l^{\mathrm{w}}; x_n^{\mathrm{b}}, y_n^{\mathrm{b}}, z_n^{\mathrm{b}}) = 0, \\ v_{n,l}^{m} - f_v(\boldsymbol{\xi}^m, \boldsymbol{\eta}_l^{\mathrm{w}}; x_n^{\mathrm{b}}, y_n^{\mathrm{b}}, z_n^{\mathrm{b}}) = 0, \end{cases} \tag{9}$$

where f_u and f_v are nonlinear functions mapping 3D points to u and v image coordinates respectively.

Therefore, for N points, M cameras and L poses, a system with $2N \times M \times L$ equations can be obtained and written as [11]:

$$\mathbf{F}(\boldsymbol{\alpha}; \boldsymbol{\beta}) = 0, \tag{10}$$

where $\boldsymbol{\alpha} = [(\boldsymbol{\xi}^1)^\top, \ldots, (\boldsymbol{\xi}^M)^\top, (\boldsymbol{\eta}_1^{\mathrm{w}})^\top, \ldots, (\boldsymbol{\eta}_L^{\mathrm{w}})^\top]^\top$ includes all unknown parameters to be estimated, and $\boldsymbol{\beta} = [u_{1,1}^1, v_{1,1}^1, \ldots, u_{N,L}^M, v_{N,L}^M, x_1^{\mathrm{b}}, y_1^{\mathrm{b}}, z_1^{\mathrm{b}}, \ldots, x_N^{\mathrm{b}}, y_N^{\mathrm{b}}, z_N^{\mathrm{b}}]^\top$ includes all known coordinates. Note that the target pose indices range from $k = 2$ to L. The first camera is fixed at the origin at the world coordinates ($\boldsymbol{\eta}_{\mathrm{w}}^{1\top} = \mathbf{0}$) to eliminate ambiguity. Calibration is an optimization process to determine $\hat{\boldsymbol{\alpha}}$ when given $\boldsymbol{\beta}$ to minimize the residuals of the function $\mathbf{F}$:

$$\arg\min_{\boldsymbol{\alpha}} \|\mathbf{F}(\boldsymbol{\alpha}; \boldsymbol{\beta})\|^2. \tag{11}$$

Once calibrated, triangulation estimates the 3D world coordinates $\mathbf{P}^{\mathrm{w}} = [x^{\mathrm{w}}, y^{\mathrm{w}}, z^{\mathrm{w}}]^\top$ of an unknown point from its 2D correspondences across multiple views. Similarly, the ideal projection of $\tilde{\mathbf{P}}^{\mathrm{w}} = [x^{\mathrm{w}}, y^{\mathrm{w}}, z^{\mathrm{w}}, 1]^\top$ onto the m-th camera's image coordinates $\tilde{\mathbf{p}}^{m} = [u^{m}, v^{m}, 1]^\top$ is:

$$\tilde{\mathbf{p}}^{m} = \mathbf{K}^m \left(\frac{1}{z^{m}} \mathbf{T}_{\mathrm{w}}^{m} \tilde{\mathbf{P}}^{\mathrm{w}} + \boldsymbol{\delta}^{m} \right), \tag{12}$$

and two independent equations can be derived and rearranged as:

$$\begin{cases} u^{m} - g_u(x^{\mathrm{w}}, y^{\mathrm{w}}, z^{\mathrm{w}}; \boldsymbol{\xi}^m) = 0, \\ v^{m} - g_v(x^{\mathrm{w}}, y^{\mathrm{w}}, z^{\mathrm{w}}; \boldsymbol{\xi}^m) = 0, \end{cases} \tag{13}$$

Copyright ©2026 by ASME

where g_u and g_v are nonlinear functions mapping 3D points to image coordinates u and v respectively. Therefore, for M cameras, a system with $2M$ equations can be obtained and written as:

$$\mathbf{G}(\mathbf{P}^{\mathrm{w}};\boldsymbol{\gamma}) = 0, \tag{14}$$

where $\boldsymbol{\gamma} = [(\boldsymbol{\xi}^1)^\top, \ldots, (\boldsymbol{\xi}^M)^\top, u^1, v^1, \ldots, u^M, v^M]^\top$ contains all known calibrated camera parameters and matched image coordinates across different cameras. Triangulation aims to determine $\mathbf{P}^{\mathrm{w}}$ by minimizing the reprojection residuals:

$$\arg\min_{\mathbf{P}^{\mathrm{w}}} \|\mathbf{G}(\mathbf{P}^{\mathrm{w}};\boldsymbol{\gamma})\|^2. \tag{15}$$

Eqs. (11) and (15) form nonlinear least-squares optimization problems, which are typically solved using iterative nonlinear optimization methods (e.g., the Levenberg–Marquardt algorithm) [18].

2.3. Uncertainty Propagation

In this framework, we assume the camera imaging model to be locally linear. This assumption is justified because the uncertainties are typically small enough to ignore the second- and higher-order terms in the Taylor series expansion. According to the standard of uncertainty propagation [19], the covariance matrix $\boldsymbol{\Lambda}_{\boldsymbol{\alpha}}$ associated with all calibration parameters $\boldsymbol{\alpha}$ can determined through calibration and be calculated by:

$$\boldsymbol{\Lambda}_{\boldsymbol{\alpha}} = \mathbf{J}_{\boldsymbol{\alpha}}^{\dagger}\mathbf{J}_{\boldsymbol{\beta}}\boldsymbol{\Lambda}_{\boldsymbol{\beta}}\mathbf{J}_{\boldsymbol{\beta}}^{\top}(\mathbf{J}_{\boldsymbol{\alpha}}^{\top})^{\dagger}, \tag{16}$$

where $\mathbf{J}_{\boldsymbol{\alpha}} = \partial\mathbf{F}/\partial\boldsymbol{\alpha}$ and $\mathbf{J}_{\boldsymbol{\beta}} = \partial\mathbf{F}/\partial\boldsymbol{\beta}$ are the Jacobian matrices of $\mathbf{F}$ with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, $(\cdot)^\dagger$ is the pseudoinverse of the matrix, and $\boldsymbol{\Lambda}_{\boldsymbol{\beta}}$ is the covariance matrix associated with $\boldsymbol{\beta}$. As the input coordinates are assumed to be independent, $\boldsymbol{\Lambda}_{\boldsymbol{\beta}}$ is diagonal. The components of $\boldsymbol{\Lambda}_{\boldsymbol{\beta}}$ can be divided into two parts and determined according to the standard [19] as follows:

