TECHNIQUE FOR REMOVING UNNECESSARY SUPERIMPOSED PATTERNS FROM IMAGE USING GENERATIVE NETWORK

Kazutake Uehira and Hiroshi Unno

Department of Network Engineering, Kanagawa Institute of Technology, Kanagawa, Japan

ABSTRACT

A technique for removing unnecessary patterns from captured images by using a generative network is studied. The patterns, composed of lines and spaces, are superimposed onto a blue component image of RGB color image when the image is captured for the purpose of acquiring a depth map. The superimposed patterns become unnecessary after the depth map is acquired. We tried to remove these unnecessary patterns by using a generative adversarial network (GAN) and an auto encoder (AE). The experimental results show that the patterns can be removed by using a GAN and AE to the point of being invisible. They also show that the performance of GAN is much higher than that of AE and that its PSNR and SSIM were over 45 and about 0.99, respectively. From the results, we demonstrate the effectiveness of the technique with a GAN.

KEYWORDS

GAN, Auto encoder, Depth map, Pattern removing.

1. INTRODUCTION

3D images have been used for various applications and will attract more attention in the multimedia field as advanced 3D displays for virtual reality, augmented reality, etc. are developed. The depth map represents two-dimensional distance information from a camera to a subject, and it is key information for creating 3D image content [1-3]. There are various ways to obtain a depth map. The pattern projection method is one effective method for obtaining a depth map because it makes it possible to obtain a map even in areas where there are no textures or edges, for which it would be difficult for other methods to obtain the depth. The pattern projection method is divided into two types: the projection of a visible pattern [4, 5] and the projection of an invisible infrared pattern [6, 7]. Visible patterns are used when only a depth map is needed because the projection pattern appears as noise in a captured RGB image. An infrared pattern is used when a depth map and a RGB image are both needed because the projected pattern does not appear in RGB images. This method requires an infrared projector and an infrared camera in addition to a normal RGB camera.

We have been developing a technique for acquiring depth maps that projects visible periodic patterns on a subject [8]. When a periodic pattern is projected onto a subject under certain conditions, a depth map is obtained from the spatial frequency of the projected pattern in the captured image with a camera.

David C. Wyld et al. (Eds): ITCSE, ICDIPV, NC, CBIoT, CAIML, CRYPIS, ICAIT, NLCA - 2021
pp. 13-22, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.110902

In the applications that we are assuming, both an RGB image and a depth map are used, and depth map is embedded in the RGB image invisibly. To achieve this, first, we project a periodic pattern to obtain a depth map, and after obtaining the depth map, we remove the pattern from the RGB image. The main question was how to remove the superimposed pattern in the image of a subject. Many studies have been reported on periodic noise reduction in images [9-13]. Many of these studies applied the frequency domain approach, as periodic noise cannot be simply separated from the original image in the spatial domain. In the frequency domain, the periodic pattern is concentrated at one point, which makes it easy to remove it. As this kind of method, a method using a notch filter [9, 10] and a method using a median filter in the frequency domain [11] and so on have been reported. However, in our case, the spatial frequency of the pattern depends on the depth of the subject and changes depending on the location, so it has a spread in the frequency domain. Therefore, these conventional methods cannot satisfactorily remove the pattern.

In our previous study, we confirmed the feasibility of a technique that uses a generative adversarial network (GAN) to remove a line and space pattern from a target image [14]. In this study, we continue development on the method of our previous study, verifying the performance of the GAN under various conditions and comparing it with an auto encoder (AE) as a generative network using a deep learning method similar to a GAN. This paper describes the experiments we conducted and clarifies the condition for removing the line and space pattern from captured images.

The remainder of this paper organized as follows. Section 2 overviews the whole study where this study is a part. Section 3 describes the experiment. Section 4 presents results of the experiments and discusses them. Finally, Section 5 presents our conclusions.

2. Method of Acquiring Depth Map by Periodic Pattern Projection

Figure 1 shows the overall configuration of our study, which we are now working on. A periodic pattern is projected onto a subject. A camera captures an image of the subject on which the pattern is projected. If the projector is placed a distance away from the camera in the depth direction as shown in Figure 1 (shown as D in the figure), the spatial frequency of the projected pattern in the captured image depends on the depth of the subject. Therefore, a depth map can be obtained by obtaining the spatial frequency of the projection pattern for each small area in the image captured with the camera.

