Lambertian Reflection Separation Under High Reflectiveness for Worn Surface Reconstruction With Insufficient Samples

Qinghua Wang, Shuo Wang, *Member, IEEE*, Tonghai Wu, Tao Shao, Yue Shu, and Thompson Sarkodie-Gyan, *Member, IEEE*

***Abstract*— Compared with 2-D images, 3-D image-based worn inspection can provide more profound information regarding the in situ endoscope examination of mechanical parts. With photometric stereo vision (PSV), 3-D topography can be reconstructed from 2-D images with minimal cost. However, this promising technology is limited by the imprecision in calculating normal vector due to the high reflectiveness of worn surfaces. An innovative model is here developed to separate Lambertian reflection from photometric images. The Lambertian reflection variations are estimated by fusing multiple photometric images, while the constants are extracted from the general plane base of worn surfaces. Considering a small sample set of worn surfaces, hybrid samples, including the real and simulated images, are adopted for parameter optimization with feature-level alignment (FA). Moreover, a label-free prediction loss (LFL) function is constructed to constrain the optimization direction via real samples. The constructed model is obtained from training 20 000 simulated samples and 1000 real samples. The results indicate that high reflectiveness can be addressed by separating the Lambertian reflections. Comparatively, the model can improve the normal vector calculation with lower average error than other existing algorithms.**

***Index Terms*— Lambertian reflection separation, photometric stereo vision (PSV), semi-supervised learning, worn surface reconstruction.**

I. Introduction

WEAR of mechanical parts inevitably occurs even in normal working conditions, e.g., gears in a wind turbine. Hence, the worn surface examination with an industrial endoscope is a dominant means for regular patrol inspection [1], [2]. So far, worn surface analysis (WSA) remains an irreplaceable technique for health monitoring, even with so many advanced on-line monitoring methods. Traditional 2-D image examination could not provide depth information, which may induce erroneous judgments. Therefore, 3-D topography reconstruction is developed driven by the photometric stereo vision (PSV) technology [3], [4], in which the 3-D properties can be determined from multiple 2-D images without additional devices. Nevertheless, it remains a challenge with inaccurate normal vector calculation due to non-Lambertian reflections of metal surfaces that are very common in local grinding or cutting. It is, therefore, desirable to separate Lambertian reflection under high reflectiveness when applying 3-D reconstruction technology.

PSV only takes advantage of monocular vision compared with other 3-D imaging methods. Under the Lambert reflectance assumption, it calculates the normal vectors from photometric images and recovers the surface heights [5], [6]. However, worn surfaces usually appear with specular reflections and shadows, which deviate from the assumption and result in errors in the calculated normal vectors [7]. Considering multiple non-Lambertian reflections, various numerical algorithms have been proposed in PSV. For instance, reflectance model-based methods adopt the bidirectional reflectance distribution function (BRDF) to substitute the ideal Lambert reflectance model for dealing with various reflections [8], [9]. Although these methods can fully use photometric information, complex BRDFs for worn surfaces may introduce complex nonlinear optimization problems. Moreover, outlier removal approaches are developed to extend the normal vector calculation to general reflections by removing non-Lambertian pixels [10], [11]. In essence, these algorithms depend on more than 20 photometric images to choose Lambertian pixels with specific illumination directions. However, such dense photometric images are often difficult to be acquired in real-world WSA.

Driven by deep learning, sparse image-based PSV gradually develops and makes it possible for precise reconstructions with in situ conditions [12], [13]. Deep network models, represented by photometric stereo fully convolutional network (PS-FCN) [14] and normal vector estimation network (NE-Net) [15], are developed to handle different reflective properties and enable topography reconstruction with only

Manuscript received 13 December 2022; revised 28 April 2023; accepted 17 May 2023. Date of publication 31 May 2023; date of current version 23 June 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 52105159 and Grant 51975455, in part by the China Postdoctoral Science Foundation under Grant 2021M702594, and in part by the Open Foundation of State Key Laboratory of Compressor Technology (Compressor Technology Laboratory of Anhui Province) under Grant SKL-YSJ202102. The Associate Editor coordinating the review process was Dr. Dong Wang. *(Corresponding authors: Shuo Wang; Tonghai Wu.)*

Qinghua Wang, Shuo Wang, Tonghai Wu, and Tao Shao are with the Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China (e-mail: wangqh@stu.xjtu.edu.cn; wangshuo@xjtu.edu.cn; tonghai.wu@mail.xjtu.edu.cn; shaotao9192@stu.xjtu.edu.cn).

Yue Shu is with the State Key Laboratory of Compressor Technology (Compressor Technology Laboratory of Anhui Province), Hefei, Anhui 230031, China (e-mail: hgmri.shuyue@qq.com).

Thompson Sarkodie-Gyan is with the Laboratory for Industrial Metrology and Automation, Department of Electrical and Computer Engineering, The University of Texas at El Paso, El Paso, TX 79968 USA (e-mail: tsarkodi@utep.edu).

Digital Object Identifier 10.1109/TIM.2023.3281567

eight photometric images. However, the effectiveness of these methods is low due to complex transformation from photometric images to normal vectors with diverse surface topography. In view of this, reflection separation is introduced as an image preprocessing stage to simplify the normal vector calculation by extracting the required Lambertian reflections from photometric images. The new strategy liberates the deep learning task from multimodal mapping, and Lambertian reflection separation emerges as a critical step of surface reconstitution. Nevertheless, the existing reflection separation models may be subject to local prediction errors for worn surfaces due to high reflectiveness, which will generate specular reflections covering the Lambertian reflections [16], [17].

Apart from the high reflectivity of worn surfaces, learning-based methods are challenged by insufficient samples, which limits the optimization of network parameters [12]. Practically, sample collection is a time-consuming and laborious task involving worn surface collection, photometric image acquisition, 3-D topography measurement, manual image alignment, etc. Therefore, worn surface samples are available in a few numbers and the absence of Lambertian reflection labels. Although the rendering-based generation method has been developed to support the training of deep learning networks, the simulated images are deviated from real images due to the degradation of surface topography and imaging process [15]. This discrepancy may keep the trained models away from real worn surfaces and lead to low separation accuracy in Lambertian reflection.

To address these issues, a Lambertian reflection separation model is proposed to improve the topography reconstruction of worn surfaces. The constructed model is focused on worn regions by reflection decomposition, and photometric images are fused to estimate Lambertian reflections in a complementary way. Considering the limitation of insufficient samples, both real and simulated samples are combined for parameter optimization. In the first instance, feature alignment is introduced to extract features with uniform distribution from both types of samples. This is followed by a label-free loss to evaluate the prediction accuracy of real samples based on the Lambertian reflection characteristics. The effectiveness of the proposed model is evaluated by comparing the normal vector with that obtained from reference algorithms. This work is expected to improve the topography reconstruction of worn surfaces.

