Detection Method of Concave Defect on Specular Surfaces Based on Swin Transformer

Manuscript received October 24, 2024; revised November 18, 2024

Kazumoto Tanaka
Department of Informatics
Kindai University
Higashi-Hiroshima, Japan
kazumoto@hiro.kindai.ac.jp

***Abstract*— Shallow concave defects on mirrored surfaces are difficult to detect automatically. This paper proposes a defect detection method using a deep neural network (DNN) that learns the presence or absence of distortion in the image of a stripe pattern reflected on a mirror surface. The Swin Transformer is used as the DNN to capture global features of the edges of the reflection. In the manufacturing process, the occurrence of defects is minimized, so it is difficult to collect enough defect images for training purposes. Therefore, in this paper, we show how to generate a large number of images of stripe pattern reflections using an optical simulation method. Our Swin Transformer showed high detection performance in defect detection experiments using actual mirrored parts.**

Keywords*— *automated surface inspection, concavity defect, optical simulation, specular surface, swin transformer

I. INTRODUCTION

Automation of visual inspection of industrial products has been studied with the aim of reducing the workload of workers and improving quality. There have been many studies on image processing approaches to automation [1]. However, there are still many defects that are difficult to detect automatically and shallow concave defects on mirror surfaces are one of them. This paper proposes a newly developed simple and useful method for detecting such defects.

In the visual inspection of concave defects on mirror surfaces performed in the production process, the presence or absence of image distortion of the reflection of a stripe pattern light source is conventionally examined [2] (Fig. 1). Accordingly, the image processing approach by using stripe pattern reflection has been the most widely used approach in automated inspection research. In a study [3], a method for evaluating the density variance in multiple images taken while shifting the phase of the stripe pattern was proposed for inspecting the surface roughness of a glossy surface. In a study [4], which was aimed at inspecting scratches on a glossy surface, scratches were detected by comparing the image density distribution of the reflection with the normal image density distribution. Study [5] proposed a method to remove superfluous reflections other than stripe patterns in mirror inspection using reflections.

On the other hand, there is a technique called the phase-shift method, in which structured light with reduced stripe spacing is projected onto an object to measure the three-dimensional shape of its surface [6], and there is a study that has applied this method to the inspection of surface irregularities [7]. In this study, defects were detected by comparing the measured 3D data with normal shape data. However, the experiments were conducted on non-glossy surfaces, which is problematic for glossy surfaces because the phase-shift method is not applicable to specular surfaces. This is because the reflective properties of specular surfaces must be taken into account when applying the phase-shift method to specular objects. In addition, there is the phase-measuring deflectometry (PMD) method, which uses the reflective properties of the specular surface[8]. Studies have also been conducted to inspect mirror surface defects using this PMD method [9]-[12]. Study [9] proposed a method to extract the specular reflection components necessary for PMD from an image using a deep neural network (DNN). A study [10] applied a DNN for object recognition to 3D images obtained by PMD for defect detection. A method for surface inspection using light-intensity contrast maps, which can be calculated relatively easily from the phase-shift algorithm, instead of curvature map images, which require much processing time in PMD, has also been proposed [11],[12]. However, PMD has the fundamental problem that it requires highly accurate camera calibration and precise positioning of the object [13], making it difficult to implement in a manufacturing environment.

This study automates the conventional method of checking the image distortion of stripe pattern reflections in visual inspection. Most previous studies of this approach (e.g., [3]-[5]) have detected defects based on perturbations in image density patterns on flat mirror surfaces. However, when the object is a curved surface, the image is distorted even on a defect-free surface, making it extremely difficult to determine whether the image is defective or not. However, experienced visual inspectors can determine whether the distortion is due to curvature alone or to both curvature and concavity defects based on their knowledge gained from many years of experience. Inspired by this fact, our study addresses the problem by using a DNN that learns the features of stripe pattern reflections on both defect-free surfaces and surfaces with concave defects. A Swin Transformer was used as the DNN to capture the global features of the stripe pattern reflections. However, in general, defect images for training the DNN cannot be collected sufficiently because efforts are made to minimize the occurrence of defects in the manufacturing process. In this study, we used an optical simulation method to generate a large number of stripe pattern reflection images. This method allows the inspection of any curved surface. The proposed method was experimentally evaluated using actual industrial products with mirrored surfaces, and its high defect detection performance was confirmed.