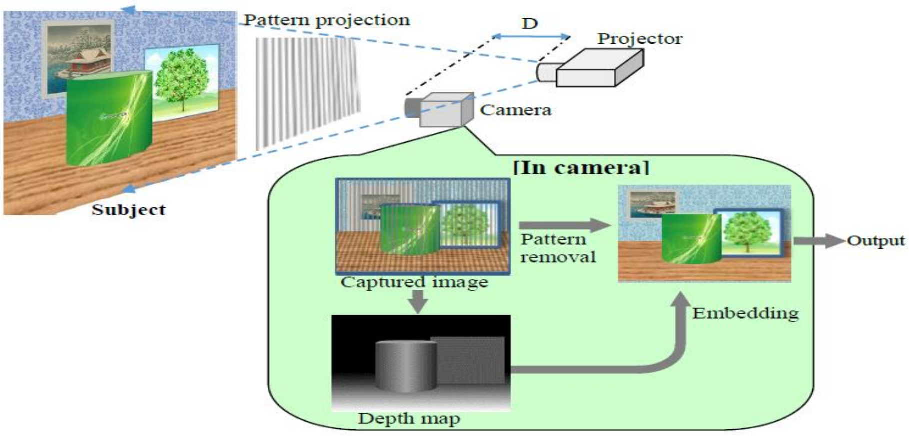


Figure 1. Overall configuration of proposed technology

The contrast of the projection pattern is set high so as to make it easy to obtain the spatial frequency of the pattern. Therefore, the pattern is clearly superimposed in the captured image. When using this image as an RGB image of the subject, this superimposed periodic pattern is obstructive and thus need to be removed. After removing this pattern, the depth map obtained is embedded in an RGB image invisibly. This can be done with technology for hiding information in images used such for as digital watermarking etc. We already confirmed that a depth map could be embedded in the image of a subject and embedded data was retained even if JPEG-compressed [15].

The final output image of this system can be treated as a normal 2D image for transmission and storage, although it includes 3D information in the form of a depth map. If necessary, we can read out the depth map to create a 3D image from the 2D RGB using the depth map.

Removing the superimposed periodic pattern is key for achieving this system. Because the periodic pattern depends on the depth of the subject and differs from place to place, that is, the pattern period is not constant and unknown in the captured image, it is difficult to remove patterns until they disappear with conventional methods mentioned above.

3. EXPERIMENTS

We conducted experiments by simulation. Figure 2 explains how to produce captured images of subjects on which periodical patterns are projected. Figure 3 shows the example images used as subjects. First, we cropped 448 images of 256 × 256 pixels from the images of subjects. We used the line and space (L/S) pattern as the periodical projection pattern. Assuming that the pattern is projected with blue light, B component images of the cropped images and projection pattern images were multiplied to produce captured images, on which the L/S pattern was superimposed. We used L/S patterns with different amplitudes and spatial frequencies. The averaged brightness was 200, and the amplitude was changed in 5 steps, of which the minimum was 10 and the maximum was 50. These values indicate the grayscale with a maximum of 255. Figure 4 shows magnified L/S pattern and the example image of the simulated captured images when the amplitude was 10 and 50. The spatial frequency was changed in 16 steps from 0.18 to 0.25 lines/pixel.

We generated 7,192 simulated captured images from the combination of 448 subject images and 16 L/S patterns with different spatial frequencies. Of these, 5,600 were used for learning, 700 were used for evaluation in learning, and 700 were used for evaluation in terms of peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM).

Figure 5 shows the configuration of the generative adversarial network (GAN) and Figure 6 shows the configuration of the encoder, decoder and discriminator of the GAN. We mainly used the GAN, and the AE was used as a reference. The encoder and decoder of the generator of the GAN consisted of 8 layers each, and the discriminator consisted of 6 layers. These layers were convolution layers, and a 4 x 4 kernel was used for convolution. Leaky-relu was used as an activation function in the encoder and the discriminator and relu was used in the decoder. Dropout was used in the decoder and dropout rate was 0.25.