The rest of this article is organized as follows: Section II contains the description of the proposed model, including the model framework, feature-level alignment (FA), and loss function. The details of the dataset and implementation are given in Section III, followed by verifications in Section IV. The discussions are presented in Section V. The conclusions are presented in Section VI.

II. Model

A Lambertian reflection separation network (LRS-Net) model is constructed, as shown in Fig. 1, to extract the desired reflection for photometric stereo reconstruction. The model structure is composed of two branches of variation estimation and constant extraction for Lambertian reflections.

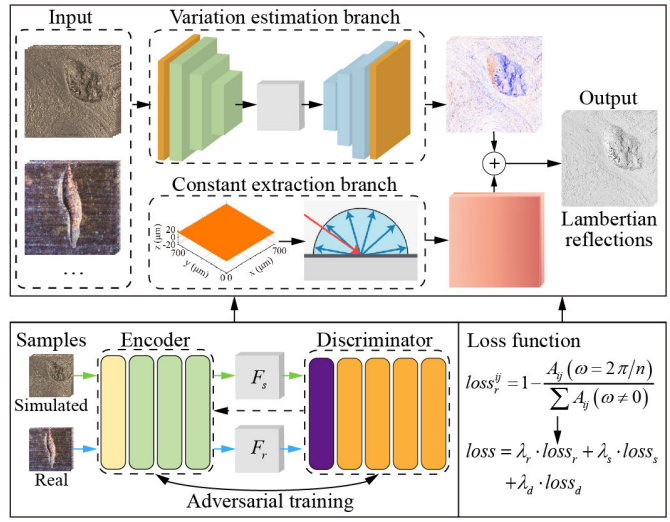


Fig. 1. Framework and optimization of the LRS-Net model.

The former branch is specified for robust reflection prediction in high-reflectiveness regions, while the latter one is designed to generate reflection constants from a specific surface topography. For model training, real and simulated samples are hybridized by FA to address insufficient samples of worn surfaces. Furthermore, a joint loss function is constructed for both kinds of samples to constraint parameter optimization according to sample labels and reflection characteristics. The details of the LRS-Net model are described below.

A. Framework of the LRS-Net Model

Worn topography can be identified as the morphological variations from an initial plain surface. Given this, the worn surface topography can be decomposed into local wear features and a plane base. Hence, the surface normal vector $\boldsymbol{N}$ is the superposition of the vertical vector $\boldsymbol{N}_v$ of the plane base and the inclined vector $\boldsymbol{N}_i$ of local features, as depicted in Fig. 2. Furthermore, the Lambertian reflection can be decomposed into two components representing the inclined and vertical vectors. The reflections from inclined vectors are varied with local height fluctuation, while the reflections from vertical vectors are invariant and equal to the Lambertian reflections from the plane base. Therefore, the separation task can be split into variation estimation and constant extraction for Lambertian reflections to simplify the implementation.

As shown in Fig. 1, the LRS-Net model is constructed with two main branches specified for variation reflection estimation and constant reflection extraction. The outputs of both the branches are summed pixel by pixel to generate the Lambertian reflection of the target surface.

1) Variation Estimation Branch: Regarding the extreme brightness from specular reflections, the variation estimation from a single image could be an ill-posed problem. For photometric images, specular reflections with the same pixel position are only observed in individual images with particular illumination directions. Moreover, Lambertian reflections from different photometric images are co-related by the sine

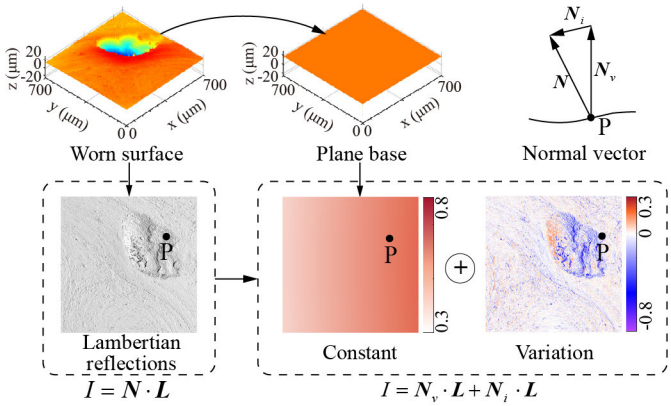


Fig. 2. Decomposition of Lambertian reflection on worn surfaces.

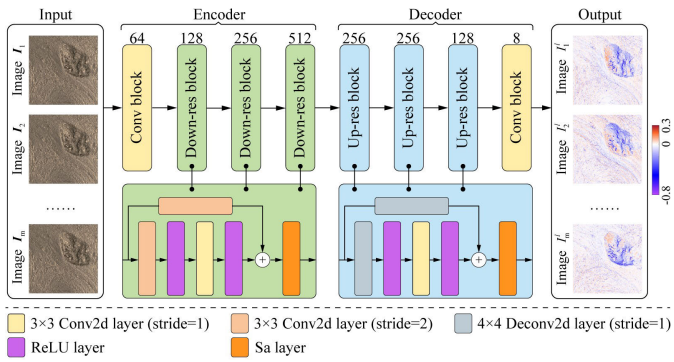


Fig. 3. Structure of the variation estimation branch (note that convolutional and deconvolutional layers in the same residual block adopt equal channel numbers. These are marked on the corresponding convolutional and residual blocks).

function [18] in the following equation:

$$I^k = \left(\rho \sin\theta^k \sqrt{N_x^2 + N_y^2}\right) \times \sin\left(\tau^k + \sin^{-1}\left(N_x / \sqrt{N_x^2 + N_y^2}\right)\right) + \left(\rho \cos\theta^k N_z\right) \tag{1}$$

where I^k is the imaging brightness at the pixel (i, j) in the kth image, $\boldsymbol{N} = [N_x\ N_y\ N_z]$ is the normal vector at the same pixel, ρ denotes the surface reflectance, θ^k is the slant angle of the kth illumination source, and τ^k is the tilt angle of the kth illumination source. Thereby, the Lambertian reflections masked by specular reflections can be inferred from additional photometric images. Given this, the variation estimation branch fuses multiple photometric images via an encoder–decoder structure, as shown in Fig. 3. The specific structures are described in the following.

The constructed variation estimation branch fuses multiple photometric images and then outputs the corresponding Lambertian reflection variations. As input to the branch, photometric images are concatenated along the channel dimension and then integrated for multistage fusion through the fully convolutional structure. The branch is composed of two segments involving an encoder to extract valuable features from photometric images and a decoder to incorporate extracted features for reflection variation generation. The encoder and decoder are composed of convolutional and residual blocks for multiple fusions with various resolutions, and spatial attention (Sa) layers are introduced in the residual blocks to focus on worn features.