- Image coordinates uncertainty: Let $\boldsymbol{\Lambda}_{\boldsymbol{\beta}_{2\mathrm{D}}}$ represent the uncertainty covariance matrix associated with the 2D coordinates of feature points across different image planes $[u_{1,1}^1, v_{1,1}^1, \ldots, u_{N,L}^M, v_{N,L}^M]^\top$. This part includes the uncertainty in digitizing images and feature extraction. While the theoretical modeling of these uncertainties is complicated [20], they can be estimated from calibration residuals, assuming the uncertainty isotropic and uniform across the image plane:

$$\boldsymbol{\Lambda}_{\boldsymbol{\beta}_{2\mathrm{D}}} = \sigma_{\mathrm{r}}^2\mathbf{I}_{Q_{\mathrm{o}}}, \quad \sigma_{\mathrm{r}}^2 = \frac{\|\mathbf{F}(\boldsymbol{\alpha};\boldsymbol{\beta})\|^2}{Q_{\mathrm{o}} - Q_{\mathrm{s}}} \tag{17}$$

  where $Q_{\mathrm{o}} = 2N \times M \times L$ is the number of the independent equations in calibration, $Q_{\mathrm{s}} = 16M + 6(L-1)$ is the length of $\boldsymbol{\alpha}$ (the number of parameters to be determined in calibration), and $\mathbf{I}$ denotes the identity matrix with the subscript indicating its dimension.

- Target Geometric Uncertainty: Let $\boldsymbol{\Lambda}_{\boldsymbol{\beta}_{3\mathrm{D}}}$ represent the uncertainty covariance matrix associated with the 3D coordinates of feature points on the calibration board $[x_1^{\mathrm{b}}, y_1^{\mathrm{b}}, z_1^{\mathrm{b}}, \ldots, x_N^{\mathrm{b}}, y_N^{\mathrm{b}}, z_N^{\mathrm{b}}]^\top$. This part is evaluated by non-statistical means (Type B uncertainty) which is derived from the manufacturing accuracy specifications of the calibration target (e.g., the lithographic printing tolerance of the checkerboard pattern) and expressed as:

$$\boldsymbol{\Lambda}_{\boldsymbol{\beta}_{3\mathrm{D}}} = \mathbf{I}_N \otimes \mathrm{diag}(\sigma_x^2, \sigma_y^2, \sigma_z^2), \tag{18}$$

  where $\otimes$ is the Kronecker product, and $\sigma_x^2, \sigma_y^2, \sigma_z^2$ are the variance in x, y, z coordinates, respectively.

The total input covariance matrix for calibration is assembled as $\boldsymbol{\Lambda}_{\boldsymbol{\beta}} = \mathrm{diag}(\boldsymbol{\Lambda}_{\boldsymbol{\beta}_{2\mathrm{D}}}, \boldsymbol{\Lambda}_{\boldsymbol{\beta}_{3\mathrm{D}}})$.

For triangulation, the uncertainty of a reconstructed 3D point $\mathbf{P}^{\mathrm{w}}$ is given by:

$$\boldsymbol{\Lambda}_{\mathbf{P}^{\mathrm{w}}} = \mathbf{J}_{\mathbf{P}^{\mathrm{w}}}^{\dagger}\mathbf{J}_{\boldsymbol{\gamma}}\boldsymbol{\Lambda}_{\boldsymbol{\gamma}}\mathbf{J}_{\boldsymbol{\gamma}}^{\top}(\mathbf{J}_{\mathbf{P}^{\mathrm{w}}}^{\top})^{\dagger}, \tag{19}$$

where $\mathbf{J}_{\mathbf{P}^{\mathrm{w}}} = \partial\mathbf{G}/\partial\mathbf{P}^{\mathrm{w}}$ and $\mathbf{J}_{\boldsymbol{\gamma}} = \partial\mathbf{G}/\partial\boldsymbol{\gamma}$. The components of covariance matrix $\boldsymbol{\Lambda}_{\boldsymbol{\gamma}}$ can also be divided into two independent parts:

- Camera parameters uncertainty: Let $\boldsymbol{\Lambda}_{\boldsymbol{\gamma}_{\mathrm{c}}}$ represent the uncertainty covariance matrix associated with all camera parameters $[(\boldsymbol{\xi})^{1\top}, \ldots, (\boldsymbol{\xi}^M)^\top]^\top$. This can be directly determined from $\boldsymbol{\Lambda}_{\boldsymbol{\alpha}}$ by extracting relevant sub-blocks.

- Stereo matching uncertainty: Let $\boldsymbol{\Lambda}_{\boldsymbol{\gamma}_{\mathrm{m}}}$ represent the uncertainty covariance matrix associated with image coordinates $[u^1, v^1, \ldots, u^M, v^M]^\top$, which quantifies the uncertainty of the stereo correspondence. In this work, we employ the zero-mean normalized cross-correlation (ZNCC) criterion to establish correspondences. Assuming $T(\mathbf{q})$ is the matching template, where $\mathbf{q}$ is within the patch window $\mathcal{W}$, the ZNCC score S^i for i-th camera is calculated as:

$$S^i = \frac{\sum_{\mathbf{q}\in\mathcal{W}}[T(\mathbf{q}) - \mu_{\mathrm{T}}]\{I^i[\mathbf{p}^i + H^i(\mathbf{q})] - \mu_{\mathrm{I}^i}\}}{\sigma_{\mathrm{T}}\sigma_{\mathrm{I}^i}} \tag{20}$$