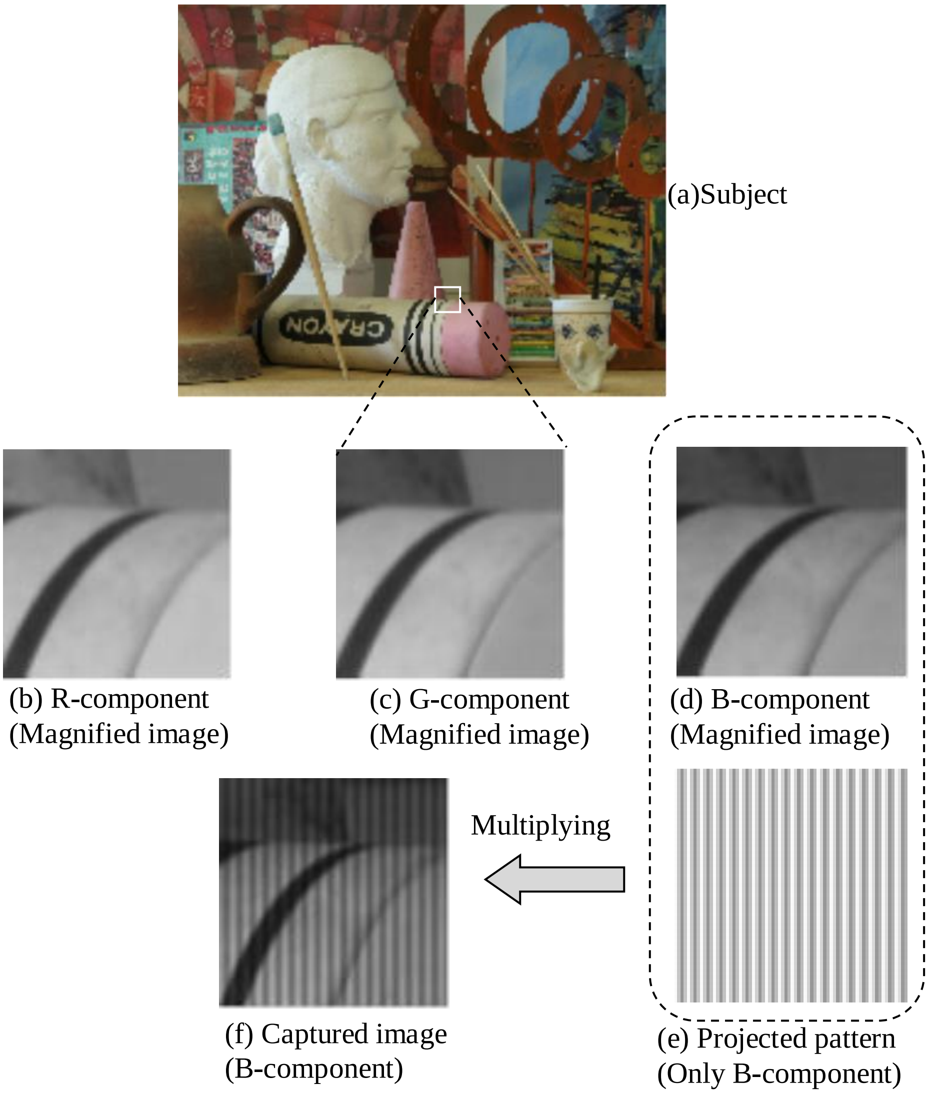


Figure 2. Simulation procedure of generating captured images

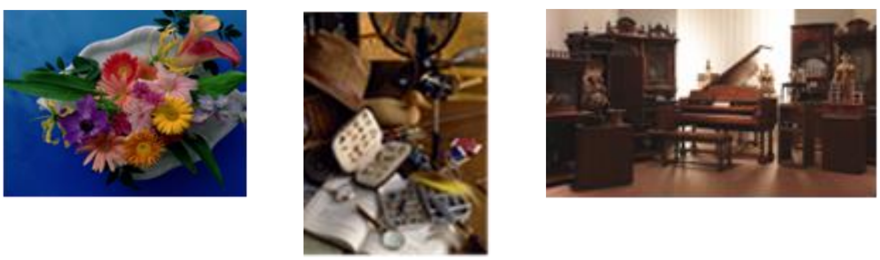


Figure 3. Example of images used as subjects