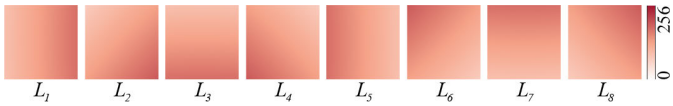


Fig. 4. Lambertian reflections of the plane base with multiple illumination sources. (Note that $\boldsymbol{L}_1$–$\boldsymbol{L}_8$ represent individually single illumination source along the clockwise direction.)

Considering the complex reflectance properties of worn surfaces, residual blocks [19] are introduced to address the accuracy degradation from over-deep network structures. Specifically, the module constructs an identical mapping by adding the shortcut path. For the brightness differences resulting from local height variations, the Sa layer [20] is embedded in residual blocks to enhance worn features by emphasizing interchannel variations in feature maps. The workflow of this layer consists of two steps. The maximum and average values of feature maps are first extracted along the channel dimension. This is followed by the difference between both the values that are regarded as weights and are further multiplied with the feature map pixel by pixel. With the weight assignment, the residual block can provide rapid localization and full characterization of worn features.

The optimized residual block is treated as basic construction components to build the variation estimation branch. Given the different image sampling demands of the encoder and the decoder, the residual block is derived into the down-residual and up-residual blocks. Down-residual blocks are used to fuse the neighborhood luminance information for local features by a convolutional layer with a stride of 2 and a kernel size of 3×3. By comparison, up-residual blocks progressively refine reflection feature maps by deconvolution layers. Furthermore, three residual blocks with the same structure and varied channels are connected in series with a convolution block to form the encoder and the decoder. Both the components are integrated to compose the variation estimation branch. With the above structures, the branch implements a multistep fusion of photometric images via multiple residual blocks and then enables the robust estimation of Lambertian reflection variations.

2) Constant Extraction Branch: A constant extraction branch is constructed to extract Lambertian reflection constants as a priori knowledge for the reflection separation task. With known surface topography, the reflection constants are determined by the illumination properties of the imaging system. Given that illumination sources are close to worn surfaces, the point light source model [3] is selected to describe the direction of incident lights. With multiple illumination sources in a circular distribution, Lambertian reflections of the plane base are shown in Fig. 4. As can be observed, the Lambertian reflection constants exhibit intensity gradients that vary with different illumination sources.

The extracted reflection constant is used to optimize the Lambertian reflection separation of worn surfaces. The reflection constants are summed together with the estimated reflection variations toward the generation of whole Lambertian

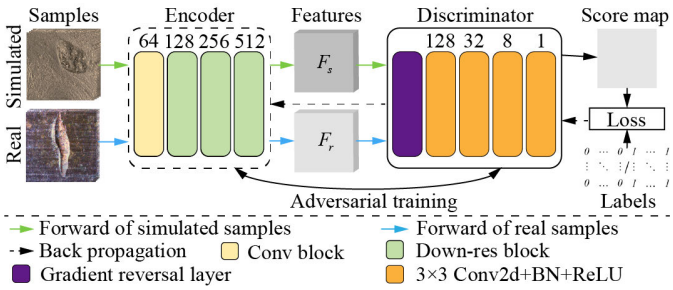


Fig. 5. Feature alignment for simulated and real samples (note the numbers of channels marked on the corresponding convolutional layers).

reflections. With the introduction of such a priori knowledge, the Lambertian reflection separation is focused on critical worn features rather than the whole measurement area. Therefore, the constant extraction branch can simplify the reflection separation task and then contributes toward improving the prediction accuracy.

B. Parameter Optimization

The LRS-Net model requires sufficient image samples to optimize its parameters. However, this may be constrained by not having enough samples due to difficulties in the inspection operation of worn surfaces. Rendering-based simulation methods can generate massive simulated samples, but significant deviations exist from real photometric images. Therefore, an FA is introduced to combine real and simulated samples for parameter optimization.

An encoder is used in the LRS-Net model to extract valuable features from photometric images. The validity of the feature extraction directly affects the applicability of reflection separation to various samples. Hence, the extracted features from simulated and real samples are aligned to jointly optimize the parameters. As shown in Fig. 5, the FA consists of three steps. In the first instance, the valuable features of simulated and real samples are extracted from photometric images with the encoder. This is followed by a domain discriminator to determine the sample categories from extracted features. Finally, the encoder and the domain discriminator are trained with the opposite gradient of the domain discrimination loss. Particularly, the domain discriminator is optimized to obtain a clear domain classification. In contrast, the encoder is inclined to extract features with the same distribution from simulated and real samples.

In the FA, the domain discriminator is used to identify sample categories from extracted features. Considering the spatial differences in extracted features, the discrimination result is represented by a score map rather than a single probability value. Hence, the domain discriminator is composed of only four convolutional layers with a stride of 1. Moreover, the gradient reversal layer [21] is embedded ahead of convolutional layers in the adversarial training for feature extraction and domain identification.

C. Loss Function

An appropriate loss function can provide the correct direction for parameter optimization. For the LRS-Net model, the optimization aims at deriving consistent features and accurate Lambertian reflections for simulated and real samples. Given this, reflection prediction loss and domain discrimination loss are constructed and integrated into training the LRS-Net model.

1) Reflection Prediction Loss for Real Samples: The loss functions are generally constructed by comparing the prediction results and the corresponding labels. However, the labels are absent for real samples since the Lambertian reflection labeling involves complicated topography measurements. To address the issue, a loss function is proposed to evaluate the prediction errors according to the Lambertian reflection characteristics.

The Lambertian reflections of photometric images are analyzed for label-free quantitative evaluation. As given in (1), the Lambertian reflections from different photometric images are satisfied with the sine function. Combined with the circumferential distribution of illumination sources, the function is further simplified to (2). It means that the sine function fit from Lambertian reflections has an equivalent angular frequency, $\boldsymbol{\omega} = 2\pi/n$

$$I^k = \alpha \sin\left(\frac{2\pi}{n} \times k + \beta\right) + \gamma \tag{2}$$

where n is the number of illumination sources in the imaging system, k is the serial number of the illumination source, and β and γ are the bias terms.