The main contributions of this paper are as follows:

 EISSN: 2961-872X ● PISSN: 2961-8924 ● DOI:10.56741/IISTR.esl.00743

- To propose a simple method for detecting shallow concave defects on mirror surfaces with high accuracy using only the reflection image of the stripe pattern.
- To show that the Swin Transformer can detect anomalous edge positioning of stripe patterns with high accuracy.
- To propose a highly versatile method because it can generate training images no matter what kind of curved surface is being inspected by using optical simulation.

The remainder of this paper is organized as follows. Section 2 describes the proposed method and Section 3 presents the experimental evaluation. Section 4 provides a discussion, and Section 5 concludes the paper.



Fig. 1. Stripe pattern reflected on a glossy and elongated curved surface. A stripe pattern light source is presented at the back.

II. METHOD

In this section, we describe the configuration of the proposed DNN and the training method of the DNN.

A. Swin Transformer for Concavity Defect Detection

Visual inspections performed by skilled inspectors focus on the spatial location of the stripe edges (Fig. 2) in the reflected image of the stripe pattern. To utilize such spatially remote and global information, we use the Swin Transformer [14] as our DNN framework. The Swin Transformer is a transformer encoder [15]-based image recognition model that can extract global features of an image at various scales using a shifted window-based self-attention mechanism.

The input image to our Swin Transformer is not the reflected stripe pattern image itself, but an image of the edges extracted from the reflected stripe pattern image. The reasons for this are as follows. As described in Section 1, the training images for the Swin Transformer are generated by simulation. However, it is very difficult to accurately simulate the density distribution of the reflected image and the background of the object to be inspected in an actual inspection. However, edge images are binary images, so it has the advantage of simplicity in that only edge shapes need to be simulated. In actual inspection images, if the background is darkened and the image is binarized, only the reflected areas can be extracted and the edges can be detected. Therefore, the Swin Transformer trained on simulated edge images is expected to work equally well on real edge images. However, the edge image input to our Swin Transformer is blurred with a Gaussian filter. The blur filtering was introduced to reduce the strictness of the pixel-wise loss [16]. That is, our Swin Transformer takes blurred edge images as input and outputs the probability of defects. Fig. 3 shows the inspection method using our Swin Transformer.

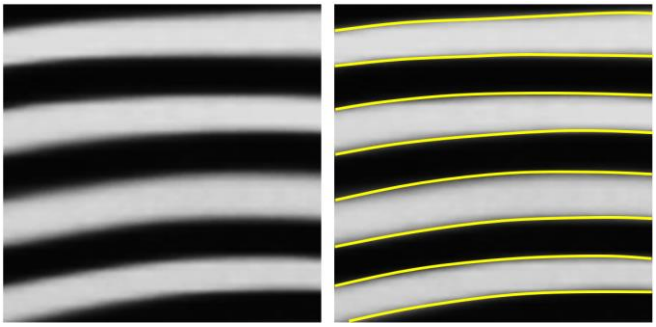

Fig. 2. (Left) Actual inspection image with stripe pattern reflection. (Right) The actual visual inspection focuses on the edges (yellow lines).

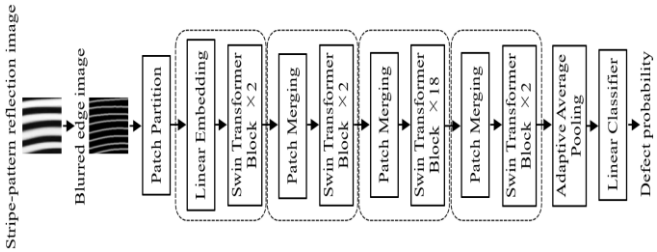

Fig. 3. Flow of the proposed inspection method.

B. Generation of Training Images

As described in Section 1, this study generates edge images of stripe pattern reflections to train the Swin Transformer by optical simulation for the problem of lack of real defect images. Fig. 4 shows the simulation model.

The stripe pattern light source, the test surface, and the camera were positioned so that the reflection of the stripe pattern was in the camera's field of view as shown in Fig. 4. The number of concavity defects was set to 0 (i.e. defect-free) or 1, as it is considered unlikely that more than one concavity defect could occur on the same surface on a tightly controlled production line. In the case that the number of defects was one, a concave defect was placed at a random position in the camera field of view.