  where $\mu_{\mathrm{T}}, \mu_{\mathrm{I}^i}$ and $\sigma_{\mathrm{T}}, \sigma_{\mathrm{I}^i}$ are the mean and standard deviations of the intensity of the template and the corresponding patch in the i-th camera view, respectively, $\mathbf{p}^i = [u^i, v^i]^\top$, and $H^i(\cdot)$ represents the homography transformation mapping image coordinates from reference view to camera i. The global consistency objective function is calculated as follows:

$$S = \sum_i S^i. \tag{21}$$

  Consequently, $\boldsymbol{\Lambda}_{\boldsymbol{\gamma}_{\mathrm{m}}}$ is computed as follows [21, 22]:

$$\boldsymbol{\Lambda}_{\boldsymbol{\gamma}_{\mathrm{m}}} = -2Q_{\mathrm{w}}\sigma_{\mathrm{r}}^2\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}\left(\nabla^2 S\right)^{-1}, \tag{22}$$

  where $\boldsymbol{\Lambda}_{\mathrm{S}} = \mathrm{diag}(S_{\mathrm{I}^1}, S_{\mathrm{I}^1}, \ldots, S_{\mathrm{I}^M}, S_{\mathrm{I}^M})$, Q_{w} is the number of pixels in the patch, and $\nabla^2 S$ is the Hessian matrix of S with respect to $[u^1, v^1, \ldots, u^M, v^M]^\top$.

The final input covariance matrix for triangulation is assembled as $\boldsymbol{\Lambda}_{\boldsymbol{\gamma}} = \mathrm{diag}(\boldsymbol{\Lambda}_{\boldsymbol{\gamma}_{\mathrm{c}}}, \boldsymbol{\Lambda}_{\boldsymbol{\gamma}_{\mathrm{m}}})$.

Copyright ©2026 by ASME

3. METHOD AND MATERIAL

3.1. System Hardware Setup

The architecture of the proposed MVS system is illustrated in Fig.1. The system is built around three Orbbec Gemini 335 sensing modules arranged symmetrically around the workpiece with an angular separation of 120°. Each module integrates two infrared (IR) cameras and an 850 nm VCSEL speckle projector within a compact and pre-aligned hardware package. In this work, only the raw IR images from the stereo cameras are utilized, while the manufacturer's proprietary depth-processing pipeline is bypassed. This design allows the implementation of a custom multi-view stereo reconstruction framework and the subsequent uncertainty analysis.

The three sensing modules are mounted at a height of 180 mm above the center of the workpiece, with a horizontal distance of 150 mm from the object surface. This configuration results in six distinct viewpoints surrounding the target object, providing sufficient geometric redundancy to mitigate occlusions and ensure that most surface regions are observed from multiple perspectives.

All six IR cameras operate at a resolution of 1280×800 pixels with a frame rate of 30 frames per second. To maintain photometric consistency across views, the automatic exposure and gain controls are disabled, and the exposure time is fixed at 2 ms. A hardware-trigger synchronization mechanism is employed to ensure simultaneous image acquisition across all cameras. The trigger signal activates all six cameras and the three VCSEL projectors concurrently, guaranteeing temporal alignment between the captured images.

During acquisition, the three projectors simultaneously illuminate the workpiece with their speckle patterns. The superposition of these patterns produces a denser and more complex composite texture on the object surface compared with a single projector. This enriched texture improves feature distinctiveness and reduces correspondence ambiguity, which is particularly beneficial when reconstructing low-texture metallic surfaces [23]. All experiments are conducted indoors under controlled lighting conditions to minimize interference from ambient infrared sources.

Prior to data acquisition, all cameras are calibrated using a multi-view calibration procedure with a planar checkerboard target. The intrinsic and extrinsic parameters are estimated using the multi-camera calibration toolbox in MATLAB, and the calibration reprojection error is below 0.3 pixels for all cameras. The resulting camera parameters are subsequently used for multi-view triangulation and uncertainty propagation in the reconstruction pipeline.

3.2. MVS Matching and Reconstruction

To validate the proposed uncertainty framework, a complex free-form metal surface was measured. Matching pixels across images is a challenging problem. traditional multi-view matching methods are difficult to apply to our metal sheet measurement scenario due to the highly reflective surface and the strong similarity between the projected speckle patterns. To address these challenges, we propose a multi-stage approach to establish reliable correspondences across views (Fig. 2):

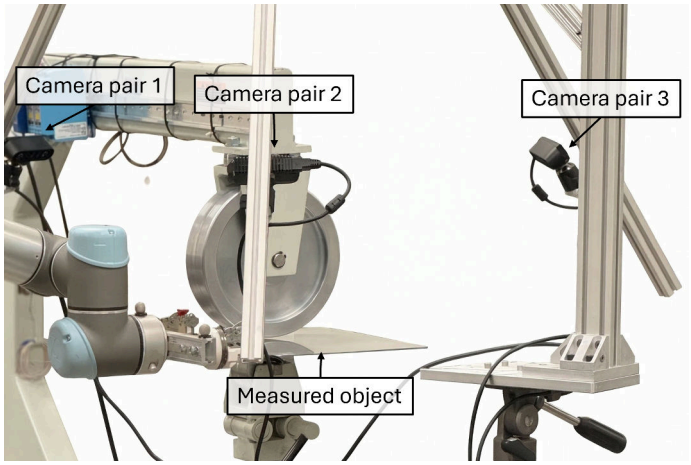


FIGURE 1: Schematic of the proposed multi-view stereo imaging system consisting of three stereo camera modules arranged around the workpiece.