The idea of fixed angular frequency is used to evaluate the Lambertian reflection separation for real samples. Specifically, the reflection prediction results at the same pixels are ordered to form a sequence and further solved for the amplitude spectrum by Fourier transform. Based on the amplitude distribution, a label-free loss function is constructed to quantify the reflection separation errors, as illustrated in (3). The loss function is equal to zero for accurate Lambertian reflections, while larger loss function values mean larger prediction errors. The evaluation is independent of the label references and thus able to be extended to real samples without labels for model optimization

$$\text{loss}_r = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} \left(1 - \frac{A_{ij}(\omega = 2\pi/n)}{\sum A_{ij}(\omega \neq 0)}\right) \tag{3}$$

where $H \cdot W$ denotes the resolution of photometric images, $A_{ij}(\omega = \omega_0)$ is the amplitude with the angular frequency $\omega = \omega_0$ of the sequence $L_{ij} = [I_{ij}^1\ I_{ij}^2, \ldots, I_{ij}^n]$, and I_{ij}^k is the predicted value of the Lambertian reflection intensity at the point (i, j) in the kth image.

2) Reflection Prediction Loss for Simulated Samples: In contrast to real samples, simulated samples are effortlessly assigned reflection labels by the determined surface topography and Lambert reflectance model. Hence, the reflection prediction loss of simulated samples is defined by the comparison of the prediction results with the corresponding labels. Given this, the L_1 loss function is selected to quantify the prediction error of Lambertian reflections pixel by pixel, as shown in the

following equation:

$$\text{loss}_s = \frac{1}{n \cdot H \cdot W} \sum_{k=1}^{n} \sum_{i=1}^{H} \sum_{j=1}^{W} |I_{ij}^k - \tilde{I}_{ij}^k| \tag{4}$$

where $\tilde{I}_{ij}^k$ is true value of Lambertian reflection intensity at the point (i, j) in the kth image.

3) Domain Discrimination Loss for Sample Categories: In FA, the domain discriminator is constructed to discriminate whether the extracted features are from real or simulated samples. For the discrimination, binary cross-entropy is selected as the domain discrimination loss and quantifies the difference between the discrimination results and domain labels, as shown in the following equation:

$$\text{loss}_d = -\frac{1}{H_s \cdot W_s} \sum_{i,j}^{H,W} y \log\left(S_r^{i,j}\right) + (1 - y) \log\left(S_s^{i,j}\right) \tag{5}$$

where $H_s \cdot W_s$ is the resolution of the score map, y denotes the domain label, and labels of simulated and real samples are set to 0 and 1, respectively, $S_r^{i,j}$ is the discrimination result at the point (i, j) in the score map for real samples, and $S_s^{i,j}$ is the discrimination result at the point (i, j) in the score map for simulated samples.

4) Complete Loss Function: The complete loss function is acquired by the weighted sum of the two reflection prediction losses and the domain discrimination loss, as defined in the following equation:

$$\text{loss} = \lambda_r \cdot \text{loss}_r + \lambda_s \cdot \text{loss}_s + \lambda_d \cdot \text{loss}_d \tag{6}$$

where λ_r is the coefficient of loss_r, λ_s is the coefficient of loss_s, and λ_d is the coefficient of loss_d.

III. Model Training

A. Training Dataset Acquisition

For the LRS-Net model, both real and simulated samples of worn surfaces are integrated for parameter optimization. Two types of samples are collected by experiment and simulation, respectively.

1) Real Sample Collection: Since the Lambertian reflection labeling entails laborious topography measurement, real samples only contain photometric images of worn surfaces. The samples are collected from actual friction pairs via an in situ imaging system, shown in Fig. 6(a). As the crucial component of the system, the illumination sources are composed of eight LEDs distributed circumferentially and equidistantly. By sequentially switching on/off LEDs, photometric images are captured with different illumination directions, as shown in Fig. 6(b). The acquisition process is repeated to acquire 1000 sets of real samples from pin-on-disks and bearings, which have experienced a long-term wear process with the pin-on-disk tribometer and the roller-ring test rig. These real samples are used for parameter optimization to ensure that the LRS-Net model can adapt to the imaging properties and morphological characteristics of worn surfaces.

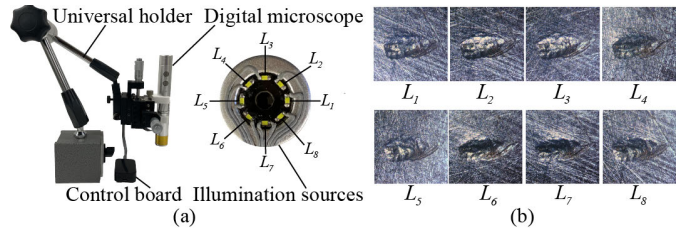


Fig. 6. Real sample collection from worn surfaces. (a) In situ imaging system. (b) Photometric images.

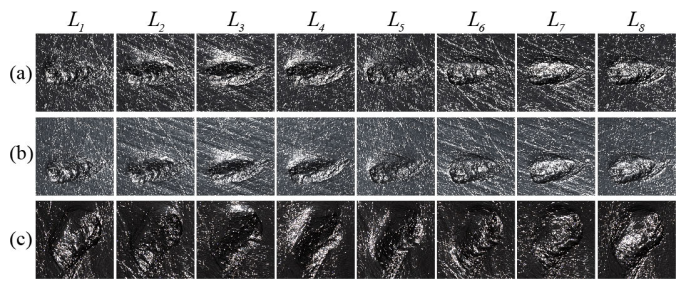


Fig. 7. Photometric images rendered with (a) grease-covered steel, (b) aluminum, and (c) chrome-steel.

2) Simulated Sample Generation: Given the small number and absence of labels in real samples, simulated samples are generated for parameter optimization of the LRS-Net model. As required for supervised training (ST), the simulated samples are composed of photometric images and the corresponding Lambertian reflections.

For vision-based reconstruction, the high-quality simulated samples are defined from two aspects. The first is surface topography with worn features, such as scratches and pits. This is followed by the high reflectiveness of metallic materials. Based on this, the simulation method is developed by expanding existing reflectance data to worn surface topographies. First, with the help of laser scanning confocal microscopy (LSCM, OLS4000, Japan), 100 sets of worn topography with multiple worn features are selected and collected from the bearing surfaces. Second, acquired topographies are augmented with affine transformation and node-based shape transformation to relieve the time-consuming surface measurement [22]. Via the augmentation method, generated topographies are characterized by various typical features similar to the worn surface but with varied sizes. Third, photometric images are generated by rendering augmented surface topographies with the Mitsubishi Electric Research Laboratories (MERL) database containing 100 materials [15]. By adopting various surface topographies and rendering materials, 20 000 sets of photometric images have been generated. As depicted in Fig. 7, generated surfaces are presented with morphological features consistent with real worn surfaces. Moreover, the photometric images exhibit typical reflection components of metallic materials, such as local highlights and shadows. Furthermore, the reflective properties of photometric images are varied with the rendered material to simulate worn surfaces of diverse materials.