The shape of the defect was represented by a two-dimensional Gaussian function, by previous study of the simulation of concave defect visualization [17]. Introducing a three-dimensional coordinate system with a test surface as the xz-plane and the plane containing the optical axis and the normal of the light source as the yz-plane (Fig. 4), the defect is represented by (1).

$$y = -d_{max} e^{-\frac{1}{2}(x,\ z)\begin{pmatrix}{\sigma_x}^2 & \sigma_{xz}\\ \sigma_{zx} & {\sigma_z}^2\end{pmatrix}^{-1}(x,\ z)^{\mathrm{T}}} \quad (1)$$

where ${\sigma_x}^2$ and ${\sigma_z}^2$ are the variances, σ_{xz} and σ_{zx} are the covariances, and d_{max} is the maximum depth of the defect.

The shape of the test surface is expressed in the form of (2) so that flat and free-form surfaces can be represented.

$$y = f(x, z). \quad (2)$$

The sum of (1) and (2) thus represents the test surface containing the defect.

The camera image was simulated using backward ray tracing [18]. Specifically, the camera was a pinhole model, and

 EISSN: 2961-872X ● PISSN: 2961-8924 ● DOI:110.56741/IISTR.esl.00743

when a ray connecting a pixel of the image sensor and the principal point reached the light source by specular reflection on the test surface, the brightness at the point of arrival was used as the pixel density. The brightness distribution of the stripe pattern of the light source was approximated by a Gaussian distribution. The Blinn-Phong reflection model [19] was used to model specular reflection on the surface. It was assumed that there was a blackout curtain in the background of the light source and that the rays that strayed from the light source reached the curtain. Finally, edge images are extracted from the simulated images and further blurred to generate training images as described in sub-section 2.1. Examples of the simulation results when the test surface is cylindrical are shown in Fig. 5.

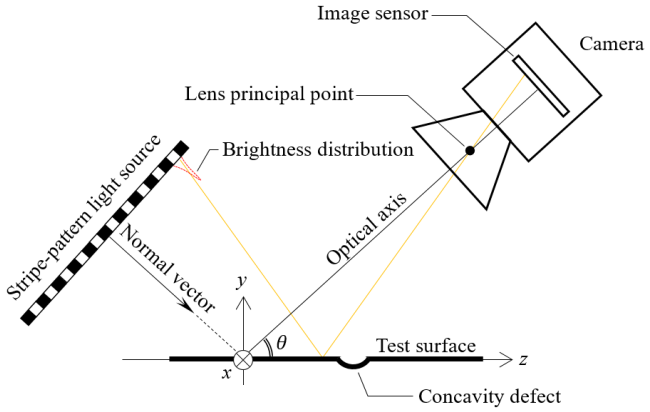

Fig. 4. Simulation model.

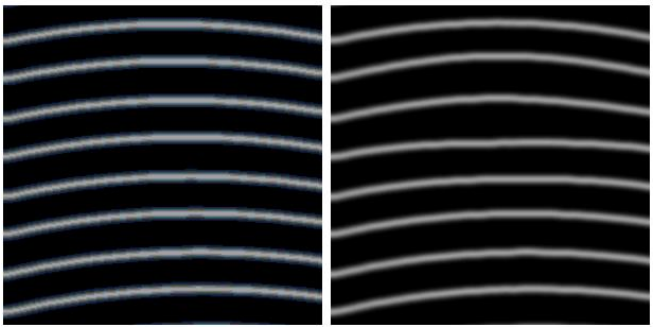

Fig. 5. Examples of images obtained from optical simulation. (Left) Defect-free edge image. (Right) defect edge image.

III. EXPERIMENTS

A. *Training of Swin Transformer*

The experiment was carried out on mirrored steel used for car window frames (size: 150 mm x 17 mm). In the simulations, the geometry of this object was represented by (3) instead of (2), because it took into account the positioning errors of the real object.

$$y = f\left(\begin{pmatrix}\cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}x \\ z\end{pmatrix} + \begin{pmatrix}\Delta_x \\ \Delta_z\end{pmatrix}\right), \tag{3}$$

where Δ_x, Δ_z and α are the misalignment in the x-direction, misalignment in the y-direction, and rotational misalignment around the y-axis, respectively. In the experiment, Δ_x, Δ_z and α were randomly selected from the ranges $[-3.0, 3.0]$, $[-3.0, 3.0]$, and $[-1.0, 1.0]$, respectively. The unit for Δ_x and Δ_z is mm, and that for α is degree.

