

MIN: Moiré Inpainting Network with Position Adaptive Mask for 3D Height Reconstruction

Tae-Jung Kim¹, Min-Ho Ha², Saba Arshad³ and Tae-Hyoung Park⁴

Abstract—In the AI-driven computer vision industry, Moiré patterns are used for rapid height measurement of components in 2D Printed Circuit Board (PCB) images. However, shadows and light reflections on metallic surfaces can cause errors. To address this, we propose a Moiré Inpainting Network that combines Moiré height measurement with inpainting model features to remove these artifacts. Using a Generative Adversarial Network (GAN) structure, our model accurately identifies and reconstructs shadow and reflection regions. The network inputs 2D Moiré images and outputs images with corrected heights. We assess performance using the Height Reconstruction Rate and GAN performance metrics. Comparative experiments show that our approach significantly outperforms current state-of-the-art inpainting models for Moiré images, proving its effectiveness in the computer vision industry.

Note to Practitioners—This paper was motivated by the problems encountered during defect inspection of Printed Circuit Board (PCB) in the manufacturing process and is applicable not only to 2D PCB images, but also to images with shadows and light reflections that require 3D shape information. The 3D shape information of a 2D PCB image provides height values, which can be used to detect defects in the components of the PCB. In this paper, we propose a method to solve the problems of the moire method through AI algorithms and apply it to the actual process by using the moire method, which has a very fast inspection speed. A network Moire Inpainting Network (MIN) is designed by mixing the inpainting model using Generative Adversarial Network (GAN), moire method, and anomaly detection model. The trained MIN network is designed to reconstruct the shadow and light reflection areas when a moire image is input and output the height value restored by the network. We have configured the UI with MFC in the actual manufacturing process and are preparing to apply it after several demonstrations and challenges with the company.

Index Terms—Artificial intelligence, Computer vision, Generative adversarial networks, Image inpainting, Anomaly detection, Moiré, Printed circuit board.

I. INTRODUCTION

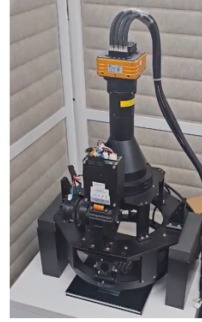
In the industrial vision field, it is important to measure the height of objects in 2D images to obtain 3D shape information

¹T. J. Kim is with the Department of Intelligent Robotics Engineering, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, Republic of Korea (28644). (e-mail: ktajejung@chungbuk.ac.kr)

²M. H. Ha is with the Department of Control and Robot Engineering, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, Republic of Korea (28644). (e-mail: minho6988@chungbuk.ac.kr)

³Saba Arshad is with the Industrial Artificial Intelligence Research Center, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, Republic of Korea (28644). (e-mail: sabarshad@chungbuk.ac.kr)

⁴T. H. Park is with the Department of Intelligent Systems and Robotics, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, Republic of Korea (28644). (e-mail: taehpark@cbnu.ac.kr)



(a) Moiré optical system



(b) VHX-7000N microscope

Fig. 1. (a) Moiré optical system for projection Moiré pattern, (b) Microscope for calculating height value of object

for visual inspection of objects. Visual inspection of objects can reveal defects based on height that are not visible in 2D images. Therefore researches have been conducted on measuring the height of objects in 2D images to obtain 3D shape information [1-9]. Among these, studies using Moiré patterns [5,9] have been carried out. Moiré refers to interference patterns created by overlapping periodic wave patterns [10].

To measure the height of components on the PCB using this method, Moiré optical system is used to capture Moiré patterns shifted at angles of 0, 90, 180, and 270 degrees. However, the PCB Moiré images projected at 0, 90, 180, 270 degrees result in shadows extending behind the components due to the tilted angle of the Moiré optical system. Additionally, since the PCB components are made of metal, light reflections occur during the projection process. These shadows and light reflections cause errors in height measurement. Consequently, accurate height measurements cannot be obtained. Therefore, it is essential to remove and reconstruct the shadows and light reflections to achieve accurate height. Also in the computer vision industry, precise height measurements of PCB components at the micrometer scale are crucial, necessitating an accurate definition of the shadow and light reflection regions to be reconstructed.

Previous studies have proposed various methods to address measure the height of objects in 2D images to obtain 3D shape information. For instance, the laser method [11] suffers from slow processing speeds, while the stereo method lacks accuracy. Other studies [12-14] have attempted to adjust the brightness of the lighting for each component, add hardware specifically for height measurement [15-17], or incorporate additional projected Moiré grids, all of which require hardware

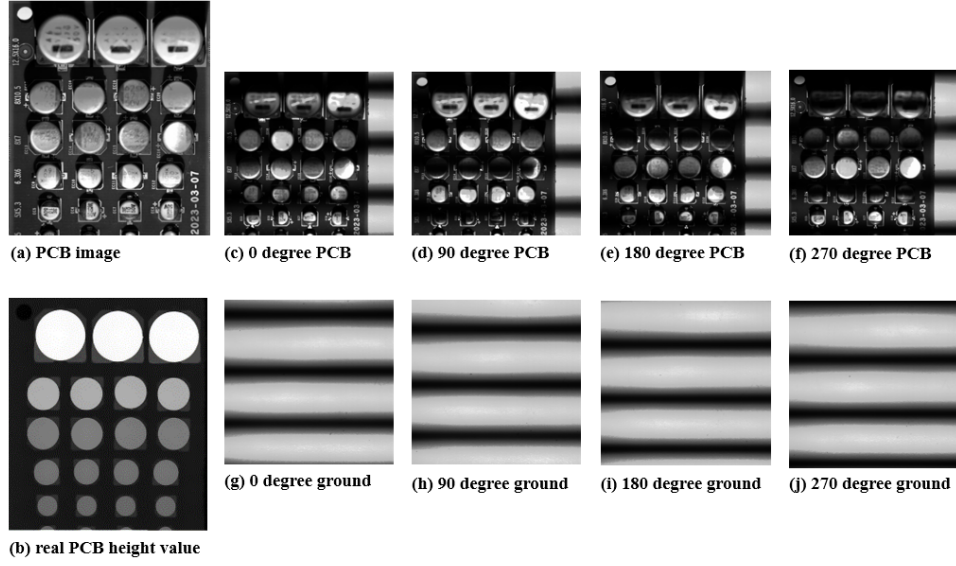


Fig. 2. (a) Printed circuit board image, (b) real Printed circuit board height value image, (c)-(f) Printed circuit board Moiré images in 0, 90, 180, 270 degrees, (g)-(j) background Moiré images in 0, 90, 180, 270 degrees

improvements. Recent research [18-21] has aimed to address this problem using artificial intelligence algorithm instead of hardware solutions, leveraging Moiré patterns. One such approach involves using the generative adversarial network (GAN) [22] model, specifically the Unpaired image-to-image translation using cycle-consistent adversarial networks (CycleGAN) [23], to reconstruct shadows and light reflections. However, this method reconstructs not only shadows and light reflections but also unnecessary regions, which is a significant limitation [20]. Moreover, researches [19,21] utilizing GAN-based inpainting model [24-29] to reconstruct shadows and light reflections has their own limitations. The limitations of these studies lack clear criteria for distinguishing shadows and light reflections, and due to the relatively simple nature of PCB images compared to face images, the two-step reconstruction process often leads to overfitting. These highlight the need for more precise and targeted approach to accurately reconstruct shadows and light reflections in PCB images.

Therefore, we propose a novel approach by designing network named Moiré Inpainting Network (MIN), which integrates the structure of existing GAN-based inpainting models with the unique characteristic of Moiré patterns. We utilize the fast unsupervised anomaly detection with generative adversarial networks (f-AnoGAN model) [30], GAN-based anomaly detection method [31-35], to accurately identify shadow and light reflection regions at the pixel level. In addition, since the shadow and light reflection regions vary for each acquisition with the Moiré optical system, we propose Position Adaptive Mask method to find fluid regions by utilizing the f-AnoGAN model.

First, the unique characteristic of Moiré patterns was referenced in the following way, and the method was applied to MIN. We will cover how to measure height via Moiré patterns in more detail later, but it requires two steps: converting the PCB Moiré images projected at 0, 90, 180, 270 degrees to a phase map, and converting the phase map to a 2D unwrap.

During these two steps, the shadow and light reflection regions will be different, and we have to set the shadow and light reflection regions accordingly. Therefore, we use this characteristic to reconstruction from phase map and 2D unwrap separately.