Furthermore, the Lambertian reflections of simulated surfaces are obtained by image rendering. Specifically, surface topographies obtained in the simulation are fed into the

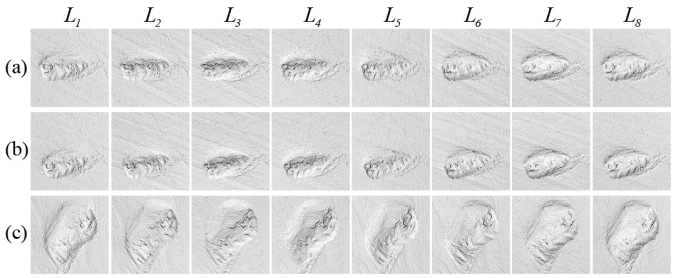


Fig. 8. Matched Lambertian reflections with photometric images in Fig. 7.

TABLE I

MEAN PREDICTION ERRORS OF THE VALIDATION SET WITH DIFFERENT COEFFICIENT COMBINATIONS

λ_r	λ_s	λ_d	Mean prediction errors
1	0.1	0.5	7.12
	0.2	0.5	7.04
	0.5	0.5	7.10
	0.2	0.1	7.11
	0.2	1.0	7.09

ideal Lambert reflectance model to generate the corresponding Lambertian reflections. Fig. 8 depicts the matched Lambertian reflections of the photometric images in Fig. 7. As may be observed, local highlights and shadows in the photometric images are replaced by moderate brightness. Furthermore, the Lambertian reflections can provide detailed topography information by interpixel and interimage brightness variations.

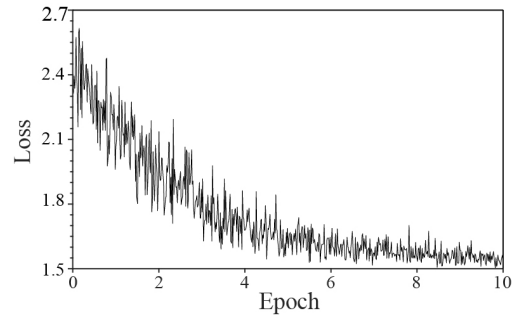


Fig. 9. Adam-based training process of the constructed LRS-Net model.

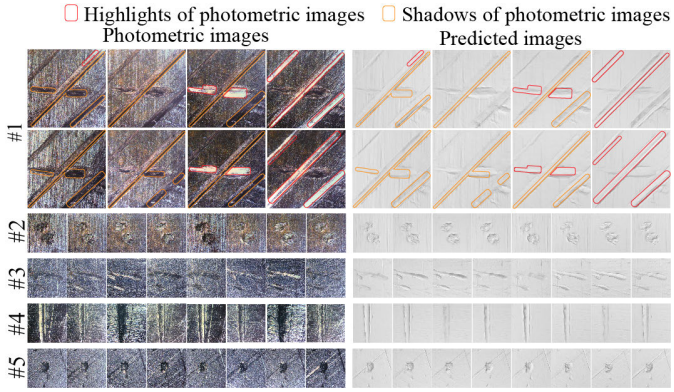


Fig. 10. Prediction results of Lambertian reflections for worn surfaces with different materials.

B. Training Details

For the LRS-Net model, network training is an imperative parameter optimization process for the Lambertian reflection separation of worn surfaces. The Adam algorithm [23] is selected as the training optimizer. The learning rate is first set to 0.001 during initial learning and then divided by two for every two epochs to enable stable convergence of network parameters in the last half of training. For the multiobjective loss function, the coefficient λ_r is set to 1 due to the priority of real samples, and then the values of the remaining two coefficients are adjusted for optimal performance. With 20 sets of real samples as the validation set, different coefficient combinations are compared, and the mean prediction errors are shown in Table I. According to the comparison results, the coefficient combination $[\lambda_s = 0.2, \lambda_d = 0.5]$ with the smallest error is selected for model training. With the above hyperparameters, the training is implemented with 100 epochs in the PyTorch platform [24]. The whole training process is illustrated in Fig. 9. As may be observed, the model loss is gradually becoming stabilized and converged, indicating that the LRS-Net model has been fully trained with the real and simulated samples.

IV. VERIFICATION AND ANALYSIS

A. Verification of the LRS-Net Model

The LRS-Net model is constructed to separate Lambertian reflections for accurate normal vector calculation of worn surfaces. Therefore, the model is evaluated from Lambertian reflections and normal vectors. To ensure a comprehensive evaluation, test samples with different worn features are selected, as shown in Fig. 10. These samples are characterized by scratches and pits in different feature scales, locations, and orientations to represent the morphological characteristics of worn surfaces. With the trained LRS-Net model, Lambertian reflections are separated from photometric images of test surfaces, as illustrated in Fig. 10. The prediction results of Lambertian reflections are characterized by continuous brightness variations in the highlight and shadow regions of photometric images. It indicates that the LRS-Net model can effectively separate Lambertian reflections for worn surfaces with high reflectiveness.

The LRS-Net model is validated by calculating the normal vectors from the separation results of Lambertian reflections. The calculation is based on least squares [5]. The ground truth of normal vectors is obtained by the topography measurement with LSCM and the height–gradient normal transformation. The results and error distribution of normal vectors are shown in Fig. 11. As may be observed, the calculated normal vectors are highly consistent with the ground truth except for feature edges and some minor areas. The differences may be attributed to the restricted characterization of the digital microscope at tiny features.

Furthermore, the normal vectors of the LRS-Net model are compared with LS [5], L_1 [25], robust principal component analysis (R-PCA) [26], constrained bivariate regression (CBR) [27], and NE-Net [15]. For the quantitative comparison, the mean angular error (MAE) is used to evaluate the calculation errors of normal vectors given in (7). As shown in Table II, the normal vector calculation with the LRS-Net model exhibits a significantly higher solution accuracy than the existing algorithms. It further indicates that the constructed

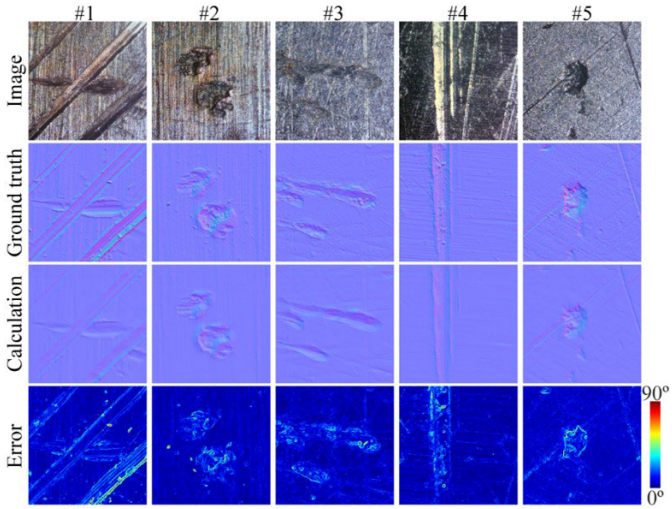


Fig. 11. Calculated normal vector map of worn surfaces. (Note: the color bar at right indicates angular error corresponding to specific colors.)