For the defect size, d_{max}, σ_x and σ_z were randomly selected from the ranges [0.01, 0.10], [1.0, 5.0], and [1.0, 5.0], respectively. $\sigma_{xz} = \sigma_{zx} = 0$ were assumed. The unit for these is mm. The positions of the camera and light source in the simulation were arranged as in the actual experimental setup (Fig. 6). With the above settings, the simulation was run 7,000 times to generate images. 1,500 of these were used for testing. Our Swin Transformer is a model that changes the number of outputs of the last layer of the Swin Transformer-v2 image classification model [20] to 1 and adds a sigmoid function, and we fine-tuned it using these training images.

The training was performed 50 times with 128 mini-batches. The learning rate was 0.0001 and the Adam optimizer [21] was used. The loss function used was cross-entropy. The test error in estimating the defect probability is shown in Fig. 7 (left).

For comparison, we also trained another Swin Transformer to estimate defects from stripe pattern images instead of edge images, and the test error for this is shown in Fig. 7 (right).

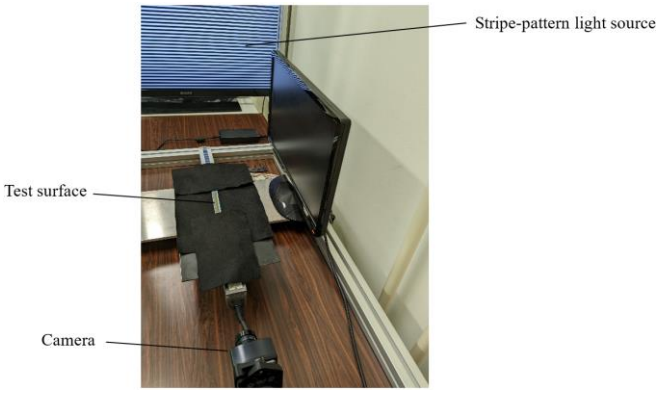

Fig. 6. Experimental equipment.

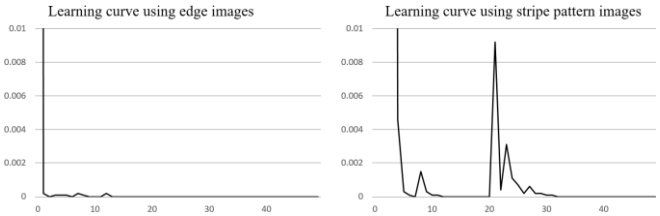

Fig. 7. Learning curve on the test set. (Left) Learning curve using edge images. (Right) Learning curve using stripe pattern images.

B. *Defect Inspection Experiment*

Experiments on defect inspection were carried out using 540 images obtained by photographing the window frame with the equipment shown in Fig. 6. 270 of these images were defect images. Examples of the obtained image are shown in Fig. 8. The images are processed by edge extraction and blurring and used as input to the Swin Transformer.

Defects were determined by whether the defect probability of the output of the Swin Transformer exceeded 0.5. Accuracy (4) and Recall (5) were used as metrics to evaluate the detection results.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \tag{4}$$

$$Recall = \frac{TP}{TP+FN}, \tag{5}$$

where *TP*, *FP*, *FN*, and *TN* are True Positive which represents the number of defective images correctly identified, False Positive which means the number of defect-free images mistakenly detected as defective, False Negative is the number of defective images mistakenly classified as defect-free, and True Negative which denotes the number of non-defective images correctly identified, respectively.

The results are shown in Table 1. The results show that *Accuracy* = 0.996, *Recall* = 1.000. We also conducted an experiment on the Swin Transformer trained using striped pattern images for comparison. The results are shown in Table 2. From the results, Accuracy = 0.928, Recall = 0.941.

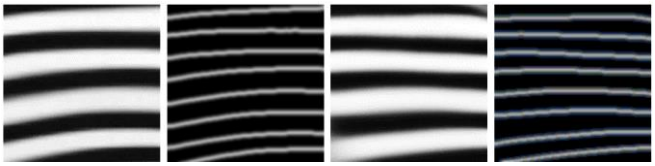

Fig. 8. Examples of real test samples. Left to right: defect-free image, defect-free edge image, defect image, and defect edge image.