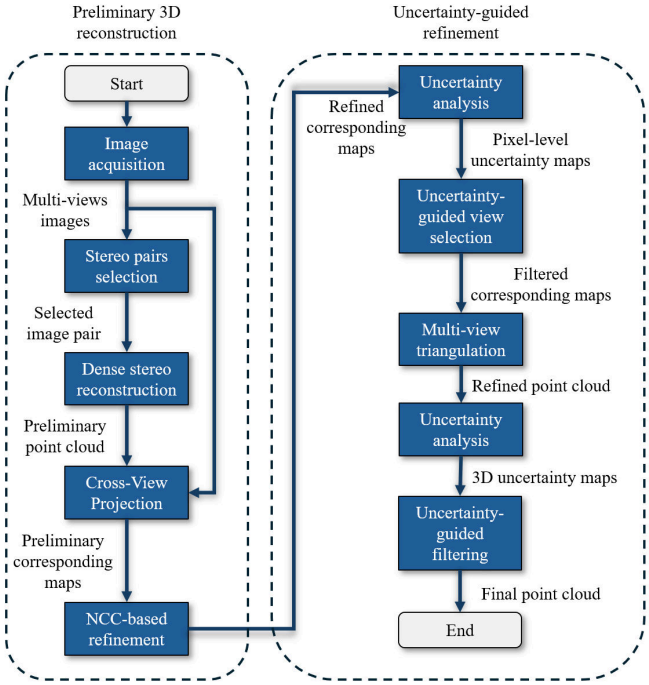


FIGURE 2: Workflow of the proposed uncertainty-aware MVS reconstruction method.

Copyright ©2026 by ASME

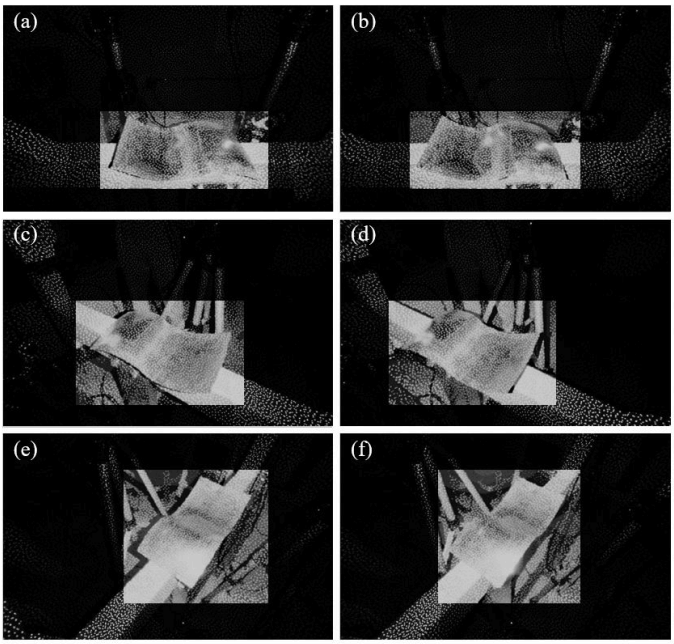


FIGURE 3: Captured images of a deformed metal sheet from three pairs of stereo cameras. Each pair of the images, (a, b), (c, d), and (e, f), corresponds to one stereo camera pair. The measured object is highlighted for better visualization.

1. **Initial stereo reconstruction:** We first select a pair of Orbbec Gemini stereo images and perform epipolar rectification. Using stereo disparity and triangulation, we obtain an initial 3D point cloud. Background points are removed, leaving only the metal sheet region. When a single pair cannot fully observe the object or is affected by occlusions, multiple camera pairs can be employed to obtain a more complete initial point cloud. Although this initialization may exhibit slight layering or misalignment due to extrinsic calibration uncertainties, these effects will be corrected during the refinement in the following steps.

2. **Cross-view projection and reference view selection:** The reconstructed 3D points are then projected into the remaining camera views, providing an initial estimate of their corresponding pixel locations in each image. For each point, we select the view with the smallest angle between the surface normal and the optical axis, and extract a local patch as the reference patch template. In cases where the primary view is obscured or affected by highlights, we iteratively evaluate the remaining views in ascending order of their angular deviation.

3. **Patch-based refinement using ZNCC:** For each other view, we perform a constrained search around the initial guess and compute ZNCC to determine the optimal matching location. We assume that each local patch can be approximated as a planar surface. This assumption allows us to use the projective transform to warp the reference patch into the other camera views [24], thereby reducing geometric distortions caused by viewpoint changes and improving matching robustness.

TABLE 1: Input uncertainty of image coordinates and calibration target geometry.

Parameter	Value
σ_r	0.132 pixels
σ_x	0.010 mm
σ_y	0.010 mm
σ_z	0.005 mm

After establishing dense pixel correspondences across all camera views, we perform multi-view triangulation to obtain the dense 3D reconstruction. In principle, using multiple views provides stronger geometric constraints than a stereo pair, since additional viewing angles should improve depth accuracy and robustness. However, in practice, multi-view triangulation is highly sensitive to outlier or low-quality measurements. Even a single corrupted correspondence, caused by specular highlights, shadows, occlusions, or weak texture, can introduce large bias to the estimated 3D location. As a result, simply fusing all available views may produce reconstructions that are less accurate.

To overcome this issue, we quantify the matching uncertainty using Eq. (22) first and filter the views lacking reliability. Regions affected by specular reflection typically exhibit high uncertainty because the resulting consistency objective function will have a flat peak due to ambiguity. Triangulation is then performed using only the subset of views that provide stable geometric support. This adaptive view-selection strategy ensures that the reconstruction is dominated by reliable observations. Finally, we assess the total uncertainty of the resulting 3D point cloud by accounting for both matching residuals and calibration errors. Points exceeding an uncertainty threshold are pruned to eliminate artifacts and ensure a high-fidelity output.

To validate the metrological performance of the proposed system, a high-precision industrial scanner (Hexagon Romer Absolute Arm, model RA7525SI) was used to acquire ground truth data. This metrology-grade instrument has a specified scanning system accuracy of 50 µm. This level of precision provides a reliable benchmark for evaluating the accuracy of our reconstruction.

4. RESULTS

4.1. Theoretical Uncertainty Evaluation

We first performed multi-view camera calibration. Because the lens distortion is small, only two distortion parameters, $[k_1, k_2]^\top$, are used. In addition, the skew coefficient s is assumed to be negligible. The input uncertainties in the image coordinates and in the calibration target geometry are listed in Table 1.