TABLE II
MAEs OF DIFFERENT ALGORITHMS FOR WORN SURFACES

Algorithms	#1	#2	#3	#4	#5
LS	9.94°	12.64°	12.47°	12.96°	12.20°
L1	10.21°	13.21°	12.91°	13.69°	12.55°
SBL	8.49°	11.55°	10.02°	10.12°	10.33°
R-PCA	11.27°	10.65°	7.14°	7.50°	8.09°
CBR	6.36°	9.62°	6.90°	7.11°	7.71°
PS-FCN	8.24°	10.97°	7.55°	7.91°	8.65°
LRS-Net	5.95°	9.28°	6.55°	6.06°	7.41°

LRS-Net model can achieve an accurate Lambertian reflection separation for worn surfaces

$$\mathrm{MAE} = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} \arccos\left(\boldsymbol{N}_{ij} \cdot \tilde{\boldsymbol{N}}_{ij}\right) \tag{7}$$

where N_{ij} and $\tilde{N}_{ij}$ are the calculated and true values of the normal vector at the point at the point (i, j), respectively.

B. Ablation Studies for the Model Architecture

The reflection decomposition guides the LRS-Net model to focus on worn features by a priori knowledge of Lambertian reflection constants. To verify its validity, an ablation study is conducted by analyzing the training process. For comparison, a contrast model is constructed by removing the reflection decomposition of the LRS-Net model. Specifically, the constant extraction branch is removed, and the variation estimation branch directly outputs the Lambertian reflection from photometric images. The prediction results of Lambertian reflections and the reflection prediction losses loss_s in different epochs are shown in Fig. 12. Compared with the contrast model, the LRS-Net model consistently focuses on worn features and obtains smaller prediction errors for training samples. Furthermore, the prediction errors are quantified with the L_1 function, and the mean and maximum values of various test samples are combined to evaluate the discussed models comprehensively. The comparison results for test samples are shown in Table III. It may be observed that the LRS-Net model presents higher prediction accuracy for worn surfaces. Therefore, the reflection decomposition facilitates the convergence of the LRS-Net model and then improves the reflection prediction accuracy under insufficient samples.

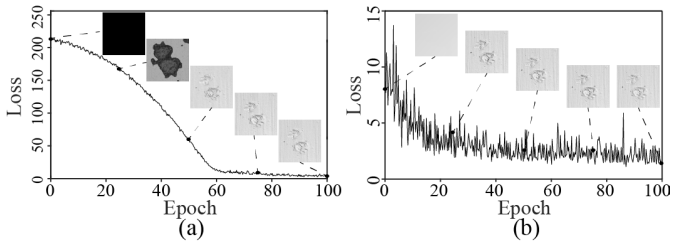


Fig. 12. Prediction results of Lambertian reflections and reflection prediction losses of (a) contrast model and (b) LRS-Net model.

TABLE III
REFLECTION PREDICTION ERRORS OF THE LRS-NET MODEL AND THE CONTRAST MODEL

Models	Mean	Maximum
LRS-Net	7.08	10.34
Contrast model	8.52	11.41

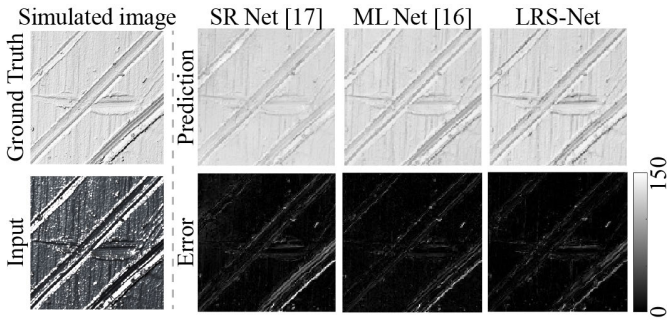


Fig. 13. Mean prediction errors of simulated test samples with different reflection separation models.

TABLE IV
MEAN PREDICTION ERRORS OF SIMULATED TEST SAMPLES WITH DIFFERENT REFLECTION SEPARATION MODELS

Models	Mean	Maximum
SR Net	5.93	13.46
ML Net	5.54	12.00
LRS-Net	4.10	11.76

Furthermore, the LRS-Net model is compared with the existing reflection separation models. To eliminate the effects of the domain gap, 500 sets of simulated images with 100 types of materials are selected as the test set. The prediction results of different models are illustrated in Fig. 13, and the L_1 mean prediction errors are shown in Table IV. It may be observed that the LRS-Net model exhibits smaller prediction errors than SR Net [16] and ML Net [17], especially in the local highlight regions.

C. Ablation Study for the Optimization Approaches

The FA is used to enable the Lambertian reflection separation for worn surfaces by the constraint on the feature extraction. To verify its role, a corresponding ablation study is carried out by comparing extracted features from simulated

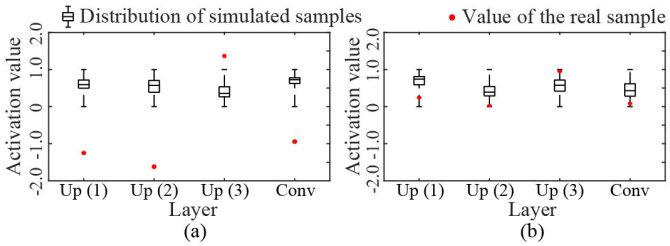


Fig. 14. Activation values of the LRS-Net model with (a) ST and (b) combined ST and FA. (Note that up-residual blocks are named Up(1), Up(2), and Up(3) along the forward propagation direction.)