Second, the structure of existing GAN-based inpainting models was referenced in the following way, and the method was applied to MIN. During the reconstruction process for removing and reconstructing shadows and light reflections, we refer to the structure of inpainting models but avoid the two-step reconstruction process to prevent overfitting. Instead, we perform a single reconstruction for each. The first reconstruction is performed on the phase map for shadow and light-reflection regions, and the second reconstruction is performed on the 2D unwrap in the same way.

This MIN structure maintains the existing inpainting model size while significantly enhancing reconstruction performance. Also in this way, our method allows direct height measurement from the reconstructed image. Therefore the designed MIN takes Moiré images as input, removes and reconstructs shadow and light reflection regions, and output the actual height. MIN handles the entire process from Moiré images to actual height, allowing to obtain 3D height information directly from 2D Moiré images.

The contributions of our proposed method are as follows. In the computer vision industry, when it comes to acquiring 3D shape information for visual inspection of objects, the height measurement method using Moiré acquires information very quickly but suffers from errors due to shadow and light reflection regions. In this paper, we propose a method to reconstruct shadow and light reflection regions by utilizing inpainting, an AI algorithm. MIN, which is designed by mixing the existing inpainting model and Moiré characteristic, shows high efficiency in acquiring 3D shape information from 2D Moiré images.

In addition, by applying an anomaly detection method to the

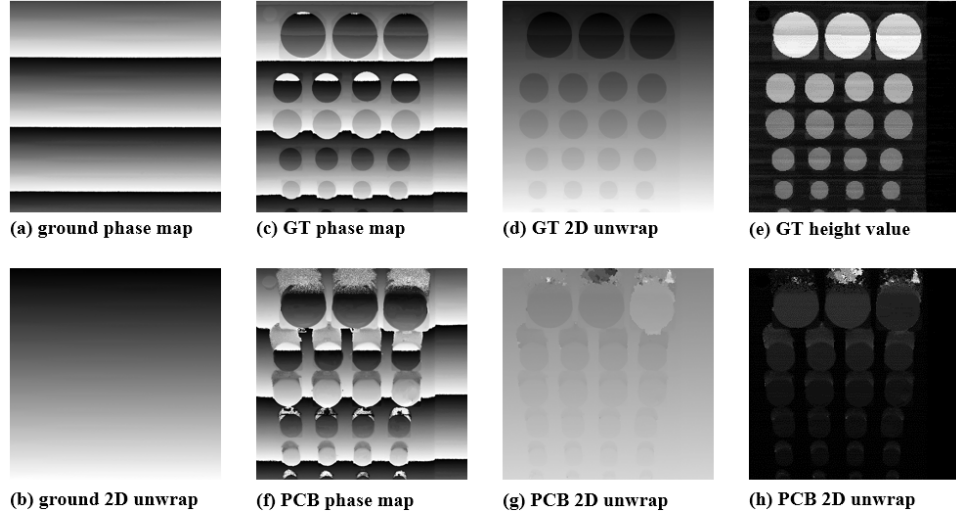


Fig. 3. (a) Background phase map image, (b) Background 2D unwrap image, (c) Real height phase map image, (d) Real height 2D unwrap image, (e) Real height rate, (f) PCB Phase map with shadow and light reflection, (g) PCB 2D unwrap image, (h) PCB height rate

regions to be reconstructed in the existing inpainting model, it is possible to set the regions accurately on a pixel-by-pixel basis. The PAM method utilizing this method can flexibly cope with errors that may occur in a dynamic manner.

In Section II, we provide a detailed explanation of the height measurement process using Moiré patterns and the data preprocessing steps. Section III describes the reconstruction networks. Section IV presents a detailed explanation of the proposed MIN, while Section V discusses the experimental setup and evaluation the results. Finally, Section VI concludes this research.

II. MOIRÉ PATTERN

The process of height measurement using Moiré patterns can be summarized as follows. Firstly, four Moiré patterns are captured and converted into phase map. Then this phase map is transformed into 2D unwrap through an unwrapping process, enabling height measurement. To acquire the dataset, the 2D image of the PCB acquired through the Moiré optical system is shown in Fig. 2(a), and the actual height value of the PCB acquired through the microscope is shown in Fig. 2(b).

First, four Moiré images are captured at angles of 0, 90, 180, and 270 degrees using 2D image of the PCB by the Moiré optical system shown in Fig. 1(a). The Moiré images projected onto the PCB, denoted as $I_n(x, y)$, (where $n = 1, 2, 3, 4$), are shown in Fig. 2(c)-(f). The Moiré images projected onto an empty background, denoted as $I_{n_{ref}}(x, y)$ (where $n_{ref} = 1, 2, 3, 4$), are shown in Fig. 2(g)-(j). Here, x and y represent the coordinates in 2D space, n and n_{ref} indicate the sequence of images from left to right.

$$\varphi_p(x, y) = \arctan \frac{I_4(x, y) - I_2(x, y)}{I_1(x, y) - I_3(x, y)} \quad (1)$$

The phase map is conducted using phase-shifting Moiré interferometry [36], which features a repeating pattern of values between $0-\pi$, demonstrates a shift in the pattern at the component locations on the PCB, as seen when comparing

$\varphi_{p_{ref}}(x, y)$ and $\varphi_p(x, y)$ in Fig. 3 and described by Eq. (1). This shift is due to the height of the components on the PCB, and by measuring the extent of this shift, the height of the components can be determined. Subsequently, the process of unwrapping the repetitive pattern in the phase map into a single continuous pattern, known as 2D unwrap, is described by Eq. (2). As a result, $\varphi_p(x, y)$ and $\varphi_{p_{ref}}(x, y)$ are transformed into $\varphi_{u-p}(x, y)$ and $\varphi_{u_{ref}}(x, y)$, respectively, which correspond to Fig. 3(g)-(b).

$$\varphi_{u-p}(x, y) = \varphi_p(x, y) + 2N\pi \quad (2)$$

By using the 2D unwrap results $\varphi_{u-p}(x, y)$ and $\varphi_{u_{ref}}(x, y)$ obtained through the formula, we can derive H_{rate} , the height rate value of the components on the PCB, as shown in Fig. 3(h). Since H_{rate} reflects height rate value in the presence of shadows and light reflections on the PCB, it inherently includes many errors. To convert H_{rate} from a height rate value to an actual height value, we need both the actual height values and the corresponding rate values derived from the actual heights. To achieve this, we measure the actual height of the PCB using a microscope, as depicted in Fig. 1(b), and convert this measurement into an image, resulting in H_{real} , the actual height values shown in Fig. 2(b). Using H_{real} and $\varphi_{p_{ref}}(x, y)$ from Fig. 3(a), we create $\varphi_{p_{gt}}(x, y)$, the phase map of the actual heights, as shown in Fig. 3(c). By similarly unwrapping $\varphi_{p_{gt}}(x, y)$, we obtain $\varphi_{u_{gt}}(x, y)$, the unwrapped phase map, shown in Fig. 3(d). By comparing $\varphi_{u_{gt}}(x, y)$ with $\varphi_{u_{ref}}(x, y)$ from Fig. 3(b), we can determine $H_{real-rate}$, the actual height rate values. Now, with $H_{real-rate}$ containing the actual height rate values and H_{real} representing the actual heights, we use a Multi layer perceptron (MLP)[37] to calculate the relationship between the height rate value and the actual height values. By applying the weights and biases w and b obtained from the MLP, as shown in Eq. (3), we can adjust the error height rate values H_{rate} to obtain corrected height values, as shown in Fig. 3(h). Additionally, by converting $H_{real-rate}$ back to actual height values using the same method, we achieve the