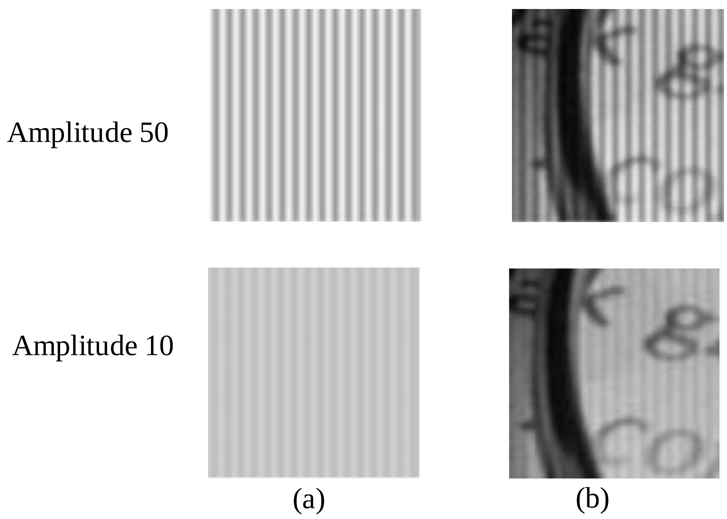


Figure 4. Magnified projected pattern (a) and simulated captured image (b)

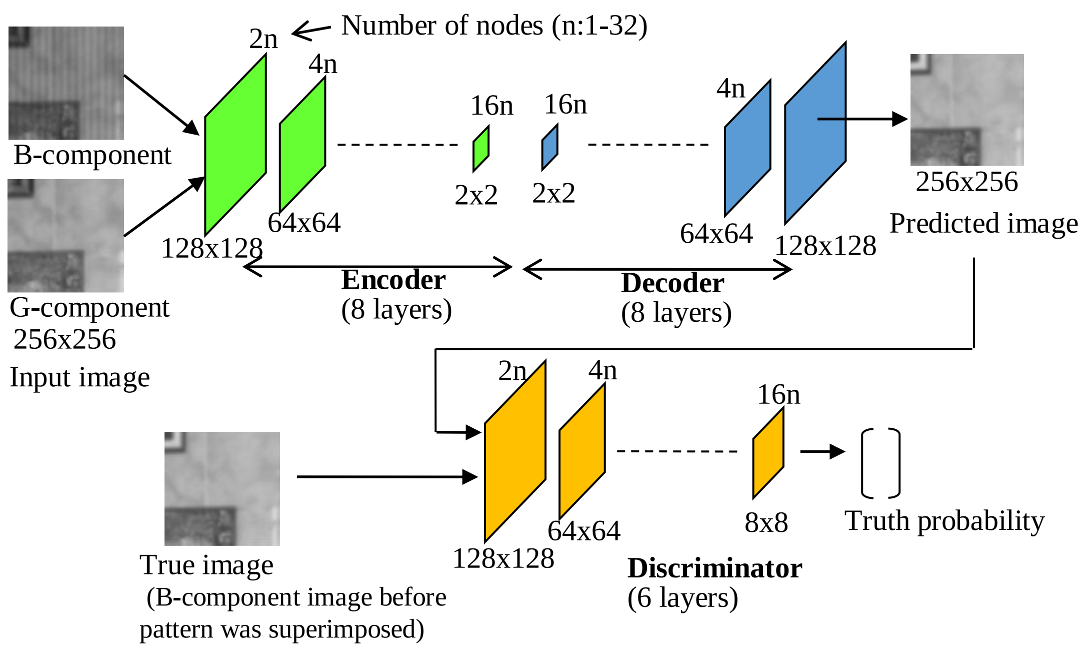
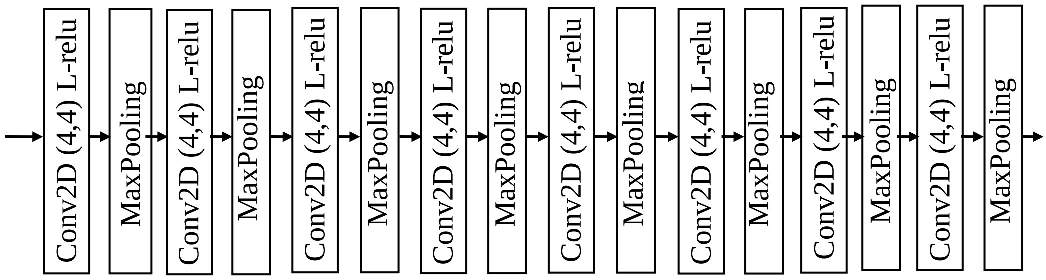