TABLE V

REFLECTION PREDICTION ERRORS OF THE LRS-NET MODEL WITH VARIOUS OPTIMIZATION APPROACHES

Optimization approach	Mean	Maximum
ST	10.88	13.06
ST+FAl	7.48	10.62
ST+LFL	7.73	10.65
ST+FA+LFL	7.08	10.34

TABLE VI

MEAN PREDICTION ERRORS OF REAL TEST SAMPLES WITH DIFFERENT SEMI-SUPERVISED METHODS

Optimization approach	Mean	Maximum
CF-Tran [29]	9.59	13.29
GAN [30]	7.23	10.62
DANN [31]	7.43	11.19
LRS-Net	7.08	10.34

and real samples. Since extracted features are used in the subsequent reflection variation regression, the comparison is conducted by the activation values of the convolutional layers in the decoder [28]. In the ablation study, the contrast model adopts ST with 20 000 simulated samples. Moreover, 100 simulated test samples are generated with MERL reflectance data and identical topography as the #1 real sample, and the involved samples are pooled together as the test set. The activation values of the test samples are shown in Fig. 14. As may be observed, the contrast model is characterized by distinct activation values for simulated and real samples. In contrast, two domain samples are confused according to activation values of the LRS-Net model. The consistent response indicates that the domain gap between simulated and real samples has been closed.

Besides the FA, the label-free prediction loss (LFL) is constructed for real samples to constrain parameter optimization. The two optimization approaches are validated by reflection prediction errors of real test samples, and the comparison results are shown in Table V. It may be observed that the two optimization approaches contribute to the accuracy improvement in Lambertian reflection separation and combine for optimal prediction results. Furthermore, the proposed optimization method is compared with the existing semi-supervised methods via real test samples. The comparison results are shown in Table VI. It may be observed that the combination of FA and LFL presents higher prediction accuracy on worn surfaces than the existing methods.

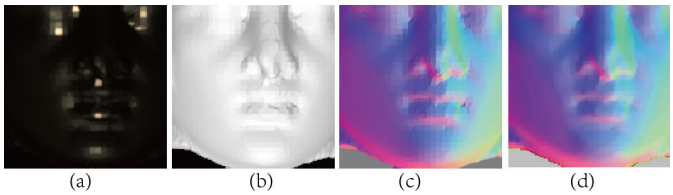


Fig. 15. Test sculpture samples including (a) photometric image, (b) prediction results of Lambertian reflections, (c) calculation results of normal vectors, and (d) ground truth of normal vectors.

TABLE VII

MAEs OF DIFFERENT ALGORITHMS FOR SCULPTURE SAMPLES

Algorithms	MAEs
LS	12.90°
L1	12.72°
R-PCA	10.01°
LRS-Net	8.55°

D. Evaluation of Generalization Performance

The LRS-Net model is further evaluated on other datasets besides worn surfaces. Considering the specific distribution of illumination sources, the evaluation is conducted with rendered images from shape datasets rather than the existing photometric image datasets. Specifically, with the MERL database, the blobby shape dataset [32] and the sculpture shape dataset [33] are rendered to generate photometric images, namely, blobby and sculpture samples. Furthermore, the blobby samples are labeled and combined with sculpture samples to optimize the model parameters.

With 100 sets of sculpture samples as the test dataset, the optimized LRS-Net model is evaluated, and the predicted Lambertian reflections and calculated normal vectors are shown in Fig. 15. As may be observed, the calculated normal vectors are highly identical with the ground truth. Furthermore, the LRS-Net model is compared with LS [5], L_1 [25], and R-PCA [26], and the MAEs of various algorithms are shown in Table VII. It may be observed that the LRS-Net model exhibits more minor calculation errors for sculpture samples than the existing algorithms.

V. DISCUSSION

Lambertian reflection separation is achieved by an LRS-Net for the calculation of worn surface normal vector. With multiple photometric images as input, reflection separation is implemented by integrating the variation estimation and constant extraction. Moreover, FA and the label-free loss function are introduced to combine simulated and real samples for model optimization. The LRS-Net model is evaluated on the prediction results of Lambertian reflections and normal vectors, and its framework and optimization approaches are further verified with ablation studies. A comparison between the constructed model and existing reflection separation algorithms is described below.

The existing outlier removal approaches select Lambertian pixels to reconstruct surface topography. However, they may fail due to inadequate valid pixels, especially for high-reflectiveness surfaces with sparse photometric images.

In contrast, the LRS-Net model can estimate Lambertian reflections of highlight and shadow regions to ensure reliable topography reconstruction. Compared with the existing image translation networks for reflection separation [16], [17], the LRS-Net model enables robust prediction for highlight regions by fusing multiple photometric images. Moreover, Lambertian reflection decomposition contributes to the focus of the LRS-Net model on worn regions and thus improves the reflection prediction accuracy. Although sample simulation methods have been developed for parameter optimization, the differences between simulated and real samples limit the application of the learning-based model on worn surfaces [15]. Compared with the existing semi-supervised methods [30], [31], the LRS-Net model adopts the FA and the label-free loss function to constrain the encoding and decoding jointly and thus achieves higher prediction accuracy on worn surfaces. Hence, the proposed model can effectively separate Lambertian reflections and then contributes to improving normal vector calculations.

VI. Conclusion

An LRS-Net model for Lambertian reflection separation is developed and used for the photometric stereo reconstruction of worn surfaces. Its properties include: 1) for the high reflectiveness of worn surfaces, reflection variations are estimated by fusing multiple photometric images and then integrated with reflection constants extracted from the plane base to enable robust reflection prediction; 2) FA is used to combine simulated and real samples for parameter optimization, and hence, the reflection separation can be extended to worn surfaces with insufficient samples; and 3) according to the Lambertian reflection characteristics, a label-free loss function is constructed to constrain parameter optimization via real samples and then further improve the reflection prediction accuracy on real surfaces. With the Lambertian reflection separation, the accuracy in normal vector calculation is improved, and the MAEs of the results are less than 10°.

References

[1] Z. Liu and L. Zhang, "A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings," *Measurement*, vol. 149, Jan. 2020, Art. no. 107002.

[2] M.-X. Shen et al., "Effect of particle size on tribological properties of rubber/steel seal pairs under contaminated water lubrication conditions," *Tribol. Lett.*, vol. 68, no. 1, pp. 1–15, Mar. 2020.

[3] C. Xu, T. Wu, Y. Huo, and H. Yang, "In-situ characterization of three dimensional worn surface under sliding-rolling contact," *Wear*, vols. 426–427, pp. 1781–1787, Apr. 2019.

[4] V. Argyriou and M. Petrou, "Photometric stereo: An overview," in *Advances in Imaging and Electron Physics*, vol. 156. Amsterdam, The Netherlands: Elsevier, 2009, ch. 1, pp. 1–54.

[5] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.*, vol. 19, no. 1, Feb. 1980, Art. no. 191139.

[6] J. Ren, Z. Jian, X. Wang, M. Ren, L. Zhu, and X. Jiang, "Complex surface reconstruction based on fusion of surface normals and sparse depth measurement," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.

[7] X. Fang, Q. Luo, B. Zhou, C. Li, and L. Tian, "Research progress of automated visual surface defect detection for industrial metal planar materials," *Sensors*, vol. 20, no. 18, pp. 1–35, 2020.

