A Neural Network Based Quantitative MURA Evaluation Method Capable of Evaluating Multiple MURA on a Screen

Satomi Kidoguchi[1], Yusuke Bamba[1]

satomi.kidoguchi@eizo.com, yusuke.bamba@eizo.com
[1]EIZO Corporation, Ishikawa, Japan
Keywords: Mura, Evaluation, Deep Learning, Machine Learning, Artificial Intelligence

ABSTRACT

We developed a deep neural network-based method for evaluation of display Mura. We defined an evaluation indicator that is highly correlated with human visual evaluation. However, it could not evaluate multiple Mura on a screen properly. Therefore, we improved the evaluation indicator so that it can handle multiple Mura.

1 INTRODUCTION

Flat panel displays have luminance and color Mura. Conventionally, the evaluation of Mura has depended on human eyes. However, such a human visual evaluation has some problems that it requires proficiency of inspectors and much time and some inspectors give different evaluations. There are various shapes and positions of Mura, and we feel various Mura degree for the variation. Therefore, it is difficult to evaluate the Mura degree simply. To overcome the difficulty, many researches had conducted. For example, applying visual sensitivity characteristic to a display image [1, 2, 3, 4] or edge detection and differential filtering [3, 5] to emphasis or extract Mura, then, calculating feature quantities of Mura have performed. However, these approaches had difficulties in optimizing the feature quantity to evaluate Mura degree or must be performed under restricted conditions even if the feature quantity was decided well. So, we had proposed a machine learning approach [6, 7]. In [6], we developed a deep neural network-based Mura evaluation method. In [7], we conducted a performance test of the method comparing with subjective evaluation by human observers. In addition, we developed a new evaluation indicator to improve the correlation with human visual evaluation. However, there was a problem in that the evaluation indicator could not evaluate properly when there were multiple Mura on a screen. In order to solve the problem, we improved the evaluation indicator so that it can handle multiple Mura on a screen. In addition, we conducted a performance test of the new evaluation indicator and confirmed that the correlation with human visual evaluation was improved by using the new evaluation indicator.

2 PREVIOUS METHOD

We developed an evaluation system using a convolution autoencoder (CAE) to detect Mura automatically [6]. The CAE is one of the unsupervised learning methods and it can be used for an abnormality detection by learning with a lot of normal data [8, 9]. The CAE consists of 2 parts of an encoder and a decoder. The encoder extracts features of input data while decreasing dimensions of the input data. The decoder reconstructs input data from encoded data. In [6], we applied Mean Squared Error (MSE, Equation 1) function to the reconstruction loss provided that we performed l_2 normalization to the encoded data, and optimized parameters while decreasing the loss function. After we optimized the parameters, non-defective Mura data could be reconstructed well by the CAE. On the other hand, defective Mura data would be reconstructed worse because the CAE could not deal with the defective Mura. These properties of the CAE are useful for the abnormality detection. Therefore, we used the MSE as a main evaluation indicator.

In Equation 1, J is batch size in training phase or 1 in evaluating phase, I is input to the CAE, $E()$ is encode function, and $D()$ is decode function.

$$MSE = \frac{1}{J}\sum_{j \in J}(I_j - D(E_c(I_j)))^2 \qquad (1)$$

$$\left(E_c(I) = \frac{E(I)}{\|E(I)\|_2}\right)$$

MSE is calculated by using full screen's data. Therefore, the larger defective Mura area is, the worse the MSE becomes. On the other hand, human observers tend to focus on intensity of Mura, not area of Mura. In order to adjust the evaluation indicator of the CAE to the human visual evaluation, in the previous method [7], we modified the evaluation indicator.

First, we set to zero the squared errors at positions where the errors were lower than a threshold. Second, we calculated the summation of the squared errors after the threshold processing. Then, we divided the summation calculated at the second step by the areas where the squared errors were higher than the threshold at the first step. Finally, in order to minimize the difference between our quantitative evaluation and the human visual evaluation, we optimized the threshold while calculating the Pearson correlation coefficients to various thresholds. We defined the new evaluation indicator as $Effective\ Error$ and show the equation in

Equation 2. In Equation 2, th means the threshold and m,n mean coordinate values of the input or output.