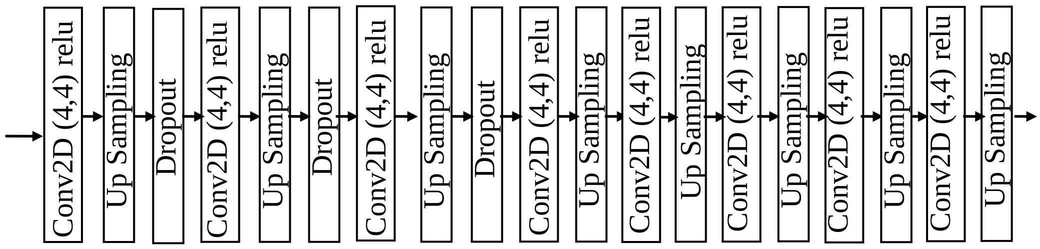
Figure 5. Configuration of the generative adversarial network (GAN)

B component images on which the L/S pattern was superimposed were input to the generator, and the generator output predicted images that were expected to contain no L/S pattern. True images were input to the discriminator as training data.

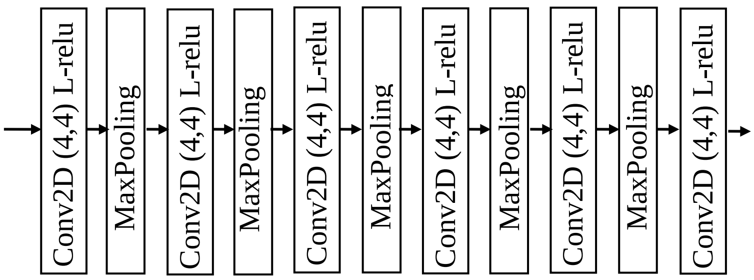
Figure 7 shows the configuration of the AE and Figure 8 shows the configuration of the encoder, decoder of the AE. The encoder and decoder of the AE consisted of five convolution layers, and a 3 x 3 kernel was used for convolution. Also in the AE, Leaky-relu was used in the encoder and relu was used in the decoder as an activation function.



(a) Encoder



(b) Decoder



(c) Discriminator

Figure 6 Configuration of the encoder decoder and discriminator of the GAN

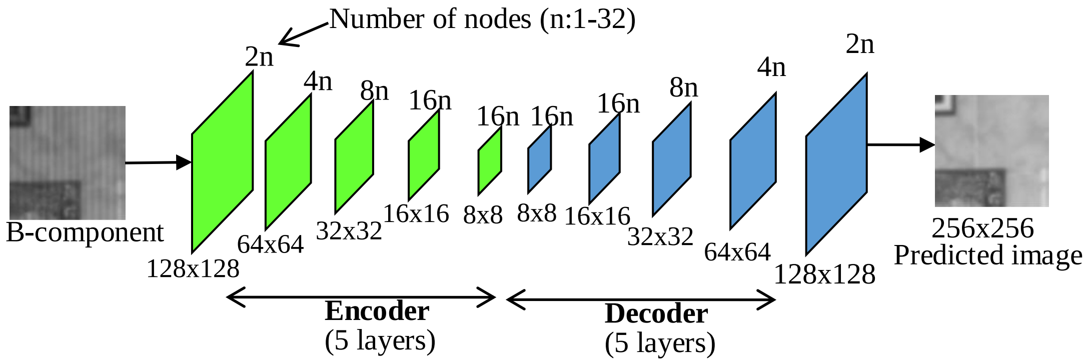


Figure 7. Configuration of the auto encoder (AE)

