VNU Journal of Science:
Computer Science and Communication Engineering

Journal homepage: http://www.jcsce.vnu.edu.vn/index.php/jcsce

Original Article

Developing an Objective Visual Quality Evaluation Pipeline for 3D Woodblock Character

Le Cong Thuong[1], Viet Nam Le[1], Thi Duyen Ngo[1], Seung-won Jung[2], Thanh Ha Le[1]*

[1] *VNU University of Engineering and Technology, 144 Xuan Thuy, Cau Giay, Hanoi, Vietnam*
[2]*Department of Electrical Engineering, Korea University, Seoul, South Korea*

Received 05 February 2024
Revised 12 June 2024; Accepted 10 December 2024

Abstract: Vietnamese feudal dynasty woodblocks are invaluable national treasures, but many have been lost or damaged due to wars and poor conditions of preservative environments. Fortunately, 2D-printed papers of the damaged or lost woodblocks have been well-preserved, allowing for reconstruction their 3D digital version. To ensure accurate reconstruction of 3D woodblocks, it is essential to have a reliable alignment method that closely matches human visual perception. In this paper, we introduce an automatic pipeline for objective visual quality evaluation of woodblock characters. The pipeline includes two components: the first shifts the quality evaluation from 3D domain to 2D domain by employing orthogonal projection to transform a 3D mesh woodblock character model into a 2D depth map image with minimum information loss. The second utilizes established 2D perceptual metrics, which closely align with human visual perception, to evaluate the 2D depth map. Our evaluation demonstrates that features of these proposed perceptual metrics employed in the pipeline can effectively characterize the visual appearance of the woodblock character based on our prepared dataset. Additionally, the experiments presented in the paper also demonstrate that these metrics can sensitively detect degradation levels in both the foreground and background components of a woodblock character.

Keywords: depthmap image, learned perceptual metrics, woodblock evaluation

1. Introduction

Woodblocks from feudal dynasties, illustrated in Figure 1, are a national treasure, particularly in East Asia, including Korea, China, Vietnam, and Japan. In Vietnam, the Nguyen Dynasty's printing woodblocks have gained international recognition, as UNESCO included

*Corresponding author.
E-mail address: ltha@vnu.edu.vn
https://doi.org/10.25073/2588-1086/vnucsce.2248

them in the World Memory Program in 2009 [1]. However, many have been lost or damaged due to war, weather, or poor physical conditional preservation. Fortunately, the corresponding 2D-printed papers are well-preserved, providing an opportunity for 3D digital reconstruction. Quality assessment of 3D character woodblocks, illustrated in Figure 2, is imperative, with the primary challenge being the accurate measurement of the reconstructed woodblock's quality in comparison to the ground truth.

Figure 1. Examples of physic 3D woodblock and corresponding 2D prints.

Subjective assessments of woodblocks [1] require human experts and cannot be automated. In contrast, current objective metrics such as Chamfer Distance (CD) [2] and Hausdorff Distance (HD) [3] are widely used for 3D reconstruction problems. These metrics measure the similarity between two point sets by assigning point pairs based on the nearest neighbor search. Other metrics incorporate low-level attributes such as differences in normal orientations [4] or curvature information [5]. However, these low features may not fully align with human perception as shown in [4, 6].

Therefore, the paper proposes an objective visual quality evaluation pipeline for measuring the quality of 3D woodblock characters. The pipeline includes 2 main components: the first component maps a 3D woodblock character to a depth map using orthogonal projection, and the second component employs perceptual 2D metrics such as Deep Image Structure and Texture Similarity (DISTS) [7] or Learned Perceptual Image Patch Similarity (LPIPS) [8] to assess the quality of the resulting depth map. The approach utilizes available perceptual metrics for images designed to align with human perception, ensuring an objective evaluation of the reconstructed woodblock's visual appearance. The pipeline presents that the set of features from the DISTS or LPIPS metric used is sufficient to capture the visual appearance of woodblock character depth under our prepared dataset. Additionally, our experiments show that the pipeline is also sensitive to different types of degradation in both foreground and background regions.

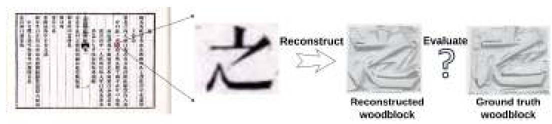

Figure 2. The evaluation problem: How to measure the quality of the generated character woodblock with the ground-truth one?

2. Related work

This section provides an overview of common objective and subjective metrics used for evaluating 3D reconstructed objects. As our pipeline involves evaluating 3D objects through 2D images, we also examine common metrics used for 2D image evaluation.