TABLE 1 Results for Swin Transformer trained using edge images.

	Predicted class	
Actual class	**Defect**	**Defect-free**
Defect	*TP* = 270	*FN* =0
Defect-free	*FP* = 2	*TN* =268

TABLE 2 Results for Swin Transformer trained using stripe pattern images.

	Predicted class	
Actual class	**Defect**	**Defect-free**
Defect	*TP* = 254	*FN* = 16
Defect-free	*FP* = 23	*TN* = 247

IV. Discussion

From Fig. 7 (left) it can be seen that the training of our Swin Transformer (i.e. the network trained on edge images of reflected stripes) converged to a test error of 0 after 13 epochs, indicating that it had acquired the ability to accurately identify abnormal reflected edges. As a result, in experiments with real products, the *Accuracy* (= 0.996) calculated from Table 1 showed that defect and defect-free products could be distinguished with high accuracy. In addition, the *Recall* (=1.000) calculated from Table 1 shows that there is a 0% error in defect detection. From these results, we can highly evaluate the detection performance of our Swin Transformer as being at a practical level that can be introduced to manufacturing sites.

However, the *FP* (= 2) in Table 1 shows that there are cases where defect-free products are detected as defective. In this case, these images were those in which the misalignment when the test surface was installed. If the surface is curved, the reflection will be distorted by the misalignment. The training images were generated assuming a variation of ±1 degree in orientation, but it is assumed that the misrecognition occurred for images outside the distribution of the training images. A possible solution to this problem would be to increase the range of variation and create more training images, but this would increase the diversity of reflection distortion, which could lead to a decrease in recognition performance. Therefore, it would be more effective to consider ways of accurately positioning the test products.

On the other hand, the Swin Transformer trained on the reflected stripe pattern images converged to 0 after 33 epochs, as shown in Fig. 7 (right), but in the experiment with the real product, the *Accuracy* was 0.928 and the *Recall* was 0.941 (from Table 2). The reason for the reduced performance compared to training on edge images is likely to be due to the difference between the actual reflected light intensity distribution and the intensity distribution in the simulated image, as discussed in Subsection 2.1. In other words, it is believed that the defect detection performance trained on the simulation images was not fully demonstrated on the actual images due to the difference in density distribution. It is therefore considered that the proposed method of training using the edge images of reflected stripes is a suitable one.

V. Conclusion

To automate the inspection of concave defects on mirror surfaces, we proposed a detection method using a Swin Transformer that has learned the normal and abnormal reflections of stripe patterns. In quality-controlled production environments, defects are rare and it is therefore difficult to collect a large number of training images. In this paper, we also proposed a method for generating images using optical simulation. Evaluation experiments using real mirror components showed that the proposed method has a high defect detection performance.

References

[1] V. Azamfirei, F. Psarommatis, and Y. Lagrosen, “Application of Automation for In-Line Quality Inspection, A Zero-Defect Manufacturing Approach,” Journal of Manufacturing Systems, 67, 1-22, 2023. https://doi.org/10.1016/j.jmsy.2022.12.010

[2] A. Tjolleng, J. Chang, J. Park, W. Lee, M. Cha, J. Park, and K. Jung, “Development of A Human-friendly Visual Inspection Method for Painted Vehicle Bodies,” Applied Ergonomics, 106(103911), 2023.

[3] Y. Roh, S. Kim, H. Shim, D. Lim and D. Jeong, "A Patterned Illumination and Phase Integration Imaging Method for High-Glossy Surface Evaluation," proc. 2010 International Symposium on Optomechatronic Technologies, Toronto, Canada, 2010. doi: 10.1109/ISOT.2010.5687315

[4] P. M. F. Forte, P. E. R. Felgueiras, F. P. Ferreira, M. A. Sousa, E. J. Nunes-Pereira, B. P. J. Bret, and M. S. Belsley, “Exploring Combined Dark and Bright Field Illumination to Improve the Detection of Defects on Specular Surfaces,” Optics and Lasers in Engineering, 88, 120-128, 2017. https://doi.org/10.1016/j.optlaseng.2016.08.002