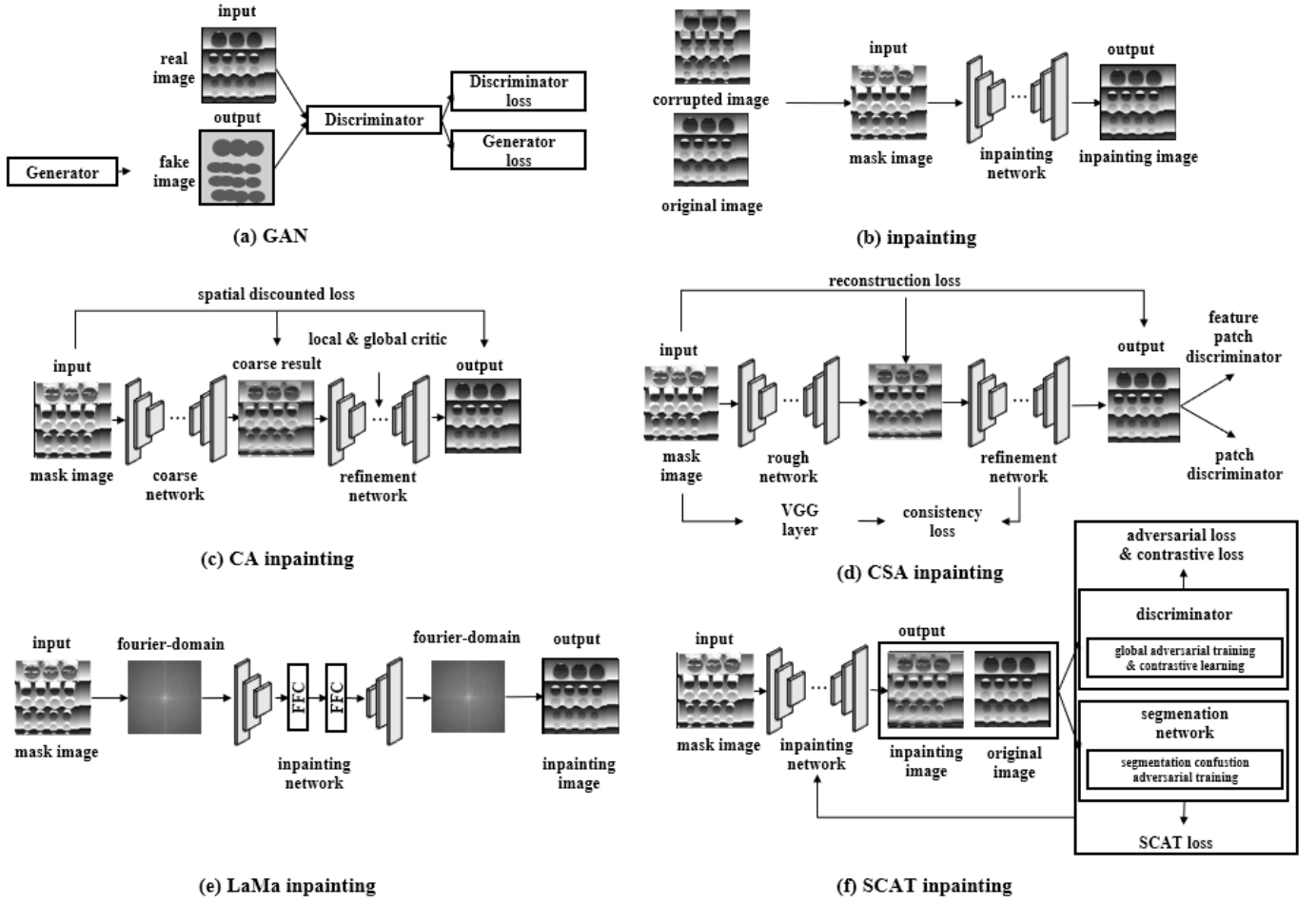


Fig. 4. (a) GAN model architecture, (b) inpainting model architecture, state-of-the-art inpainting models with (c) CA inpainting model architecture, (d) CSA inpainting model architecture, (e) LaMa inpainting model architecture, (f) SCAT inpainting model architecture

reconverted actual heights, as depicted in Fig. 3(e). However, there is a discrepancy between the actual heights in Fig. 2(b) and the reconverted heights in Fig. 3(e), which is attributed to errors arising during the projection of the Moiré pattern.

$$H_{real} = H_{rate} * w + b \quad (3)$$

However, comparing the height values in Fig. 3(h)-(e) reveals significant errors due to the interference from shadows and light reflections. Therefore, it is necessary to remove these factors in advance. To achieve this, we design a mixed reconstruction approach called MIN, which combines the height reconstruction method of Moiré patterns with an inpainting model to remove and reconstruct the shadow and light reflection regions. Furthermore, to accurately identify the shadow and light reflection regions, we define a Position Adaptive Mask to create mask images. This method helps in isolating the problematic regions, allowing for their effective removal and reconstruction.

III. RECONSTRUCTION NETWORK

A. Generative adversarial network

The existing state-of-the-art inpainting models[38-41], such as Generative image inpainting with contextual attention (CA

inpainting)[38], Coherent semantic attention for image inpainting (CSA inpainting)[39], Resolution-robust large mask inpainting with fourier convolutions (LaMa inpainting)[40] and Generative image inpainting with segmentation confusion adversarial training and contrastive learning (SCAT inpainting)[41], are all based on the GAN framework. The foundational structure of the GAN[22] model, as shown in Fig. 4(a), aims to generate images that appear real. The GAN model consists of two networks: the Generator network (G) and the Discriminator network (D). The role of the generator is to create images, while the discriminator's role is to distinguish between real images and those created by the generator. Through a process of adversarial training, these two networks are pitted against each other. The generator continually improves its ability to produce realistic images, while the discriminator enhances its ability to distinguish between real and generated images.

B. Inpainting

The inpainting model, depicted in Fig. 4(b), aims to reconstruct corrupted images to their original state. Although based on the GAN model, the inpainting model not only generates images but also reconstructs the corrupted regions

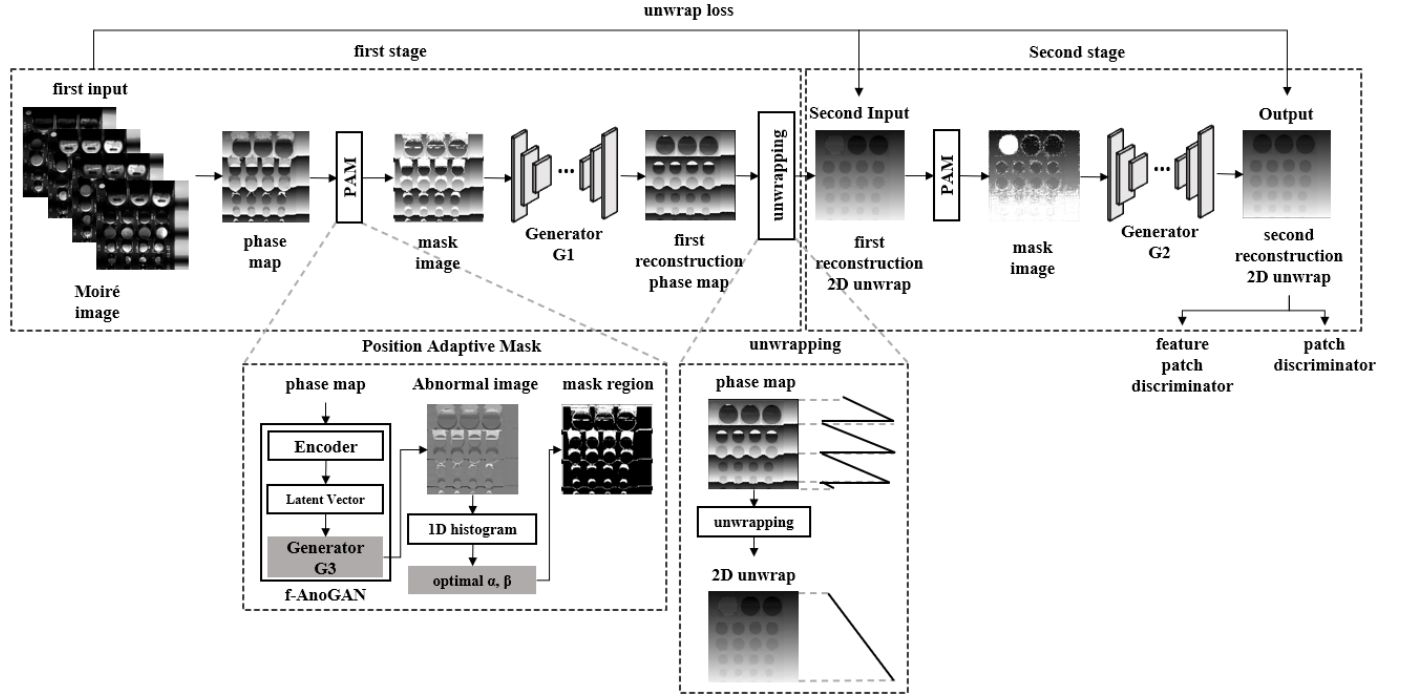


Fig. 5. Moiré Inpainting Network architecture with Position Adaptive Mask and unwrapping architecture

to resemble the original image. The corrupted regions are defined as mask regions, which are marked with white pixels. The input to the model is the original image overlaid with the mask image, and the model focuses on reconstruction the masked regions to create a natural resemblance to the original image. While traditional inpainting models reconstruct damaged images to their original form, this study focuses on reconstructing images considering the regions of shadows and light reflections. Using a PCB phase map dataset, the original image is without shadows and light reflections, whereas the corrupted image includes these elements. The inpainting image output shows the shadow and light reflection regions removed and reconstructed to appear as the original regions.