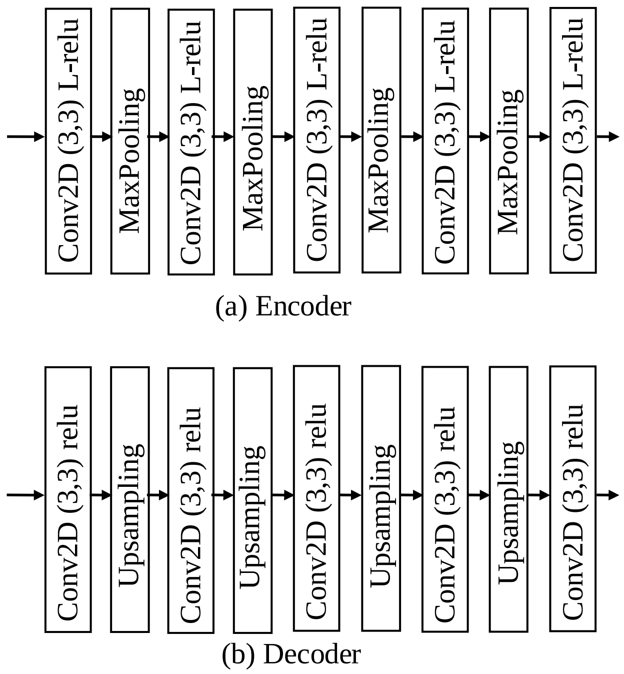


Figure 8. Configuration of the encoder and decoder of the AE

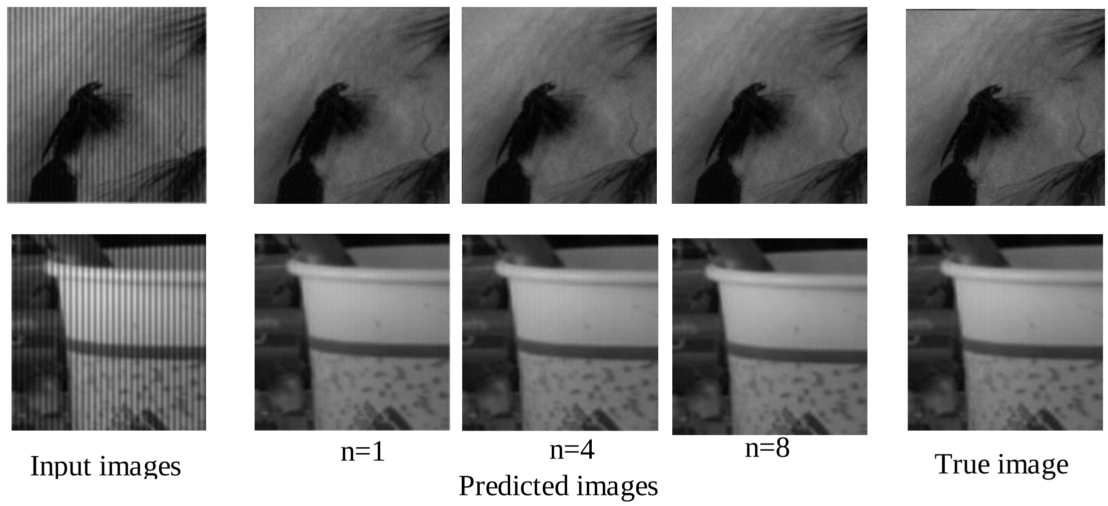


Figure 9. Examples of input, predicted, and true image

4. RESULTS AND DISCUSSION

Figure 9 shows examples of the input, predicted, and true images. As these figures show, we cannot see the L/S patterns at all in all of the predicted images, and the predicted images look almost the same as the true image.