Subjective 3D quality assessment. Assessing the subjective quality of 3D models relies on human observers participating in experiments where they evaluate the perceived quality of distorted or modified models. Subjective quality assessment for 3D models requires human observers to participate in experiments where they evaluate the perceived quality of distorted objects. In the woodblock domain, Ngo et al. [1] proposed assessment criteria encompassing both objective and subjective aspects, ensuring preservation requirements for the entire woodblock. However, this approach necessitates expert involvement and lacks automation capabilities. Additionally, Apollonio et al. [9] suggest a validation pipeline that comprises six distinct categories: data collection, data acquisition, data analysis, data interpretation, and data representation. The pipeline's success depends on adhering to different standards and assumptions at each stage to minimize subjective results in the 3D reconstruction output. Overall, subjective quality assessment for woodblocks requires human experts and cannot be automated.

Objective 3D quality assessment. Chamfer Distance and Hausdorff Distance are commonly used as purely geometric error measures in point cloud tasks. These metrics use the nearest neighbor search to establish point pairs. Other metrics incorporate low-level features like normal orientation differences [4] or curvature information [5], but they may not fully align with human perception. As a result, these metrics often yield poor results when used to evaluate subjective datasets [4, 6].

Objective 2D quality assessment. Objective 2D quality assessment, or objective image quality assessment (IQA) is a field that focuses on developing computational models to predict the perceived quality of visual images. Full-reference IQA methods compare a distorted image to a complete reference image. According to [10], full-reference methods can be divided into five categories: error visibility, structural similarity, information-theoretic, learning-based, and fusion-based methods. Error visibility methods apply a distance measure directly to pixels, such as mean squared error (MSE). Structural similarity methods are constructed to measure the similarity of local image structures, often using correlation measures, and typical metrics include the Structural Similarity Index (SSIM) [11]. Information-theoretic methods measure some approximations of the mutual information between the perceived reference and distorted images, typical visual information fidelity (VIF) metrics [12]. Learning-based methods learn a metric from a training set of images and corresponding perceptual distances using supervised machine learning methods. Fusion-based methods combine existing IQA methods. Recently, learning-based methods have been rapidly developed, and deep neural network-based metrics such as LPIPS [8] and DISTS [7] have offered the best overall performance based on learned perceptual features [10].

3. Methods

As demonstrated in Section 2, conventional 3D quality metrics fall short of woodblock character evaluation. However, based on both practical woodblock creation and observations from [13], we have discovered a key property: due to the process of constructing printing woodblocks, the woodblock-making technique in Asia begins with a smooth rectangular block. This block is then carved to create the slope of the characters. From the direction of the carving, which is orthogonal to the surface of the woodblock characters, all the information can be observed. This allows us to represent them as 2D depth maps with minimal detail loss. This opens the door to leveraging established 2D perceptual metrics for evaluation, which closely align with human perception.

Figure 3 illustrates our automatic objective visual pipeline, which consists of two key steps:

- **3D-to-Depthmap mapping**: The input mesh woodblock characters $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$ undergo orthogonal projection and quantization through module $\boldsymbol{P}$, resulting in the generation of orthographic depth map images of woodblock characters $\boldsymbol{D_1}$ and $\boldsymbol{D_2}$, respectively. This process facilitates the transition from the 3D to the 2D domain for shift evaluation.

- **Perceptual evaluation**: The depth map images of woodblock characters $\boldsymbol{D_1}$ and $\boldsymbol{D_2}$, are subject to evaluation using the learned perceptual metric module $\boldsymbol{M}$. This module is specifically designed for 2D images and generates a score $\boldsymbol{S}$, which quantifies the similarity between the two woodblock characters.

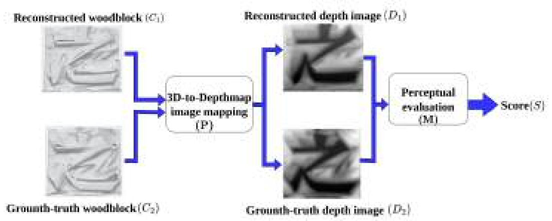

Figure 3. The comprehensive automatic objective visual evaluation pipeline and notation system for woodblock analysis.

3.1. 3D-to-Depthmap Mapping

The mapping process is carried out in a simulation environment, primarily using the Open3D library, when depth maps are generated using orthographic projection. In this environment, depth maps are generated using orthographic projection. The projection direction is carefully aligned to be perpendicular to the surface of the woodblock character. To ensure consistency and accuracy, the distance from the camera surface to the character surface is maintained at a fixed 12 units. Figure 4 illustrates the orthographic projection. The resulting depth map assigns a distance value to each pixel, representing the distance between the camera plane and the surface in the orthogonal direction. This distance value is then mapped and quantized to the range of 0 to 255. The significance of this 2D depth map representation is that it enables evaluation using 2D perceptual metrics, shifting the focus from 3D to 2D analysis.