The CA inpainting model[37], shown in Fig. 4(c), improves on the traditional inpainting model structure with a two-stage process to enhance reconstruction performance. For the PCB dataset, the input is the masked image, and the output is the reconstructed image with shadows and light reflections removed and reconstructed. reconstruction occurs in two stages: initially in the coarse network and subsequently refined in the refinement network. This model emphasizes image reconstruction around the mask region, filling in the surroundings naturally for focused reconstruction.

The CSA inpainting model[38], illustrated in Fig. 4(d), also uses a two-stage structure like the CA inpainting model but significantly enhances performance by introducing the CSA layer. Additionally, it includes a VGG Layer and uses both feature patch discriminator and patch discriminator in the discriminator network to further improve reconstruction performance. The input and output for the PCB dataset are the same as for the CA inpainting model.

The LaMa inpainting model[39], shown in Fig. 4(e), differs

from previous inpainting models by using fast fourier convolution (FFC) instead of traditional convolution operations in the image domain. This approach transforms images into the Fourier domain, allowing for broader contextual information and reduced computational complexity, resulting in faster training and higher-quality image reconstruction. Unlike previous models, it achieves superior performance without requiring two stages of reconstruction. The input and output for the PCB dataset are the same as for the CA inpainting model.

The SCAT inpainting model[40], depicted in Fig. 4(f), introduces segmentation confusion adversarial training (SCAT) and contrastive learning. The segmentation network labels the mask and non-mask regions, aiming to distinguish the masked regions. The generator works to reconstruct the mask regions realistically, disrupting the segmentation network's ability to identify them. The SCAT loss enhances reconstruction performance. contrastive learning trains the model to make the reconstructed image more distinct from the damaged image while closely resembling the original image. The input and output for the PCB dataset are the same as for the CA inpainting model.

The role of the inpainting model in this paper is to reconstruct the image in the desired style for the mask region of the image, but there is no proper criterion for the mask region. For the purpose of this paper, we need criterion for shadow and light reflection regions in PCB Moiré images. In addition, we need a fluid mask image for the shadow and light reflection regions that occur during the phase map and 2D unwrap process in the Moiré height measurement method. Therefore, we propose a novel network, the Moiré Inpainting Network, to serve as an additional role of the inpainting model.

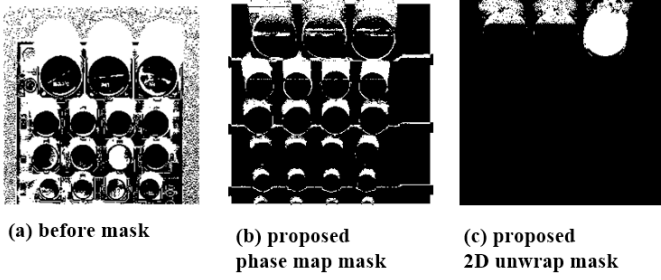


Fig. 6. (a) mask without PAM, (b) phase map mask with PAM, (c) 2D unwrap mask with PAM

IV. MIN : MOIRÉ INPAINTING NETWORK

A. Two Stage Architecture

The architecture of the MIN model is shown in Fig. 5. It is designed as a two-stage structure, inspired by the CSA inpainting and CA inpainting models. In the Generator ($G1$) of first stage, reconstruction is performed on the phase map of Moiré image, and in the Generator ($G2$) of second stage, reconstruction is performed on the first reconstructed 2D unwrap of phase map. The Position Adaptive Mask (PAM) method is applied in both stages to dynamically update the mask image during each training iteration. In unwrapping between the first and second stages converts the phase map to 2D unwrap.

This two-stage structure, with each stage performing reconstruction once, is designed considering the characteristics of the Moiré height measurement method discussed in Section II. Moiré height measurement method requires two conversion processes, and during the conversion from phase map to 2D unwrap, slight pixel differences can often lead to image distortion. Additionally, since the phase map and 2D unwrap are relatively simple datasets and require relatively short training times due to the smaller dataset size, a single reconstruction at each stage is sufficient, because performing reconstruction twice could lead to overfitting. This approach allows the first stage to adequately reconstruct the shadow and light reflection regions in the phase map, and the second stage to further reconstruction any remaining error regions in the 2D unwrap. Consequently, the outputted 2D unwrap image can be directly converted into height value. Ultimately, the MIN network is designed to take a 2D Moiré image as input and output the reconstruction real height. This ensures that the model effectively addresses and corrects the distortions caused by shadows and light reflections, providing accurate height reconstruction in PCB phase map and 2D unwrap.

B. PAM : Position Adaptive Mask

In traditional methods [20, 22] for defining shadow and light reflection regions in PCB 2D Moiré images, these regions are determined based on pixel brightness. Specifically, in the 2D PCB Moiré images, pixels with values greater than 230 are defined as light reflection regions, and pixels with values less than 30 are defined as shadow regions. In resulting mask, with shadow and light reflection regions set to 255, remaining are set to 0, is shown in Fig. 6(a) as the without Position

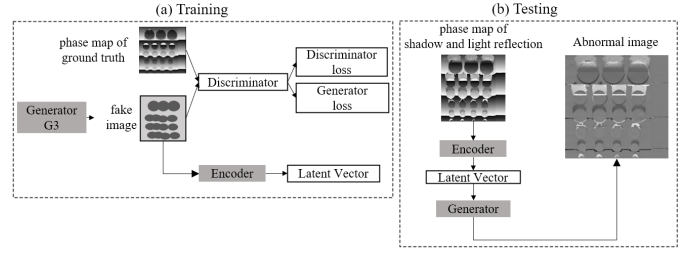


Fig. 7. f-AnoGAN model architecture, (a) model training and (b) model testing

Adaptive Mask. However, this method is imprecise as it relies on subjective visual criteria. Moreover, the images inputted for reconstruction are not Moiré images but phase map or 2D unwrap, leading to different damaged regions due to shadow and light reflection during the conversion process. Consequently, the reconstructed height using the conventional mask affects unnecessary regions and fails to properly reconstruct fine details requiring reconstruction.

To address these issues, this paper proposes Position Adaptive mask that uses an anomaly detection model to accurately identify shadow and light reflection regions and generate precise pixel-level mask image through additional preprocessing. In image processing, an anomaly detection model distinguishes between a normal image and an abnormal image, identifying the anomalous regions, shadow and light reflection regions in this paper, as a mask. When there is no ground truth data for the mask regions, a GAN-based anomaly detection model can be used. This type of model can detect anomalies using only images with and without anomalies. Therefore, we use f-AnoGAN, a GAN-based model, to detect anomalies between phase map of ground truth and phase map of shadow and light reflection. The detected anomalies are then post-processed to convert them into mask images. By using this method, the phase map mask with Position Adaptive Mask is shown in Fig. 6(b). Similarly, the 2D unwrap mask with Position Adaptive Mask, which detect anomalies between 2D unwrap of ground truth and 2D unwrap of shadow and light reflection, is shown in Fig. 6(c).

1) *f-AnoGAN model*: The anomaly detection model used in this paper is the f-AnoGAN model, depicted in Fig. 7. The model is trained using images without anomalies, and once training is complete, it detects the anomalous regions, such as shadows and light reflections. During training process in Fig. 7(a), as shown in the figure, the f-AnoGAN model train the Generator ($G3$) through GAN training to generate phase map of ground truth. Then, through the Encoder (E) training, the generated phase map is converted into latent vector. By extracting the image features into latent vector using only the E , fast computation is achieved.

During the testing process in Fig. 7(b), the phase map of shadow and light reflection is given as the model. The E converts this image into latent vector, and the $G3$ uses this vector to regenerate the image, detecting anomalies in the process. The detected anomalies produce an abnormal image, which features notably bright or dark pixels in the anomaly regions.