$$Effective\ Error = \frac{1}{Area_{th}} \sum_{m,n} SE_{th}(m,n) \qquad (2)$$

$$\left(\begin{array}{l} SE_{th}(m,n) = \begin{cases} 0\,, & ((I(m,n) - D(E_c(I(m,n))))^2 < th) \\ (I(m,n) - D(E_c(I(m,n))))^2\,, & (oherwise) \end{cases} \\ Area_{th} = \sum_{m,n} \begin{cases} 0\,, & ((I(m,n) - D(E_c(I(m,n))))^2 < th) \\ 1\,, & (otherwise) \end{cases} \end{array} \right)$$

3 PROPOSED METHOD

The $Effective\ Error$ was highly correlated with the results of human visual evaluation when there was single defective Mura on a screen but it could not evaluate properly when there were multiple defective Mura on a screen. This is because the $Effective\ Error$ is calculated by using an average value of the squared errors above a threshold and it was affected by multiple Mura which have some intensities on a screen. As a result, even an originally defective Mura may be evaluated as non-defective Mura.

Therefore, we improved the $Effective\ Error$ so that it can evaluate multiple Mura properly. First, as with the conventional process, we set to zero the squared errors at positions where the errors are lower than a threshold. Second, we applied connected-component labeling (8-connectivity) for extracting multiple Mura. Finally, we calculated the $Effective\ Error$ for each labeled region. The calculation equation is shown in Equation 3. In Equation 3, as with Equation 2, th means the threshold. $\mathrm{m_i}, n_i$ mean coordinate values of the each labeled region, and $\mathrm{Area}_{\mathrm{th}_i}$ means the area of the labeled region. By calculating the $Effective\ Error$ for each region, the results of each Mura are not affected by each other.

$$Effective\ Error_i = \frac{1}{Area_{th_i}} \sum_{m_i,n_i} SE_{th}(m,n) \qquad (3)$$

$$\left(Area_{th_i} = \sum_{m_i,n_i} \begin{cases} 0\,, & ((I(m,n) - D(E_c(I(m,n))))^2 < th) \\ 1\,, & (otherwise) \end{cases} \right)$$

4 EXPERIMENT

We prepared pseudo Mura images and conducted human visual evaluation by our inspectors and Neural Network evaluation same as the previous study [7]. We prepared 52 pseudo Mura images, including images which include multiple Mura. These multiple Mura images were the most difficult to evaluate for the previous method. An example of a pseudo multiple Mura images are shown in Figure 1. We displayed pseudo Mura images on an LCD display (27.0 inches, 2,560x1,440 pixels, with maximum brightness of 400cd/ m^2 manufactured by EIZO Corporation) in a dark room.

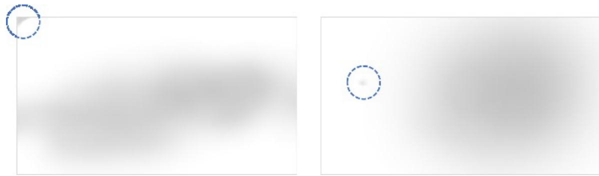


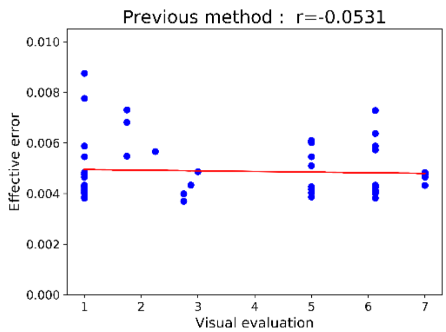
Fig. 1 Some examples of pseudo multiple Mura
The pseudo Mura image had small Mura (in blue circle) and large Mura.

For human visual evaluation, we recruited 4 observers. They are each licensed to inspect displays quality that our corporation defined have experience of between 6 years and 25 years. Each observer scored from 1 to 12 points for each displayed image where 1 is the worst and 12 is the best. Then we averaged the scores each of the 4 observers achieved when they evaluated the various Mura.

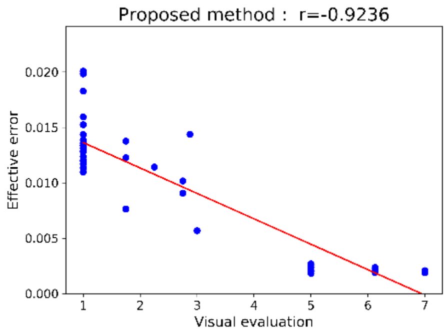
For quantitative evaluation, we displayed the pseudo Mura images on the LCD display and captured them with a two-dimensional luminance colorimeter to get luminance and color images in a dark room. Then, we input the data to the learned CAE to evaluate the pseudo Mura images. Finally, we calculated the $Effective\ Error$ of the previous method and the proposed method for comparison.