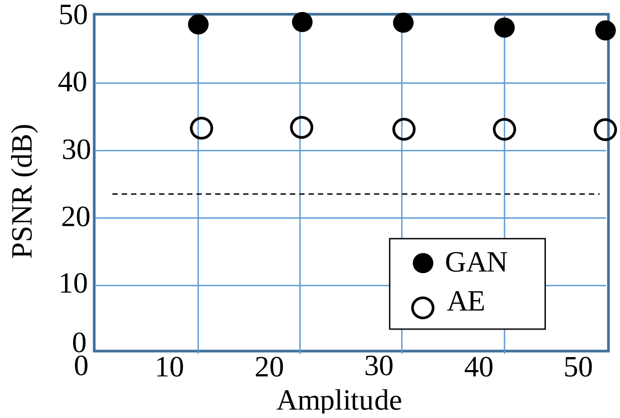


Figure 10. PSNR of image after pattern removal against original image

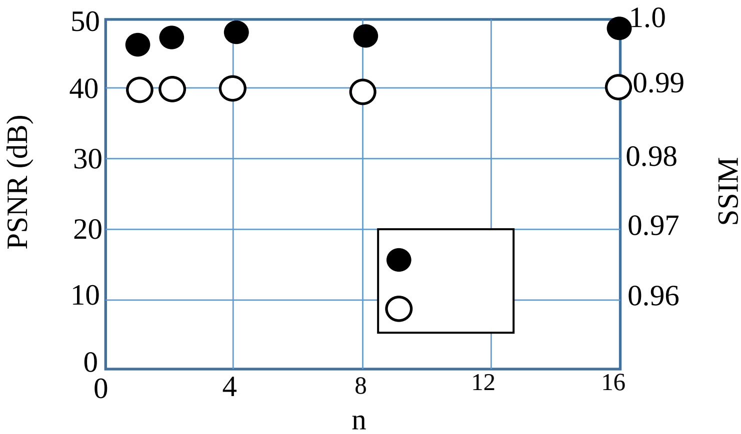


Figure 11. Dependence of PSNR and SSIM on number of nodes in middle layers

Figure 10 shows the PSNR of images after pattern removal using GAN and AE against the original image. Original image means the B component image before the pattern was superimposed. Figure 10 shows the dependence on the amplitude of the L/S pattern. The dashed line indicates the values before removing the pattern. It can be seen that the PSNR improved greatly, especially when the GAN was used. It exceeded 45, and this is a high enough value for the purpose of this study. In this figure, the PSNR was obtained as the dependence on the amplitude of the L/S pattern, and it is also seen that the PSNR decreased as the amplitude of the pattern increased. However, this decrease was slight considering the difference in amplitude of the L/S patterns. This might be because the pattern is not complicated, so the GAN can be easily trained and output an image close to the true image regardless of the amplitude of the L/S. We can see that the PSNR for the GAN was higher than that for the AE. Therefore, a GAN is better than an AE for our system.

Figure 11 shows the dependence of the PSNR and SSIM when using GAN on the number of units of the middle layers. In this figure, n for the horizontal axis is the number shown in Figure 5, and it is proportional to the number of nodes in the middle layers. From Figure 11, it is seen the PSNR and SSIM increased as the number of nodes increased; however, even when n was one, the PSNR exceeded 45, and SSIM was 0.99; both are a high enough for our system.

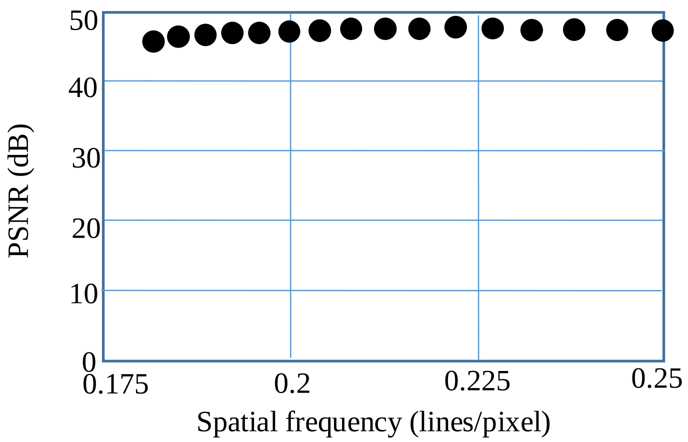


Figure 12. Dependence of PSNR on spatial frequency of L/S pattern

Figure 12 shows the dependence of the PSNR when using GAN on the spatial frequency of the L/S pattern. It shows that they are almost independent from the frequency and high enough at any frequency.

5. CONCLUSION

We studied a technique to remove periodic patterns from the RGB image of the subject taken by the camera. These patterns are projected onto the subject when its image is captured for the purpose of acquiring a depth map of the subject. Since these patterns become unnecessary after acquiring the depth map and it is just noise for RGB images, we attempted to remove these patterns using generative network, GAN and AE. From the experimental, it was shown that these periodic patterns were effectively removed by using GAN and AE to the point of being invisible. They also show that the performance of GAN is much higher than that of AE and that its PSNR and SSIM were over 45 and about 0.99, respectively. From the results, we demonstrate the effectiveness of the technique with a GAN.

In future work, we will perform similar experiments on other periodic patterns such as checkered patterns to confirm the effectiveness of the GAN method.