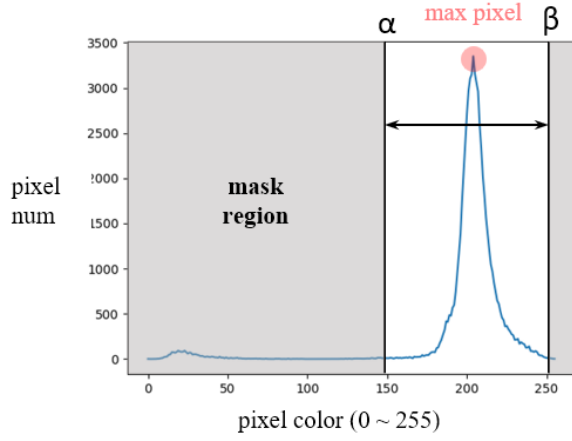


Fig. 8. Abnormal image from f-AnoGAN to 1D histogram

Algorithm 1 Position Adaptive Mask

Input: Image $I_{original}$
Output: Mask image I_{mask}
encoder: $Z_{latent} = E(I_{original})$
generator: $I_{abnormal} = G(Z_{latent})$
Initialize H_{hist} of size I_{width} with zeros
for each $x \leftarrow I_{width}$ **and** $y \leftarrow I_{height}$ **do**
 $H_{hist}[I_{abnormal}(x, y)] ++$
end for
 $H_{max} = \max(H_{hist})$
for each $x \leftarrow I_{width}$ **do**
 if $H_{hist}[x] == H_{max}$ **then**
 $P_{max} = x$
 end if
end for
for each $x \leftarrow I_{width}$ **and** $y \leftarrow I_{height}$ **do**
 for each $\alpha \leftarrow 30$ **and** $\beta \leftarrow 30$ **do**
 if $I_{abnormal}(x, y) > P_{max} + \beta$ **and** $I_{abnormal}(x, y) < P_{max} - \alpha$ **then**
 $M_{region}(x, y) = 255$
 else
 $M_{region}(x, y) = 0$
 end if
 end for
end for
Initialize I_{mask} of size (I_{width}, I_{height}) with zeros
for each $x \leftarrow I_{width}$ **and** $y \leftarrow I_{height}$ **do**
 if $M_{region}(x, y) == 255$ **then**
 $I_{mask}(x, y) = M_{region}(x, y)$
 else
 $I_{mask}(x, y) = I_{original}$
 end if
end for

2) *1D Histogram Computation:* the generated abnormal image is converted into a 1D histogram, resulting in a graph Fig. 8. The x-axis represents the pixel color (0-255), and y-axis represents the total number of pixels in the image. From this graph, it is evident that the regions without anomalies have the highest pixel counts around the max pixel. By setting α

for threshold of darker pixel and β for threshold of brighter pixel centered on the max pixel, it is possible to isolate the anomaly regions. The anomaly regions are set to a pixel of 255, and the remaining regions are set to 0, creating the mask image.

However, to set appropriate α and β , a criterion is needed. This paper uses the Height Error Rate (HER), which is the inverted value of Height Reconstruction Rate (HRR). It is a percentage of how much height is reconstructed, and when flipped, we get a percentage of the height difference, which we define as the HER. For α and β , mask images are obtained within the range of 0-30, and real height of ground truth and real height of reconstruction image datasets are prepared. We take our datasets and use HER to find how much the actual heights differ. The range of α and β is limited to 30 because mask regions become excessively large and meaningless beyond this point. As illustrated in Fig. 8, the mask region is set based on the α and β , resulting in smaller mask regions for smaller values and larger regions for larger values. Observing the recorded values shows that as α and β increases, the HER initially increases and then decreases. A higher HER indicates a higher error of height between two images, which means the mask region requiring reconstruction is larger. However, when the HER value starts to decrease after reaching its peak, it implies that the mask region has become unnecessarily large, covering regions that do not require reconstruction. Ultimately, the point where the HER value first reaches its maximum indicates the optimal mask region, and the corresponding α and β are chosen to generate the optimal mask image. Although the same PCB image is used, the shadow and light reflection regions differ slightly with each projection from Moiré optical system, resulting in different optimal α and β .

By implementing this approach, we aim to more accurately define and mask the shadow and light reflection regions, leading to better reconstruction result for the PCB phase map and 2D unwrap. This will enhance the accuracy of the height reconstruction process by focusing on the actual regions requiring correction. In the MIN model, this method allows for the real-time generation and updating of mask images during training, ensuring that the mask accurately correspond to the error regions.

The PAM can be summarized with Algorithm 1. In the first stage, $I_{phasemap}$ of phase map is input to generate the I_{mask} of mask image. In Algorithm 1, the encoder and generator refer to the $G3$ and E of pre-trained f-AnoGAN model. By using $I_{abnormal}$ of the abnormal image obtained through f-AnoGAN model, H_{hist} of 1D histogram is computed, and P_{max} of the max pixel from Fig. 8 is determined. The optimal α and β are then found using the previous method to identify M_{region} of the mask region. M_{region} is combined with $I_{phasemap}$ to generate I_{mask} . In the second stage, the same process is applied with the 2D unwrap of first reconstruction input to generate the I_{mask} of mask image.

C. Unwrapping

The Moiré height measurement method requires conversion to a phase map and then to a 2D unwrap. The phase map,

Algorithm 2 unwrap loss**Input:** 2D unwrap image I_u of shape $[height, width]$ **Output:** unwrap loss U_{loss} $A = |I_u[: -2, :]|$ $C = |I_u[1 : -1, :]|$ $B = |I_u[2 :, :]|$ $U_{loss} = |A - C| - |C - B|$

as shown in Fig. 3(a), consists of repetitive periodic patterns, while the 2D unwrap, as shown in Fig. 3(b), unfolds these periodic patterns into a single pattern. In the 2D unwrap, the pattern's slope is linear, meaning that the pixel difference between specific pixels $y + n$ and $y + (n + 1)$ along the y-axis of a 256x256 image is constant, where $n = 0, 1, \dots, 255$. The parts of the 2D unwrap where the pixel differences are not constant are considered errors, and the sum of these errors is defined as unwrap loss U_{loss} . By applying U_{loss} during training, the pixel differences can be minimized, thus enhancing the reconstruction performance in the 2D unwrap. In training part, as shown in Fig. 5, U_{loss} is calculated from both the second input and output of the 2D unwrap, and parameter updates are performed accordingly.

Algorithm 2 illustrates the steps involved in U_{loss} computation. For a 2D unwrap image I_u , we generate three images A, B and C as follows. Where A is the image I_u with the last two rows removed, B is the one with the first two rows removed, C is with the first and last rows removed. The process involves computing the absolute differences between A and C, and C and B, respectively, and then finding the difference between these absolute values. This results in the loss for the 2D unwrap, U_{loss} .

D. HRR : Height Reconstruction Rate

The goal of this paper is to reconstruct images to measure the heights of components on a PCB, comparing the reconstructed height to the actual height. Therefore, we define the HRR as the performance metric to compare the degree of height reconstruction between the proposed MIN network and various inpainting models. To compare height reconstruction using HRR, we need to define the comparison region, the subjects to be compared, and the method for comparison.

First, we must define the comparison region to compare the degree of height reconstruction. The data used in this paper consists of three types of Printed circuit board image datasets, defined as PCB1, PCB2, and PCB3. We obtain the actual height data for PCB1, PCB2, and PCB3 using the same method as the Moiré pattern in Section II shown in Fig. 9(a)-(c). By removing pixel values greater than the height of the PCB, the non-component regions will have a pixel value of 0. Assigning a pixel value of 255 to all non-zero pixel values will extract only the component regions of the PCB. The objective of this paper is the height reconstruction of PCB components, so extracting the component regions for each PCB results in Fig. 9(d)-(f).

Next, to compare the degree of height reconstruction, we must define the two datasets to be compared. For the actual

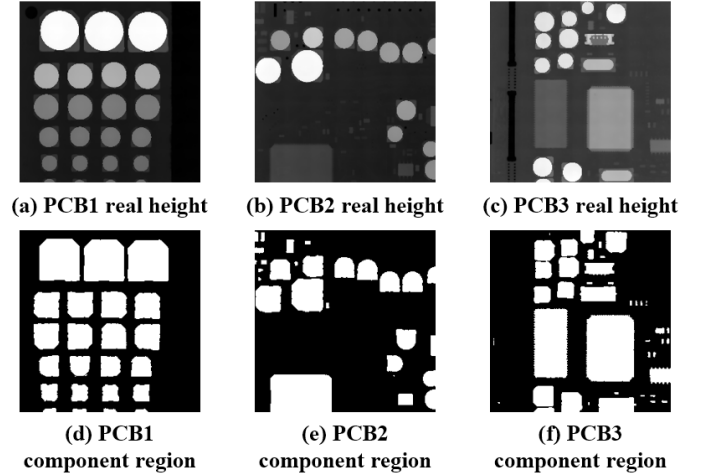


Fig. 9. (a)-(c) real height value of PCB1, PCB2, PCB3, (d)-(f) only component region of PCB1, PCB2, PCB3, consisting of pixel 255

height data, we use phase map of ground truth and define H_{real} . For the reconstructed height data, we use phase map of shadow and light reflection and define H_{recon} .