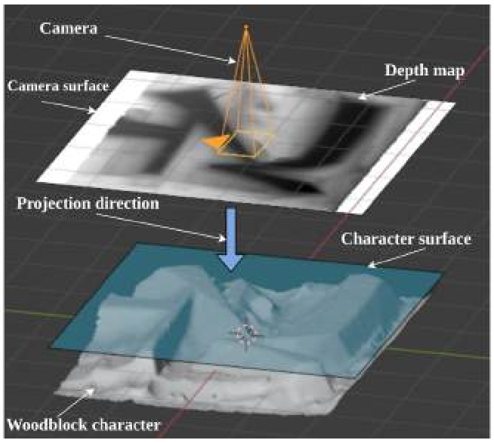

Figure 4. Illustration of orthographic projection of the woodblock character, and notation system

3.2. Perceptual Evaluation

Deep perceptual metrics are becoming increasingly dominant in the full reference metric problem. These metrics are the features in a neural network that is trained with millions of data samples, allowing it to obtain better statistics of the object compared to other metric types such as error visibility metrics like MSE and SSIM [11] or information measures such as visual information fidelity (VIF) [12]. Step 2 of our proposed pipeline focuses on two commonly used deep perceptual metrics: DISTS [7] and LPIPS [8], which offer more comprehensive descriptions that closely align with human visual perception.

- **DISTS metric**. The Deep Image Structure and Texture Similarity (DISTS) metric combines structural and textural information to evaluate image quality. It transforms the reference and distorted images using a variant of the VGG network [14], which is pre-trained for object recognition on the ImageNet database. The transformation involves spatial convolution, half-wave rectification, and downsampling, with modifications to ensure the representation is aliasing-free and injective. The resulting representation captures both local structural distortions (such as noise or blur) and textural resampling. DISTS uses learnable weights optimized to balance structural and texture similarities provides a robust assessment of image quality that correlates well with human perception.

- **LPIPS metric**. The Learned Perceptual Image Patch Similarity (LPIPS) metric uses a pre-trained deep neural network, typically VGG, to extract features from reference and distorted images. It computes the perceptual distance by calculating a weighted sum of the distances between the feature maps of image patches. The weights are learnable and optimized to match human perceptual judgments. LPIPS is trained on the Berkeley-Adobe Perceptual Patch Similarity (BAPPS) dataset, which includes a wide variety of image distortions, allowing it to generalize well across different types of visual impairments. This method leverages the deep features' ability to capture complex aspects of human perception, providing a more accurate assessment of image quality compared to traditional metrics.

In the paper, we use the DISTS and LPIPS because of the following reasons:

- Compared to traditional metrics like PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index), these chosen metrics leverage deep learning and capture high-level features, and robustness across diverse image types compared to other metrics as shown [7, 10].
- Compared to other deep perceptual metrics, despite these metrics not latest metrics, these metrics have been extensively validated in the literature and adopted by the computer vision community for perceptual similarity [7, 10, 15, 16].

4. Evaluation

In this section, we evaluate important aspects of 3D problem reconstruction metrics such as degradation sensitivity, the impact of woodblock components, and capturing the visual appearance of woodblock characters. We do not aim to prove the effectiveness of DISTS or LPIPS metrics, widely recognized as reliable metrics for evaluating image quality, as they are designed to align with human perception. Our goal is to strengthen the pipeline and provide a more solid analysis of the problem.

4.1. Datasets

For the evaluations in subsequent sections, a dataset, denoted as HMI-TRIPLE-WB, utilizing a subset of the 3D digital woodblock collection from the Human-Interaction Laboratory at the University of Engineering and Technology. This original collection comprises high-resolution mesh models of entire woodblocks, each with a resolution ranging from 40 million to 100 million vertices, captured using the ATOS Q 12M machine. From this extensive collection, individual characters were carefully selected and cropped to compose the HMI-TRIPLE-WB dataset. Specifically, this dataset includes a total of 24 samples, each derived from five historical woodblocks, which are part of the Dai Nam Thuc Luc historical records. Each sample within the dataset contains three woodblock characters and fulfills the following two requirements:

- Characters in each sample share the same textual symbol.
- All characters within each sample originate from the same whole woodblock.