[5] A.-C. Wei, Y-.C. Chang, and J.-R. Sze, “Study on Reduction of Background Fringes for Defect Detection of Specular Surface,” proc. 2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference, Taipei, Taiwan, 2023. doi: 10.1109/APSIPAASC58517.2023.10317452

[6] C. Zuo, S. Feng, L. Huang, T. Tao, W. Yin, and Q. Chen, “Phase Shifting Algorithms for Fringe Projection Profilometry: A review,” Optics and Lasers in Engineering, 109, 23-59, 2018. https://doi.org/10.1016/j.optlaseng.2018.04.019

[7] S. Ordones, M. Servin, M. Padilla, I. Choque, J. L. Flores, and A. Muñozc, “Shape Defect Measurement by Fringe Projection Profilometry and Phase-Shifting Algorithms,” Optical Engineering, 59(1), 014107, 2020. doi: 10.1117/1.OE.59.1.014107

[8] M. C. Knauer, J. Kaminski, and G. Hausler, “Phase Measuring Deflectometry: A New Approach to Measure Specular Free-Form Surfaces,” proc. Optical Metrology in Production Engineering, Strasbourg, France, 2004. https://doi.org/10.1117/12.545704

[9] S. Bazeille, A. Meguenani, K. Tout, S. Kohler, O. Jrad, J.-P. Chambard and C. Cudel, “A Scanning Deflectometry Scheme for Online Defect Detection and 3-D Reconstruction of Specular Reflective Materials,” The International Journal of Advanced Manufacturing Technology, 131, 245–259, 2024. https://doi.org/10.1007/s00170-024-13034-8

[10] M. Song, J. Yi, S. Wei, and Z. Wen, "Mirror-like Defect Detection Technology based on YOLOv5 and Phase Measuring Deflectometry," proc. 2nd International Conference on Robotics, Artificial Intelligence and Intelligent Control, Mianyang, China, 2023. 10.1109/RAIIC59453.2023.10280830.

[11] J. Guan, J. Li, X. Yang, X. Chen, and J. Xi, “Defect Detection Method for Specular Surfaces based on Deflectometry and Deep Learning,”

Optical Engineering, 61(6), 061407, 2022. https://doi.org/10.1117/1.OE.61.6.061407

[12] J. Guan, J. Fei, W. Li, X. Jiang, L. Wub, Y. Liub, and J. Xi, “Defect Classification for Specular Surfaces based on Deflectometry and Multi-Modal Fusion Network,” Optics and Lasers in Engineering, 163(107488), 2023. https://doi.org/10.1016/j.optlaseng.2023.107488

[13] 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, 20(18), 5136, 2020. https://doi.org/10.3390/s20185136

[14] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” Proc. ICCV2021, (virtual), Oct. 2021. doi:10.1109/ICCV48922.2021.00986

[15] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. G., Ł. Kaiser, and I. Polosukhin, “Attention is All you Need,” Proc. NIPS 2017, Long Beach, USA, Dec. 2017.

[16] A. Wang, M. Ren, and R. S. Zemel, “SketchEmbedNet: Learning Novel Concepts by Imitating Drawings,” Proc. ICML2021, (Virtual), 2021.

[17] O. Hirose, A. Ishii, S. Hata and I. Washizaki, “Detection of Small Convex and Concave Defects on Optical Films by Patterned Illumination (2nd Report) - Estimation of Defect Detection by using Ray Tracing Method -,” Journal of the Japan Society of Precision Engineering, vol. 67, no. 7, pp. 1135-1139, 2001.

[18] A. S. Glassner (ed), An Introduction to Ray Tracing, Academic Press, New York, USA, 1989.

[19] J. F. Blinn, "Models of Light Reflection for Computer Synthesized Pictures," in Proc. SIGGRAPH '77, California, USA, 1977.

[20] Z. Liu, H. Hu, Y. Lin, Z. Yao, Z. Xie, Y. Wei, J. Ning, Y. Cao, et al., "Swin Transformer V2: Scaling Up Capacity and Resolution," Proc. CVPR2022, New Orleans, USA, Jun. 2022. doi: 10.1109/CVPR52688.2022.01170

[21] D.P. Kingma and J. Ba, “Adam: A method for stochastic optimization, Proc. ICLR2015, San Diego, USA, May, 2015.