Finally, we define the method for comparing the degree of height reconstruction. The comparison region in the compare region has a pixel value of 255, while other regions have a pixel value of 0. The comparison method involves dividing H_{real} by H_{recon} for the 255-pixel comparison region and expressing it as a percentage, known as the HER. HRR is then defined as inverted value of HER, as shown in Eq. (4). HRR allows us to evaluate the degree of height reconstruction by comparing the reconstructed height to the actual height.

$$HRR = \begin{cases} (1 - H_{recon}/H_{real}) * 100, & Area \neq 0 \\ 0, & otherwise \end{cases} \quad (4)$$

This paper presents a new network, MIN, designed by combining the Moiré height measurement method with features of traditional inpainting models. Unlike traditional inpainting models that use random or fixed rectangular masks, MIN generates Position Adaptive Mask using anomaly detection. In the first and second stage of Fig. 5, MIN continuously updates the mask for dynamically changing error regions, allowing for flexible reconstruction. This approach ensures precise pixel-level mask region settings and captures additional errors through Position Adaptive Mask.

E. Training

The overall learning process is as follows: Firstly, the network structure in the first stage and second stage, as shown in Fig. 5, is a basic u-net structure used in the inpainting model. It consists of an encoder and decoder structure with 4x4 convolutions with skip connections. Initially, the phase map of ground truth and phase map of shadow and light reflection are input into the MIN. In the PAM, phase map of shadow and light reflection is used to accurately define the shadow and light reflection regions at the pixel level, creating a mask region. For training, the input to the first stage is phase

map of ground truth, which does not contain shadow and light reflection. The Mask image, created by overlaying the mask region on the phase map, is input into the network, focusing on image reconstruction in the masked region. This allows the model to learn the context of regions without shadow and light reflection. Once the model is trained, phase map, containing shadow and light reflection regions, is input into the first stage. Based on the previous training, the model outputs a phase map with the shadow and light reflection regions removed and reconstructed.

In the second stage, the input is the 2D unwrap of ground truth and the 2D unwrap of reconstructed phase map in the first stage. During the unwrapping of the reconstructed phase map, minor pixel breakages and unreconstructed regions can cause errors. The PAM generates mask regions in real-time for these errors and breakages. The mask image, created by combining these generated mask regions with 2D unwrap of ground truth, is input into the network. The model focuses on reconstructing the image in the masked regions, learning the context of regions without errors or breakages. Once trained, the second stage receives the 2D unwrap from the first stage and outputs a 2D unwrap of second reconstruction with the errors and breakages removed and reconstructed.

During the backpropagation process after training, Feature discriminator (F) and Patch discriminator (P) are used as referenced in [39]. The parameter update process for the Discriminator is as follows: By using the two discriminators F and P, faster and higher quality image generation is possible. Mean squared error loss is computed using the squared difference between the two images, while L1 loss is computed using the absolute difference between the two images. The predicted values of the reconstructed image and the ground truth image are input into P, and P_{loss} is calculated using Mean squared error loss. Additionally, during the two-stage process, the high-level feature space of both the ground truth and the generated images are extracted from the ImageNet-pretrained VGG-19. F_{loss} is calculated using L1 loss from these extracted values. Finally, the parameters of the discriminator are updated with D_{loss} , which is the sum of P_{loss} and F_{loss} in equal proportion. The parameter update process for the generator is similar to that of the discriminator. After calculating P_{loss} and F_{loss} , the previously defined Unwrap loss, U_{loss} , is also calculated for the generated image. All these losses are summed and adjusted by a certain lambda value to obtain G_{loss} , which is used to update the generator's parameters. Parameters for both the first stage and second stage are updated accordingly.

F. Validation

The overall inference process is as follows: In the first stage, unlike the training process, the input is the Moiré images that contain shadows and light reflections. The mask image is generated by inputting phase map of Moiré image into the PAM, where the shadow and light reflection regions are already present. Then, the first stage model, which has learned the context of the mask regions from the phase map of ground truth, removes and reconstructs these mask regions and outputs the first reconstructed phase map.

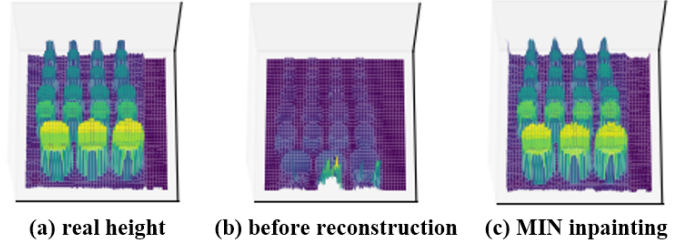


Fig. 10. height reconstruction image to 3D plot, (a) real height value of PCB components, (b) height value of PCB components with shadow and light reflection, (c) height value of PCB components after reconstruction with MIN inpainting

In the second stage, the first reconstructed phase map is unwrapped to create a 2D unwrap. The PAM then updates the mask image in real-time to address any errors or breakages caused by minor pixel differences. The second stage model, which has learned the context of the mask regions from the 2D unwrap of ground truth, references these mask regions to remove and reconstruct the 2D unwrap and outputs the second reconstructed 2D unwrap. Finally, the difference between the second reconstructed 2D unwrap and 2D unwrap of background is used to obtain the height rate values, and the actual height values of the second reconstructed image are output through the MLP using the calculated w and b .

In Fig. 10, the before and after reconstruction results are shown with 3D plots of the reconstructed height. Fig. 10(a) is a 3D plot of the actual height, Fig. 10(b) is a 3D plot of the height reconstructed using the pre-reconstruction Moiré height reconstruction method, and Fig. 10(c) is a 3D plot of the height reconstructed using the MIN model. The figure visually demonstrates that the height reconstruction result using the MIN model closely resembles the actual height value, highlighting the importance of removing shadow, light reflection regions, and error regions from the Moiré image.

V. EXPERIMENTS AND RESULTS

A. Experimental setup

The experiments were conducted using two GTX GeForce 3080 GPUs. The data was obtained by projecting four Moiré patterns with shadows and light reflections onto the images using a Moiré optical system. Each PCB was projected 100 times, resulting in a total dataset of 300 images across three types of PCBs. Multiple shots were taken for each PCB because the shadow and light reflection regions vary slightly with each shot, thus capturing multiple images increases the variety of data. The actual height values of the PCBs were collected using a microscope and converted into images from csv files. Additionally, the four Moiré images were converted into phase map and 2D unwrap to generate the necessary data. The background Moiré pattern was used to create the images of the actual PCB height values, which were then transformed into phase map and 2D unwrap for further use.

In the comparative experiments, considering the relatively simple image features and the small dataset, the number of training iterations was fixed at 500. The image size was set to 256×256 pixels. To validate the effectiveness of the approach

TABLE I
HRR OF HEIGHT VALUE IMAGE WITH MIN DEPENDING ON WHETHER UNWRAP LOSS [%]

MIN	before U_{loss}	after U_{loss}
none	67.66 \triangle 0.3%	67.91
two stage	73.60 \triangle 0.8%	74.21
PAM	71.45 ∇ 0.9%	70.84
both	84.27 \triangle 1.0%	85.12

TABLE III
HRR OF HEIGHT VALUE IMAGE WITH MIN DEPENDING ON WHETHER PAM [%]