The uncertainty of the camera parameters is then evaluated using Eq. (16). Figure 4 presents the covariance matrix of the camera parameters. For each camera, the parameter vector contains twelve parameters:

$$[f_x, f_y, u_0, v_0, k_1, k_2, \varphi_w^c, \theta_w^c, \psi_w^c, t_{w,x}^c, t_{w,y}^c, t_{w,z}^c]^\top.$$

A block-wise structure can be observed, corresponding to the parameter groups of different cameras. The diagonal elements indicate the variances of individual parameters, whereas

Copyright ©2026 by ASME

the off-diagonal elements represent parameter correlations. The presence of non-negligible off-diagonal terms shows that the estimated parameters are coupled, revealing the interdependence between intrinsic and extrinsic parameters in the calibration process.

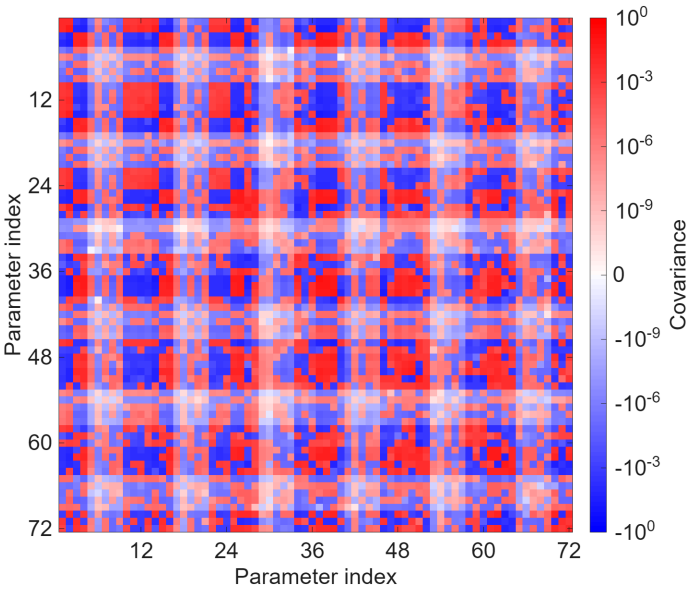


FIGURE 4: Covariance matrix of the camera parameters.

Before evaluating the proposed reconstruction method on real experimental environment, we first analyze the theoretical behavior of uncertainty in multi-view reconstruction. This analysis provides a reference for understanding how geometric factors influence reconstruction accuracy and helps explain the motivation behind the uncertainty-aware strategy introduced in this work. In particular, we examine two complementary aspects of uncertainty: the dependence of triangulation accuracy on object distance and the spatial variation of projection uncertainty across the image plane. These results provide an intuitive illustration of the uncertainty propagation framework described in Section 2.

The depth uncertainty (uncertainty in z-direction) curves in Fig. 5 characterize the theoretical limits of stereo triangulation as a function of object distance. In this analysis, the uncertainty is evaluated for stereo camera pair 1 using Eq. (19). The uncertainty in the input camera parameters, $\mathbf{\Lambda}_{\boldsymbol{\gamma}_c}$, is obtained from calibration, and the stereo matching uncertainty, $\mathbf{\Lambda}_{\boldsymbol{\gamma}_m}$, is assumed to be:

$$\mathbf{\Lambda}_{\boldsymbol{\gamma}_m} = \begin{bmatrix} 0.04 & 0.01 \\ 0.01 & 0.04 \end{bmatrix}, \tag{23}$$

which is a typical value for high-quality stereo matching.

The resulting curves show that the depth uncertainty increases nonlinearly as the object distance grows, consistent with the simplified stereo triangulation model for parallel cameras. Consequently, even a small subpixel matching error leads to significantly larger depth errors at greater distances. The curves therefore illustrate the intrinsic geometric limitation of triangulation-based reconstruction: as objects move farther away from the cameras, the achievable depth accuracy deteriorates even when image matching precision remains constant.

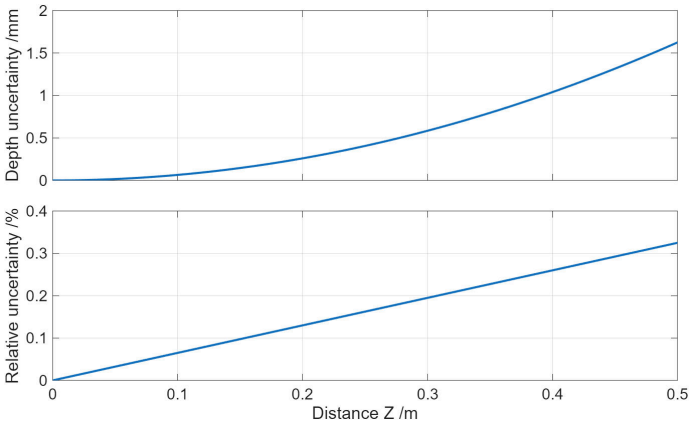


FIGURE 5: Depth uncertainty as a function of object distance for the stereo camera pair 1.

While the previous analysis describes the overall effect of distance on depth uncertainty estimation, the projection uncertainty map shown in Fig. 6 illustrates how camera intrinsics uncertainty affects the projected pixel location within a single camera view. This analysis is based on the uncertainty propagation framework developed in Section 2.3. Assuming that the covariance matrix for 3D world coordiantes is zero, the covariance of the pixel uncertainty $\mathbf{\Lambda}_{\mathbf{p}^c}$ can be caculated as:

$$\mathbf{\Lambda}_{\mathbf{p}^c} = \mathbf{J}_{\boldsymbol{\gamma}_c} \mathbf{\Lambda}_{\boldsymbol{\gamma}_c} \mathbf{J}_{\boldsymbol{\gamma}_c}^{\top}, \tag{24}$$

where $\mathbf{J}_{\boldsymbol{\gamma}_c} = \partial \mathbf{F} / \partial \boldsymbol{\gamma}_c$ the Jacobian matrices of $\mathbf{F}$ with respect to $\boldsymbol{\gamma}_c$.