ACKNOWLEDGEMENTS

This study was carried out with the support of a research grant from Kanagawa Institute of Technology.

REFERENCES

[1] L. Zhang, C. Vázquez, and S. Knorr (2011) “3D-TV Content Creation: Automatic 2D-to-3D Video Conversion”, IEEE Transactions on Broadcasting, Vol. 57, No. 2, pp372-383

[2] S. Mathai, P Mathai, and K A Divya (2015) “Automatic 2D to 3D video and image conversion based on global depth map”, Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research

[3] S.Zingera, L.Do, and P.H.N.de With (2010) “Free-viewpoint depth image based rendering”, Journal of Visual Communication and Image Representation、Vol. 21, Issue, 5-6, pp533-541

[4] D. Scharstein and R. Szeliski, (2003) “High-Accuracy Stereo Depth Maps Using Structured Light”, Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Madison, Wisconsin, vol. 1, pp195-202.

[5] I. Ishii, K. Yamamoto, K. Doi, and T. Tsuji (2007) “High-speed 3D image acquisition using coded structured light projection”, Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp925-930.

[6] A. Maimone and H. Fuchs (2011) “Encumbrance-Free Telepresence System with Real-Time 3D Capture and Display Using Commodity Depth Cameras”, The 10th IEEE International Symposium on Mixed and Augmented Reality, Washington, DC, United States, pp137-146.

[7] F. Alhwarin, A. Ferrein, and I. Scholl (2014) “IR Stereo Kinect: Improving Depth Images by Combining Structured Light with IR Stereo”, Proceedings of Pacific Rim International Conference on Artificial Intelligence, pp409-421.

[8] H. Unno, S. Isaka, Y. Takashima, K. Uehira, (2015) “New Technique for Embedding Depth Information in Captured Images Using Light Beam Containing Invisible High-Frequency Patterns-Design and Preparation of New Experimental Setup With Some Comments”, Journal of Display Technology, Vol. 11, No. 2, pp136-145.

[9] I. Aizenburg, C. Butakoff, (2008) “A windowed Gaussian notch filter for quasi-periodic noise removal”, Image and Vision Computing, Vol. 26, Issue 10, pp1347-1353.

[10] F. Sur, M. Grediac (2015) “Automated removal of quasiperiodic noise using frequency domain statistics”, Journal. of Electronic Imaging, Vol. 24, No.1, 013003.

[11] I. Aizenberg, C. Butakoff (2002) “Frequency Domain Median-like Filter for Periodic and Quasi-Periodic Noise Removal”, SPIE Proceeding, Vol. 4667, pp181-191

[12] J. Varghese (2016) “Adaptive threshold based frequency domain filter for periodic noise reduction”, International Journal of Electronics and Communications, Vol. 70, Issue 12, pp1692-1701

[13] S. Ketenci, A. Gangal (2012) “Design of Gaussian star filter for reduction of periodic noise and quasi-periodic noise in gray level images”, Proceedings of the IEEE International Symposium on Innovations in Intelligent Systems and Applications.

[14] K. Uehira and H. Unno, (2020), “Technique for removing superimposed patterns on objects using GAN,” Proceedings of IEEE ICCE2020

[15] K. Uehira, (2014), "Technique of Embedding Depth Maps into 2D Images”, Journal of Electronic Science and Technology, Vol. 2, No. 1, pp95-100

AUTHORS

Kazutake Uehira received his B.S. and M.S. in Electronics in 1979 and 1981 and his Ph.D. in 1994 all from the University of Osaka Prefecture, Japan. He joined NTT Electrical Communication Laboratories in Tokyo in 1981. Since then, he has been engaged in research on imaging technologies. In 2001, he joined Kanagawa Institute of Technology, Japan, as a professor and is currently engaged in research on information embedding technology and 3D image.

Hiroshi Unno received the B.S. degree in mathematical science from Tokyo Denki University, Japan, in 1987, and the M.S. degree in information science from the Japan Advanced Institute of Science and Technology, Hokuriku, Japan, in 1996. He joined Chuo University, Japan, in 1992. Since then, he has been engaged in research on software development environments, formal methods, and computer networks. He joined the Kanagawa Institute of Technology, Japan, in 2001 and is currently engaged in research on image processing and 3D displays.