MIN	before mask	PAM
none	67.66 \triangle 5.6%	71.45
unwrap loss	67.91 \triangle 4.3%	70.84
two stage	73.60 \triangle 14.5%	84.27
both	74.21 \triangle 14.7%	85.12

proposed by the MIN model, comparative experiments were conducted by distinguishing between before and after applying the PAM, as well as before and after applying the unwrap loss. Additionally, comparisons were made between one-stage and two-stage structures. In these comparisons, four scenarios were tested to evaluate performance using HRR. The one-stage structure reconstruct only the phase map and directly using it for real height via unwrapping. In cases where PAM was not applied, the previous mask generation method was used to create mask images by designating regions with more than 230 pixels as light reflection regions and regions with less than 30 pixels as shadow regions. For comparison with existing inpainting models, the performance of five models was evaluated: CA inpainting, CSA inpainting, LaMa inpainting, SCAT inpainting, and the proposed MIN. Performance metrics included Fréchet inception distance (FID), Learned perceptual image patch similarity (LPIPS), Structural similarity index measure (SSIM), and Peak signal-to-noise ratio (PSNR), which are commonly used for evaluating GAN models. FID [42] is a metric used to evaluate the quality of images generated by generative models, comparing the distributions of real and generated images using feature vectors from a pre-trained Inception network. LPIPS [43] measures the perceptual similarity between two images by comparing deep features extracted from trained neural networks, aiming to better align with human visual perception. SSIM [44] is a metric that assesses the similarity between two images based on their luminance, contrast, and structure, providing a value that quantifies visual similarity. PSNR [45] is a metric used to measure the quality of reconstruction of lossy compression codecs, comparing the original and compressed images by calculating the ratio between the maximum possible power of a signal and the power of corrupting noise. Additionally, HRR, defined in the paper, was used to assess the degree of height reconstruction, which is the main goal of the study. All of the above performance metrics measure the similarity between the height of ground truth image and the height of

TABLE II
HRR OF HEIGHT VALUE IMAGE WITH MIN DEPENDING ON WHETHER TWO STAGE [%]

MIN	one stage	two stage
none	67.66 \triangle 8.8%	73.60
unwrap loss	67.91 \triangle 9.3%	74.21
PAM	71.45 \triangle 17.9%	84.27
both	70.84 \triangle 20.2%	85.12

reconstruction image. The lower the value for FID and LPIPS, and the higher the value for SSIM, PSNR, and HRR, the more similar the two images are.

B. Results

The comparison results for the four different scenarios in the MIN model, based on whether unwrap loss is applied, are shown in Table I. The first row of Table I compares the MIN model before and after applying U_{loss} , dividing the comparison into before and after the loss application. The first column of Table I categorizes the scenarios based on whether both two-stage and PAM are applied, resulting in four cases: none (neither applied), one applied at a time, and both (both applied). \triangle is the percentage performance improvement of the proposed method over the existing method. ∇ is the percentage performance reduction of the proposed method compared to the existing method. The performance metric HRR indicates that HRR generally improves after applying unwrap loss compared to before its application.

The comparison results based on whether the two-stage process is applied are shown in Table II. The first row of Table II compares the MIN model in one stage and two stage configurations. The second column also categorizes the scenarios based on the application of unwrap loss and PAM, resulting in four cases. The performance metric HRR indicates that applying Two stage significantly improves HRR.

The comparison results based on whether PAM is applied are shown in Table III. The first row of Table III compares the MIN model before applying the mask and after applying PAM. The second column categorizes the scenarios based on the application of unwrap loss and the two-stage process, resulting in four cases. The performance metric HRR indicates that applying PAM significantly improves HRR. Particularly, combining PAM with the two-stage structure shows a significant improvement in reconstruction rate.

Through these results in Tables I-III, we confirmed that the proposed method in the MIN model is effective. We then conducted comparative experiments with existing inpainting models by applying all these methods. The comparison involved converting the phase map results generated by the existing inpainting models into actual height values and comparing them with the actual height values converted by the MIN model using performance metrics. The comparison results are shown in Table IV, where the first column lists CA Inpainting, CSA Inpainting, LaMa Inpainting, SCAT Inpainting, and the proposed MIN model in order. The first row shows the comparison results between the images converted into

TABLE IV
PERFORMANCE METRICS OF HEIGHT VALUE IMAGE WITH INPAINTING MODELS

model	FID ↓	LPIPS ↓	SSIM ↑	PSNR ↑	HRR ↑
CA inpainting	4388.11 Δ 7982%	0.50 Δ 625%	0.68 Δ 40%	24.60 Δ 84%	58.55% Δ 45%
CSA inpainting	458.51 Δ 834%	0.32 Δ 400%	0.86 Δ 10%	34.38 Δ 32%	71.36% Δ 19%
SCAT inpainting	1427.11 Δ 2596%	0.38 Δ 475%	0.84 Δ 13%	33.33 Δ 36%	72.18% Δ 18%
LaMa inpainting	635.81 Δ 1157%	0.36 Δ 450%	0.85 Δ 12%	33.76 Δ 34%	73.25% Δ 16%
MIN (proposed)	54.97	0.08	0.95	45.28	85.12%

height values and the actual height values using the GAN performance metrics and HRR mentioned earlier. The results indicate that the MIN model shows the best performance in height reconstruction using Moiré images.

VI. CONCLUSION

In the AI computer vision industry, Moiré patterns enable rapid height measurement of components in 2D Printed Circuit Board (PCB) images. However, shadows and light reflections from metallic surfaces often lead to errors in these measurements. To address this issue, we introduce a Moiré Inpainting Network with position adaptive mask for 3D reconstruction that integrates Moiré height measurement with inpainting model features to remove and reconstruct shadows and reflections. This network, based on Generative Adversarial Networks (GANs), utilizes fast unsupervised anomaly detection to accurately identify regions affected by shadows and reflections. The Moiré Inpainting Network processes 2D Moiré images, removing and reconstructing shadows and light reflections, and outputs the corrected height measurements. We evaluate the performance using the Height Reconstruction Rate and GAN performance metrics. Comparative experiments show that our Moiré Inpainting Network significantly outperforms existing state-of-the-art inpainting models for Moiré images. This demonstrates the effectiveness of our method for Moiré height measurement in the computer vision industry. Moreover, unlike traditional inpainting methods that focus solely on restoring the original image, this approach can reconstruct images in a desired style using flexible mask images. This versatility makes it suitable for various types of image reconstructions beyond just Moiré images.

ACKNOWLEDGMENTS

This work is supported by Innovative Human Resource Development for Local Intellectualization program through the Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT)(IITP-2024-2020-0-01462)

REFERENCES

- [1] T.-H. Park and N. Kim, "A dynamic programming approach to PCB assembly optimization for surface mounters," in *International Journal of Control, Automation and Systems*, pp. 192–199, Apr. 2007.
- [2] J.-S. Lee and T.-H. Park, "Defect classification of components for SMT inspection machines," in *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 21, no. 10, pp. 982–987, 2015.
- [3] J.-D. Song, Y.-G. Kim, and T.-H. Park, "defect classification method of PCB solder joint by color features and region segmentation," in *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 23, no. 12, pp. 1086–1091, 2017. doi: 10.5302/J.ICROS.2017.17.0187
- [4] S. Zhang, "High-speed 3D shape measurement with structured light methods:a review," in *Optics and Lasers in Engineering*, vol. 106, pp. 119–131, July. 2018.
- [5] Z. Wang, Z. Zhang, N. Gao, Y. Xiao, F. Gao, and X.Jiang, "Single-shot 3D shape measurement of discontinuous objects based on a coaxial fringe projection system," in *Applied Optics*, vol. 58, no 5, pp. A169–A178, 2019.
- [6] C. Lin, D. Zheng, K. Qian, J. Han, and L. Bai "Spatial pattern-shifting method for complete two-wavelength fringe projection profilometry," in *Optics Letters*, vol. 45, no 11, pp. 3115–3118, 2020.
- [7] W. Choi, Y. Park, and J. Lee, "The method of precision three dimension height inspection by laser scan," in *The Journal of Korean Institute of Information Technology*, vol. 9, no 12, 2021.
- [8] F. Zhong, R. Kumar, and C. Quan, "A cost-effective single-shot structured light system for 3D shape measurement," in *IEEE Sensors Journal*, vol. 19, no 17, pp. 7335–7346, Sept. 2019.
- [9] S. Tang, X. Zhao, G. Sun, and D. Zhao, "Adaptive fringe projection for 3D shape measurement with large reflectivity variations by using image fusion and predicted search," in *International Journal of Optics*, vol. 2020, pp. 1–14, 2020.
- [10] J. E. Greivenkamp, J. H. Bruning, D. Malacara, "Phase shifting in terferometry", In *Optical Shop Testing*, pp.547–666, 1992.
- [11] W. Y. Choi, Y. J. Park, and J. H. Lee, "The method of precision three dimension height inspection by laser scan," In *The Journal of Korean Institute of Information Technology*, vol. 9, no. 12, pp. 221–227, Dec. 2011.
- [12] C. Waddington and J. Kofman, "Saturation avoidance by adaptive fringe projection in phase-shifting 3D surface-shape measurement," In *International Symposium on Optomechatronic Technologies*, 2010.
- [13] J. Wang and Y. Yang, "High-speed three-dimensional measurement technique for object surface with a large range of reflectivity variations," In *Applied Optics*, vol 57, no 30, pp. 9172–9182, 2018.
- [14] H. Lin, J. Gao, Q. Mei, Y. He, J. Liu, and X. Wang, "Adaptive digital fringe projection technique for high dynamic range three-dimensional shape measurement," In *Optics Express*, vol 24, no. 7 pp.7703–7718, 2016.
- [15] Z. Qi, Z. Whang, J. Huang, Q. Xue, and J. Gao, "Improving the quality of stripes in structured-light three-dimensional profile measurement," In *Optical Engineering*, vol. 56, no 3, 2017.
- [16] B. Zhang, Y. Ouyang, and S. Zhang, "High dynamic range saturation intelligence avoidance for three-dimensional shape measurement," In *2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pp.981–990, Shenzhen, China, 2015.
- [17] L. Ekstrand and S. Zhang, "Autoexposure for three-dimensional shape measurement using a digital-light-processing projector," In *Optical Engineering*, vol. 50, no. 12, 1 December, 2011.
- [18] C.-H. Kim, et al, "AI-based reflective area restoration method for three-dimensional AOI," In *Journal of the Association for Control and Robotics Systems*, pp: 936–944, 29.11 (2023).
- [19] C.-H. Kim, T.-J. Kim, and T.-H. Park, "Deep learning based Moiré pattern restoration method using phase features," In *Journal of the Association for Control and Robotics Systems*, 30.5 (2024): 555–562.
- [20] T.-J Kim, M.-H Ha, and T.-H Park, "Phase map Reconstruction of Moiré Patterns Using GAN," In *Proceedings of the National Conference of the Control Robotics Systems Society*, (2023): 495–496.
- [21] T.-J. Kim, C.-H. Kim, and T.-H. Park, "Moiré height phase value restoration using inpainting," In *Journal of Information and Control*, (2023): 162–163.