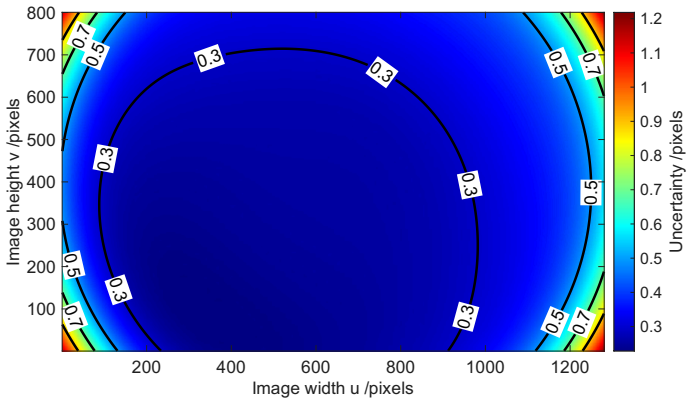


FIGURE 6: Projection uncertainty heatmap for left camera in pair 1 at a fixed object depth of z^c = 0.3 m.

For a fixed object depth of $z^c = 0.3$ m, a dense grid of image locations is evaluated across the sensor. Each pixel location is associated with a corresponding 3D point at the specified depth, and the projection uncertainty is evaluated accordingly. The scalar uncertainty value visualized in the heatmap is defined as the root-sum variance $\sqrt{tr(\mathbf{\Lambda}_{\mathbf{p}^c})}$, where $tr(\cdot)$ denotes the trace of the covariance matrix.

The resulting uncertainty map reveals a clear spatial structure across the image plane. The uncertainty is lowest near the image center and gradually increases toward the periphery, forming a

Copyright ©2026 by ASME

characteristic radial pattern. This behavior reflects the increasing sensitivity of the projection geometry to parameter perturbations at larger viewing angles, as well as the influence of lens distortion and perspective effects. The contour overlays provide quantitative markers of the uncertainty distribution across the field of view.

To further clarify the sources contributing to the reconstruction uncertainty, Table 2 summarizes the uncertainty budget in the proposed system. The total uncertainty of the reconstructed 3D points arises from three primary components: camera calibration uncertainty, stereo matching uncertainty, and the geometric sensitivity of triangulation. Calibration uncertainty affects the projection model and leads to spatially varying reprojection uncertainty across the image plane (Fig. 6). Stereo matching uncertainty originates from the image correspondence estimation process and is modeled based on the NCC matching criterion. Finally, triangulation geometry introduces additional sensitivity that increases with object distance, as illustrated in Fig. 5. Together, these components determine the overall reconstruction uncertainty.

4.2. Reconstruction Uncertainty and Comparison

The captured images of the metal sheet (Fig. 3) shows severe highlight saturation and similar projected pattern which makes conventional MVS algorithms prone to unstable or ambiguous correspondences. Based on the proposed method, we use Fig. 3(a) and 3(b) to give an initial guess for the 3D point cloud. The uncertainty contribution of each view is then evaluated for view selection. Figure 7 shows the retained and filtered pixels for each view. Pixels tend to be rejected in regions with large surface curvature or with a large viewing angle relative to the surface normal. Under these conditions, the projected pattern is significantly distorted, resulting in poor stereo-matching reliability. This behavior agrees with visual intuition and the purpose of the proposed view-selection process.

Figure 8 shows the 3D reconstruction uncertainty and accuracy maps, which provide a quantitative, per-point assessment of the reconstruction's reliability. The uncertainty map obtained from all views exhibits large high-uncertainty regions near strong specular reflections (see in Fig. 3(e) and 3(f)). These regions correspond to areas of poor photometric consistency which influence the NCC matching algorithm. In contrast, the uncertainty map obtained after applying the proposed view-selection strategy shows a clear reduction in uncertainty in those area.

The corresponding height-difference maps relative to the laser-scanner ground truth (Fig. 8(c) and 8(d)) reveal a consistent reduction in reconstruction error when using the selected views. This strong spatial correlation between the uncertainty map and the reconstruction error confirms the effectiveness of the proposed uncertainty model as a predictor of metrological

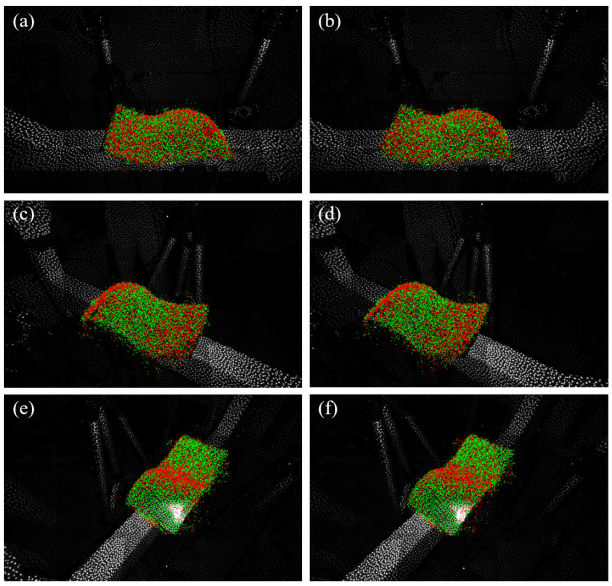


FIGURE 7: Selected and filtered pixels in each view. Pixels retained by the proposed view-selection strategy are shown in green, while pixels rejected during filtering are shown in red.

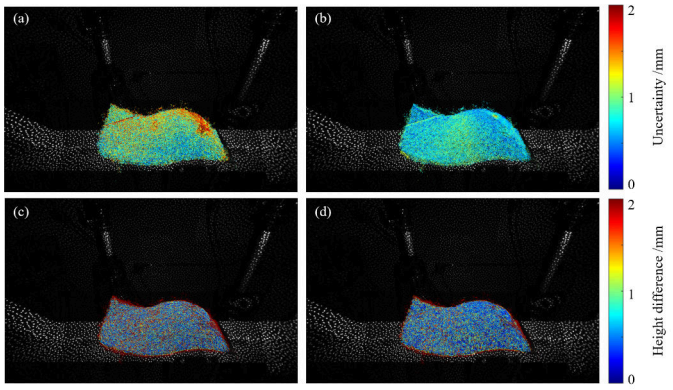


FIGURE 8: Comparison of 3D points reconstruction results. The height difference is obtained from the comparison with the results from a laser scanner. For better visualization, the uncertainty and height difference are projected to the reference image coordinates. (a) and (c) are the uncertainty and height difference of the 3D points reconstructed using all six views, while (b) and (d) are those using selected views.