[8] A. S. Georghiades, "Incorporating the Torrance and sparrow model of reflectance in uncalibrated photometric stereo," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, Oct. 2003, pp. 816–823.

[9] T. Zickler, P. Belhumeur, and D. Kriegman, "Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction," in *Proc. 7th Eur. Conf. Comput. Vis.*, May 2002, pp. 869–884.

[10] D. Miyazaki, K. Hara, and K. Ikeuchi, "Median photometric stereo as applied to the Segonko Tumulus and museum objects," *Int. J. Comput. Vis.*, vol. 86, nos. 2–3, pp. 229–242, Jan. 2010.

[11] Y. Mukaigawa, Y. Ishii, and T. Shakunaga, "Classification of photometric factors based on photometric linearization," in *Proc. 7th Asian Conf. Comput. Vis.*, Jan. 2006, pp. 613–622.

[12] H. Liu, Y. Yan, K. Song, and H. Yu, "SPS-Net: Self-attention photometric stereo network," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 5006213.

[13] Y. Ju, M. Jian, S. Guo, Y. Wang, H. Zhou, and J. Dong, "Incorporating Lambertian priors into surface normals measurement," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 5012913.

[14] G. Chen, K. Han, and K.-Y. K. Wong, "PS-FCN: A flexible learning framework for photometric stereo," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, vol. 11213, 2018, pp. 3–19.

[15] Q. Wang, S. Wang, B. Li, K. Zhu, and T. Wu, "In-situ 3D reconstruction of worn surface topography via optimized photometric stereo," *Measurement*, vol. 190, Feb. 2022, Art. no. 110679.

[16] J. Lin, M. E. A. Seddik, M. Tamaazousti, Y. Tamaazousti, and A. Bartoli, "Deep multi-class adversarial specularity removal," in *Image Analysis: 21st Scandinavian Conference, SCIA 2019, Norrköping, Sweden, June 11–13, 2019, Proceedings*. Berlin, Germany: Springer-Verlag, pp. 3–15.

[17] J. Shi, Y. Dong, H. Su, and S. X. Yu, "Learning non-Lambertian object intrinsics across ShapeNet categories," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5844–5853.

[18] F. Azhar, K. Emrith, S. Pollard, M. Smith, G. Adams, and S. Simske, "Testing the validity of Lamberts law for micro-scale photometric stereo applied to paper substrates," in *Proc. 10th Int. Conf. Comput. Vis. Theory Appl. (VISAPP)*, 2015, pp. 246–253.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[20] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.

[21] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," in *Proc. 32nd Int. Conf. Mach. Learn.*, vol. 37, Jul. 2015, pp. 1180–1189.

[22] K. A. J. Eppenhof and J. P. W. Pluim, "Pulmonary CT registration through supervised learning with convolutional neural networks," *IEEE Trans. Med. Imag.*, vol. 38, no. 5, pp. 1097–1105, May 2019.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015, pp. 1–15.

[24] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–12.

[25] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Robust photometric stereo using sparse regression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 318–325.

[26] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, "Robust photometric stereo via low-rank matrix completion and recovery," in *Proc. Asian Conf. Comput. Vis.*, 2010, pp. 703–717.

[27] S. Ikehata and K. Aizawa, "Photometric stereo using constrained bivariate regression for general isotropic surfaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2187–2194.

[28] A. Nguyen, J. Yosinski, and J. Clune, "Understanding neural networks via feature visualization: A survey," in *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning* (Lecture Notes in Computer Science), vol. 11700. Cham, Switzerland: Springer, 2019, pp. 55–76.

[29] J. Chen et al., "A transfer learning based super-resolution microscopy for biopsy slice images: The joint methods perspective," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 1, pp. 103–113, Jan. 2021.

[30] W. Hung, Y. Tsai, Y. Liou, Y. Lin, and M. Yang, "Adversarial learning for semi-supervised semantic segmentation," in *Proc. 29th Brit. Mach. Vis. Conf.*, Feb. 2018, pp. 1–12.

[31] Y. Ganin et al., "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 1–35, Jan. 2016.

[32] M. K. Johnson and E. H. Adelson, "Shape estimation in natural illumination," in *Proc. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 2553–2560.

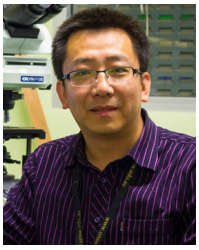
[33] O. Wiles and A. Zisserman, "SilNet: Single-and multi-view reconstruction by learning from silhouettes," in *Proc. Brit. Mach. Vis. Conf.*, 2017, pp. 1–13.

Qinghua Wang received the bachelor's degree in mechanical design manufacture and automation from the Hefei University of Technology, Hefei, China, in 2019. He is currently pursuing the Ph.D. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China.

His research interests include worn topography reconstruction and wear evolution analysis.

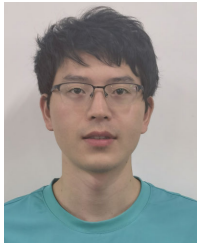
Shuo Wang (Member, IEEE) received the Ph.D. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2021.

He is currently an Assistant Professor with the School of Mechanical Engineering, Xi'an Jiaotong University. His research interests include machine condition monitoring, wear debris analysis, and micro 3-D surface analysis.



Tonghai Wu received the Ph.D. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2006.

From 2013 to 2014, he was a Visiting Scholar with the University of New South Wales, Sydney, NSW, Australia. He is currently a Professor with the School of Mechanical Engineering, Xi'an Jiaotong University. His current research interests include machinery diagnosis and prognosis, machine vision on wear particle, and surface analysis.



Tao Shao received the M.S. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2018, where he is currently pursuing the Ph.D. degree, under the supervision of Prof. Tonghai Wu.

His research interests include worn topography reconstruction, wear evolution analysis, and uncertainty quantification.

Yue Shu received the Ph.D. degree from the University of Science and Technology of China, Hefei, China, in 2014.

He is currently a Senior Engineer with the State Key Laboratory of Compressor Technology, Hefei General Machinery Research Institute Company Ltd., Hefei. His main research interests include compressor design and manufacturing, monitoring and diagnosis, and remote operation and maintenance.

Thompson Sarkodie-Gyan (Member, IEEE) received the Dr.-Ing. degree in measurement and control from Technical University, Berlin, Germany, in 1985.

He is currently an Associate Professor with the Department of Electrical and Computer Engineering, The University of Texas at El Paso, El Paso, TX, USA. His current research interests include biomedical engineering, mechatronics systems engineering, robotics and automation, electrical measurements of nonelectrical quantities, machine vision and pattern recognition, applications of soft computing, and sensors.