- [22] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, and Y. Bengio, "Generative adversarial networks," In *IEEE signal processing magazine*, pp. 35(1), 53-65, 2018.
- [23] Zhu, Jun-Yan, et al, "Unpaired image-to-image translation using cycle-consistent adversarial networks," In *Proceedings of the IEEE international conference on computer vision*, 2017.
- [24] H. Zheng, Z. Lin, J. Lu, S. Cohen, E. Shechtman, C. Barnes, and J. Luo, "Image inpainting with cascaded modulation GAN and object-aware training," In *European Conf. on Computer Vision*, pp. 277-296, October 2022.
- [25] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, and V. Lempitsky, "Resolution-robust large mask inpainting with fourier convolutions," In *Proc. of the IEEE/CVF winter conference on applications of computer vision*, pp. 2149-2159, 2022.
- [26] W. Li, Z. Lin, K. Zhou, L. Qi, Y. Wang, and J. Jia, Mat: "Mask-aware transformer for large hole image inpainting," In *Proc. of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10758-10768, 2022.
- [27] D. P. Kingma, and M. Welling, "Auto-encoding variational bayes," In *arXiv preprint arXiv*, 1312.6114, 2013.
- [28] S. Zhao, J. Cui, Y. Sheng, Y. Dong, X. Liang, E. I. Chang, Y. Xu, "Large scale image completion via co-modulated generative adversarial networks," In *arXiv preprint arXiv*, pp. 2103-10428, 2021.
- [29] P. Jeevan, D. S. Kumar, A. Sethi, "WavePaint: Resource-efficient Token-mixer for Self-supervised Inpainting," In *arXiv preprint arXiv*, pp. 2307-00407, 2023.
- [30] T. Schlegl, P. Seebeck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, "f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks," In *Medical image analysis*, pp.54, 30-44, 2019.
- [31] T. Schlegl, P. Seebeck, S. M. Waldstein, U. Schmidt-Erfurth, G. and Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," In *Int. Conf. on information processing in medical imaging*, pp. 146-157, May 2017.
- [32] K. Batzner, L. Heckler, and R. Köhniß, "EfficientAD: Accurate visual anomaly detection at millisecond-level latencies," In *Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 128-138, 2024.
- [33] A. Mousakhan, T. Brox, and J. Tayyub, "Anomaly Detection with Conditioned Denoising Diffusion Models," In *arXiv preprint arXiv*, 2305.15956, 2023.
- [34] A. Berg, J. Ahlberg, and M. Felsberg, "Unsupervised learning of anomaly detection from contaminated image data using simultaneous encoder training," In *arXiv preprint arXiv*, 1905.11034, 2019.
- [35] P. Liznerski, L. Ruff, R. A. Vandermeulen, B. J. Franks, K. R. Müller, M. Kloft, "Exposing outlier exposure: What can be learned from few, one, and zero outlier images," In *arXiv preprint arXiv*, pp. 2205-11474, 2022.
- [36] J. E. Greivenkamp and J. H. Bruning, "Phase shifting interferometry," In *Optical Shop Testing*, D. Malacara, ed., pp.547-666, Wiley, New York, 1992.
- [37] Rosenblatt, F, "The perceptron: a probabilistic model for information storage and organization in the brain," In *Psychological review*, pp. 65.6 (1958): 386.
- [38] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," In *Proc. of the IEEE conference on computer vision and pattern recognition*, pp. 5505-5514, 2018.
- [39] H. Liu, B. Jiang, Y. Xiao, and C. Yang, "Coherent semantic attention for image inpainting," In *Proc. of the IEEE/CVF International Conference on Computer Vision*, pp. 4170-4179, 2019.
- [40] Suvorov, Roman, et al, "Resolution-robust large mask inpainting with fourier convolutions," In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2022.
- [41] Z. Zuo, L. Zhao, A. Li, Z. Wang, Z. Zhang, J. Chen, and D. Lu, "Generative Image Inpainting with Segmentation Confusion Adversarial Training and Contrastive Learning," In *arXiv preprint arXiv*, pp. 2303-13133, 2023.
- [42] Heusel, Martin, et al, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," In *Advances in neural information processing systems*, 30 (2017).
- [43] Zhang, Richard, et al, "The unreasonable effectiveness of deep features as a perceptual metric," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018.
- [44] Wang, Zhou, et al, "Image quality assessment: from error visibility to structural similarity," In *IEEE transactions on image processing*, 13.4, (2004): 600-612.

- [45] Gonzalez, Rafael C, "Digital image processing," In *Pearson education india*, 2009.

VII. BIOGRAPHY SECTION



Taejung Kim Graduated from Chungbuk National University, Department of Information and Communication Engineering in 2023. Master's program in the Department of Intelligent Robotics at the same graduate school from 2023 to the present. His research interests include image processing and machine vision.



Minho Ha Graduated from Chungbuk National University, Department of Electronic Engineering, 2019. M.S. in Control and Robotics Engineering, Dongguk University, 2019-2021. Currently a PhD student at the same graduate school. His research interests include image processing and machine vision.



Saba Arshad Saba Arshad received the B.S. Hons degree in computer science from Pir Mehr Ali Shah Arid Agriculture University, Pakistan, and M.S. degree in computer science from the COMSATS University Islamabad, Pakistan, in 2015 and 2017. She received a PhD degree in control and robot engineering from Chungbuk National University, South Korea, in 2023. Her research interests include visual simultaneous localization and mapping, long-term visual place recognition, loop closure detection, visual-lidar fusion-based object detection, HD mapping for autonomous vehicle navigation, image processing, Anomaly detection and machine learning. Currently, she is a postdoctoral research fellow at Industrial AI Research Center, Chungbuk National University, South Korea.



Taehyoung Park Graduated from Seoul National University, Department of Control and Instrumentation, 1988. M.S., Dongguk Graduate School, 1990. D., Dongguk University, 1994 1994 1997 Senior Researcher, Samsung Tech One 1997 Present Professor, Department of Intelligent Robotics, Chungbuk National University. Visiting Professor, University of Toronto, 2000 2001. Chairman of the Robotics and Automation Research Group of the Institute of Electrical and Electronics Engineers, 2012-2017. Director of Grand ICT Research Center (Industrial Artificial Intelligence Research Center), designated by the Ministry of Science and ICT, 2020-present. His research interests include robotics and automation, autonomous driving, etc.