TABLE 2: Uncertainty budget of the reconstructed 3D points in the proposed MVS framework.

Source of Uncertainty	Typical Magnitude	Influence on Reconstruction
Camera calibration	~0.4 pixels (reprojection error)	dominant near image periphery
Stereo matching	~0.2 pixels (matching error)	dominant in low-texture regions
Triangulation geometry	increases with distance z^c	dominant at large z^c
Reconstructed 3D points	up to ~2 mm in the experiments	—

Copyright ©2026 by ASME

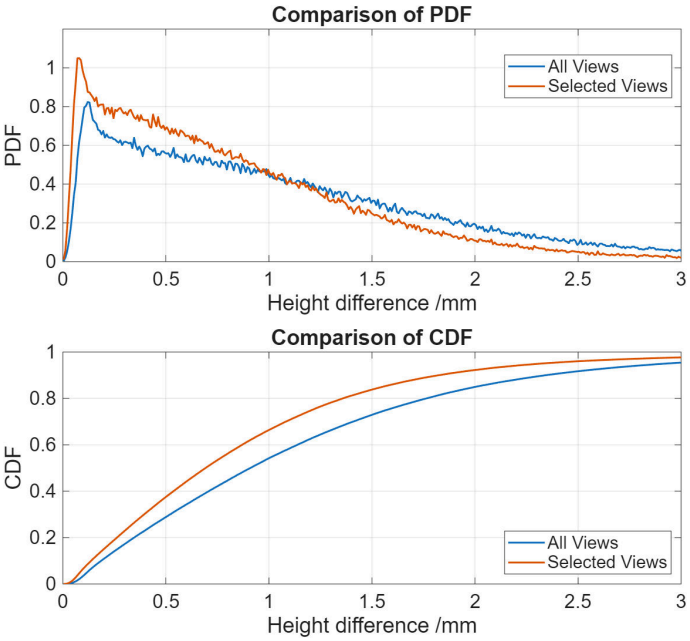


FIGURE 9: Comparison of the probability density function (PDF) and cumulative distribution function (CDF) of the height difference from all camera views and a selected subset of views.

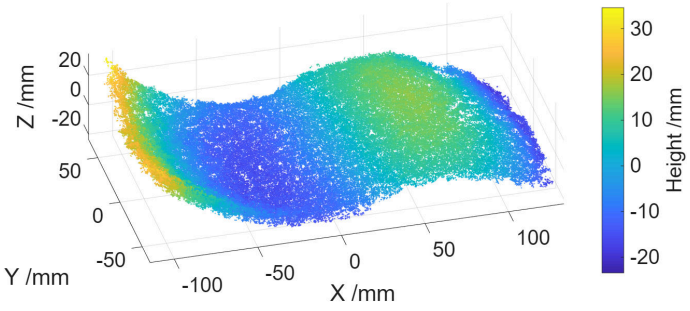


FIGURE 10: Reconstructed 3D point cloud of a deformed metal sheet from uncertainty filtering.

reliability.

Figure 9 gives statistical comparison of the height-difference error distributions obtained from all camera views versus those computed using only a selected subset of views. The probability density functions (PDFs) in the top row show that both distributions are highly skewed toward small errors, with a sharp peak near zero. However, the selected-view configuration exhibits a slightly higher peak and a steeper decay, indicating that it produces more measurements concentrated near zero error and fewer large deviations. This behavior is further confirmed by the CDF curves in the bottom plot: the selected-view CDF rises more quickly, meaning that a larger fraction of points achieve low reconstruction error compared to the all-view configuration.

Beyond guiding optimal view selection, the estimated point-wise uncertainty also enables direct quality filtering of the reconstructed 3D point cloud. Figure 10 shows the final reconstructed point cloud of the free-form metal surface. Here, we remove all points with a predicted 3D positional uncertainty greater than 2 mm, resulting in a cleaner and more reliable representation of the surface.

Overall, these results demonstrate that uncertainty-aware muti-view stereo improves the precision of 3D reconstruction, particularly under challenging imaging conditions such as high reflectivity and weak texture.

5. CONCLUSION

This study presents an uncertainty-aware framework for quantifying and managing 3D reconstruction uncertainty in MVS systems. The proposed approach explicitly models the propagation of errors originating from camera calibration and image matching, enabling the estimation of uncertainty associated with reconstructed 3D points. Through this formulation, both the geometric characteristics of triangulation and the spatial variation of projection uncertainty can be systematically analyzed.

Building on this uncertainty estimation, an uncertainty-guided view selection strategy is introduced to improve reconstruction reliability. By identifying and excluding unstable observations, particularly those affected by challenging surface conditions such as specular reflections or unfavorable viewing geometries, the proposed method reduces the influence of unreliable measurements during triangulation.

Experimental evaluations demonstrate that the uncertainty-aware strategy consistently improves reconstruction accuracy and robustness compared with conventional MVS approaches that utilize all available views indiscriminately. The results show that incorporating uncertainty information into the reconstruction process leads to more stable and reliable 3D measurements.

Overall, the presented framework establishes a principled approach for integrating uncertainty analysis into multi-view reconstruction. By providing quantitative confidence estimates for reconstructed points and enabling uncertainty-guided measurement selection, the method contributes toward more reliable and traceable 3D metrology. This capability is particularly valuable for advanced manufacturing applications, where robust in-situ monitoring and measurement reliability are critical.

ACKNOWLEDGMENTS

The authors would like to acknowledge support from the U.S. National Science Foundation Engineering Research Center for Hybrid Autonomous Manufacturing Moving from Evolution to Revolution (ERC-HAMMER) under Award Number EEC-2133630.