5 RESULTS

We checked the Pearson correlation coefficient between the $Effective\ Errors$ and the results of human visual evaluation. Figure 2 (a) shows the correlation between the previous method and human visual evaluation. Figure 2 (b) shows the correlation between the proposed method and human visual evaluation. We picked up the most difficult pseudo Mura images for the previous method for performance checks. Therefore, the correlation coefficient was quite low at -0.053. When there are multiple Mura on a screen, the previous method averages squared errors of all areas above the threshold and tend to evaluate better than actual results. On the other hand, we got a high correlation of -0.92 by the proposed method. The proposed method calculated the $Effective\ Error$ for each Mura on a screen. Therefore. it could be correlated with human visual evaluation.



(a) Results of the previous method



(b) Results of the proposed method

Fig. 2 Pearson correlation between our system and visual evaluations.
The graphs had a visual rating of up to 7 out of 12.

6 CONCLUSION

In previous study, we developed the Mura evaluating system based on an unsupervised learning method of the CAE and we used the MSE as the evaluation indicator. Evaluation results of the previous method were greatly affected by the area of Mura because the MSE was averaged the squared errors across the entire image area. On the other hand, human observers do not tend to consider the area of Mura when they evaluate Mura. In order to adjust the evaluation indicator of the CAE to the human visual evaluation, we developed the $Effective\ Error$ to suppress the difference and to improve the correlation. The $Effective\ Error$ had higher correlation than MSE and it was effective for general Mura. However, in certain cases, such as when there are multiple Mura on a screen, the $Effective\ Error$ could not handle the multiple Mura. Therefore, we modified the $Effective\ Error$ so that it can handle the multiple Mura. Through the performance test, we got higher correlation with human visual evaluations by using the modified $Effective\ Error$.

Our conclusions are as follows.

- We developed the new evaluation indicator for the Mura evaluating system based on the CAE that can be applied to multiple Mura on a screen.
- We conducted an experiment with human visual evaluations and compared correlations between the human visual evaluation and each of the previous method and the proposed method.
- The proposed method obtained higher correlation than the previous method.

REFERENCES

[1] S. Wang, Z. Jhang, and C. Wen. "A Mura Metric Based on Human Vision Models", J. SID, Vol. 37, No. 1, pp. 291-294 (2006).

[2] T. Asano, Y. Takagi, T. Kondo, J. Yao and W. Liu, "Image quality evaluation based on contrast sensitivity function," 2011 IEEE International Conference on Mechatronics and Automation, pp. 658-663 (2011)

[3] K. Nagamine, S. Tomioka, T. Tamura, and Y. Shimpuku. "A Quantitative Evaluation Method for Luminance and Color Uniformity of a Display Screen Based on Human Perceptions", Proc. IDW '11, pp. 341-344 (2011).

[4] K. Nagamine and S. Tomioka. "A Quantitative Mura Evaluation Method that Depends on Viewing Distance", Proc. IDW '12, pp. 1975-1976 (2012).

[5] T. Ichikawa and H. Nakaya. "Apparatus and method for automatically detecting non-uniformity of flat panel display", JP2014127062A (2012).

[6] K. Tsutsukawa, N. Tabata, and Y. Bamba, "Speedy and Quantitative Evaluation of Luminance Non-Uniformity Based on Deep Neural Networks", J. SID, Vol. 50, No. 1, pp. 969-972 (2019).

[7] K. Tsutsukawa, M. Kobayashi, and Y. Bamba, "Neural Network Based Quantitative Evaluation of Display Non-Uniformity Corresponds Well with Human Visual Evaluation", J. SID, pp. 1214-1217 (2020).

[8] Y. Xia, X. Cao, F. Wen, G. Hua and J. Sun, "Learning Discriminative Reconstructions for Unsupervised Outlier Removal," 2015 IEEE International Conference on Computer Vision (ICCV), pp. 1511-1519 (2015).

[9] C. Aytekin, X. Ni, F. Cricri and E. Aksu, "Clustering and Unsupervised Anomaly Detection with l2 Normalized Deep Auto-Encoder Representations," 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1-6 (2018).