REFERENCES

[1] Daehn, Glenn, Blue, Craig, Johnson-Bey, Charles, Lewandowski, John J., Mahoney, Tom, Okwudire, Chinedum, Rossman, Tali, Schmitz, Tony and Silveston, Rebecca. "Emerging Opportunities in Distributed Manufacturing: Results and Analysis of an Expert Study." *Integrating Materials and Manufacturing Innovation* Vol. 13 No. 3 (2024): pp. 688–702.

[2] Suresh, Vignesh, Balasubramaniam, Badrinath, Yeh, Li-Hsin and Li, Beiwen. "Recent Advances in In Situ 3D Surface Topographical Monitoring for Additive Manufacturing Processes." *Journal of Manufacturing and Materials Processing* Vol. 9 No. 4 (2025): p. 133.

Copyright ©2026 by ASME

[3] Chheang, Vuthea, Narain, Saurabh, Hooten, Garrett, Cerda, Robert, Au, Brian, Weston, Brian, Giera, Brian, Bremer, Peer-Timo and Miao, Haichao. "Enabling additive manufacturing part inspection of digital twins via collaborative virtual reality." *Scientific Reports* Vol. 14 No. 1 (2024): p. 29783.

[4] Nzumile, Jailos Mrisho, Mahabi, Victoria and Taifa, Ismail W. R. "Contribution of Industry 4.0 Technologies in Adopting Metrology 4.0 in Manufacturing Industries." Machado, Carolina and Davim, J. Paulo (eds.). *Smart Engineering Management*. Springer International Publishing, Cham (2024): pp. 43–72.

[5] He, Xu, Welo, Torgeir and Ma, Jun. "In-process, real-time monitoring of forming forces in rotary draw bending process." *The International Journal of Advanced Manufacturing Technology* Vol. 134 No. 9 (2024): pp. 4651–4666.

[6] Alonso, Marcos, Maestro, Daniel, Izaguirre, Alberto, Andonegui, Imanol and Graña, Manuel. "Depth Data Denoising in Optical Laser Based Sensors for Metal Sheet Flatness Measurement: A Deep Learning Approach." *Sensors* Vol. 21 No. 21 (2021): p. 7024.

[7] Claus, Felix, Hamann, Bernd, Leitte, Heike and Hagen, Hans. "Decomposing deviations of scanned surfaces of sheet metal assemblies." *Journal of Manufacturing Systems* Vol. 61 (2021): pp. 125–138.

[8] Luo, Hua, Zhang, Ke, Li, Ruifeng, Xia, Yu and Liang, Peidong. "A structured-light 3D sensor-based shape measurement method for casting allowance evaluation." *Frontiers in Physics* Vol. 10 (2022).

[9] Pasinetti, Simone, Nuzzi, Cristina, Luchetti, Alessandro, Zanetti, Matteo, Lancini, Matteo and De Cecco, Mariolino. "Experimental Procedure for the Metrological Characterization of Time-of-Flight Cameras for Human Body 3D Measurements." *Sensors* Vol. 23 No. 1 (2023): p. 538.

[10] Guo, Jian, Tan, Dingzhong, Guo, Shizhe, Chen, Zheng and Liu, Rang. "Digital Inspection Technology for Sheet Metal Parts Using 3D Point Clouds." *Sensors* Vol. 25 No. 15 (2025): p. 4827.

[11] Chen, Jiechun, Ding, Zhenliang and Yuan, Feng. "Theoretical Uncertainty Evaluation of Stereo Reconstruction." *2008 2nd International Conference on Bioinformatics and Biomedical Engineering*: pp. 2378–2381. 2008.

[12] Di Leo, Giuseppe, Liguori, Consolatina and Paolillo, Alfredo. "Covariance Propagation for the Uncertainty Estimation in Stereo Vision." *IEEE Transactions on Instrumentation and Measurement* Vol. 60 No. 5 (2011): pp. 1664–1673.

[13] Sankowski, W., Włodarczyk, M., Kacperski, D. and Grabowski, K. "Estimation of measurement uncertainty in stereo vision system." *Image and Vision Computing* Vol. 61 (2017): pp. 70–81.

[14] Zhang, Zhengyou. "A flexible new technique for camera calibration." *IEEE Transactions on pattern analysis and machine intelligence* Vol. 22 No. 11 (2002): pp. 1330–1334.

[15] Diebel, James et al. "Representing attitude: Euler angles, unit quaternions, and rotation vectors." *Matrix* Vol. 58 No. 15-16 (2006): pp. 1–35.

[16] Heikkila, Janne and Silvén, Olli. "A four-step camera calibration procedure with implicit image correction." *Proceedings of IEEE computer society conference on computer vision and pattern recognition*: pp. 1106–1112. 1997. IEEE.

[17] Tsai, Roger. "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses." *IEEE Journal on Robotics and Automation* Vol. 3 No. 4 (2003): pp. 323–344.

[18] Levenberg, Kenneth. "A method for the solution of certain non-linear problems in least squares." *Quarterly of applied mathematics* Vol. 2 No. 2 (1944): pp. 164–168.

[19] ISO., I and OIML, BIPM. *Guide to the Expression of Uncertainty in Measurement*. Aenor Madrid, Spain (1993).

[20] De Santo, Massimo, Liguori, Consolatina, Paolillo, Alfredo and Pietrosanto, Antonio. "Standard uncertainty evaluation in image-based measurements." *Measurement* Vol. 36 No. 3-4 (2004): pp. 347–358.

[21] Förstner, Wolfgang. "Reliability analysis of parameter estimation in linear models with applications to mensuration problems in computer vision." *Computer Vision, Graphics, and Image Processing* Vol. 40 No. 3 (1987): pp. 273–310.

[22] Zhang, Hongmou, Grießbach, Denis, Wohlfeil, Jürgen and Börner, Anko. "Uncertainty model for template feature matching." *Pacific-Rim Symposium on Image and Video Technology*: pp. 406–420. 2017. Springer.

[23] "Projectors for D400 Series Depth Cameras."

[24] Hartley, Richard and Zisserman, Andrew. *Multiple view geometry in computer vision*. Cambridge university press (2003).

Copyright ©2026 by ASME