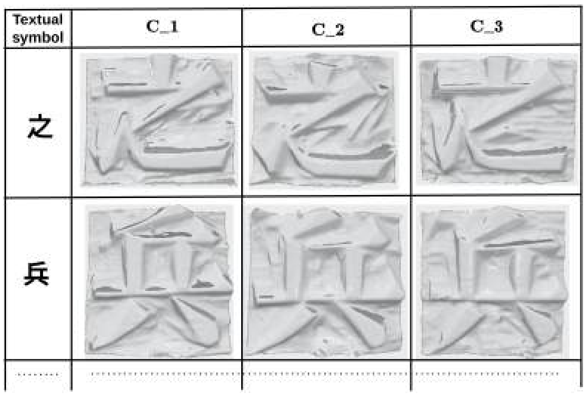

Figure 5. Each row is a sample of the HMI-TRIPLE-WB dataset.

Figure 5 displays two samples from the HMI-TRIPLE-WB dataset, showcasing all the characters contained within each sample.

Furthermore, as indicated in [1], a comprehensive quantitative analysis requires assessing both the intentional and unintentional details imparted by the carving artists. In the context of woodblock characters, these details are present in two primary regions: the foreground, which manifests the artists' intentions, and the background, which subtly reveals unintentional details. The foreground predominantly exhibits the character itself, whereas the background contributes contextual details and encompasses the area surrounding the main character. Given the absence of pre-existing labels to distinguish between these two regions, a manual labeling process was employed to. Figure 6 provides a detailed of the structural elements of a woodblock character.

4.2. Which Region in the Woodblock Affects the Score of Metrics the Most?

In this section, we emphasize the importance of the background in addition to the foreground when evaluating the quality of woodblock characters. According to a woodblock expert, the background provides the "spirit" of the woodblock. Therefore, we examine the influence of various regions of woodblock characters, including the background, on metric scores. We achieve this by removing either the foreground or background region of the characters and comparing the metric scores of the original and modified versions. We also note that discarding the foreground region removes the curvature, smoothness, and inner geometry information of the character, but not the outer geometry.

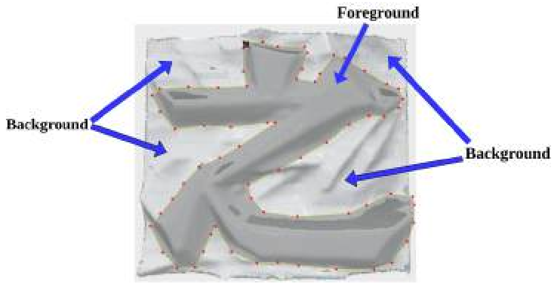

Figure 6. The main components of each woodblock character. The foreground is the region inside the red-dot borderline.

With each sample $1 \leq i \leq N$, the HMI-TRIPLE-WB can be presented to set of $\{C_1^i, C_2^i, C_3^i\}_{1\leq i\leq N}$ where C_j^i denotes the j^{th} character in the set i^{th} of our dataset. With the proposed pipeline including 3D-to-Depthmap mapping ($\boldsymbol{P}$) and perceptual evalution ($\boldsymbol{M}$), the analysis proceeds through the following steps:

1. **Whole character difference calculation:** Compute the score differences across the characters based on the whole character comparison. This is achieved by applying Equation 1:

$$\Delta s = \frac{1}{\mathbf{N}} \sum_{i=1}^{\mathbf{N}} \left| \mathbf{M}\left(\mathbf{P}(C_1^i), \mathbf{P}(C_2^i)\right) - \mathbf{M}\left(\mathbf{P}(C_1^i), \mathbf{P}(C_3^i)\right) \right|. \tag{1}$$

2. **Foreground difference calculation:** Extract and compare the foregrounds of the characters. Denote the foregrounds of C_1^i, C_2^i, and C_3^i as F_1^i, F_2^i, and F_3^i respectively. Compute the foreground

differences using Equation 2:

$$\Delta sf = \frac{1}{N}\sum_{i=1}^{N}\left|\boldsymbol{M}\left(\boldsymbol{P}(F_1^i), \boldsymbol{P}(F_2^i)\right) - \boldsymbol{M}\left(\boldsymbol{P}(F_1^i), \boldsymbol{P}(F_3^i)\right)\right|. \quad (2)$$

3. **Background difference calculation:** Similarly, for each sample, extract and compare the backgrounds of the characters, denoted as B_1^i, B_2^i, and B_3^i for each C_1^i, C_2^i, and C_3^i respectively. Compute the background differences using Equation 3:

$$\Delta sb = \frac{1}{N}\sum_{i=1}^{N}\left|\boldsymbol{M}\left(\boldsymbol{P}(B_1^i), \boldsymbol{P}(B_2^i)\right) - \boldsymbol{M}\left(\boldsymbol{P}(B_1^i), \boldsymbol{P}(B_3^i)\right)\right|. \quad (3)$$

The significance of a region can be determined by the difference between the Δs, Δsf, and Δsb values, with smaller differences indicating greater significance. Figure 7 illustrates the results for each data type. The experiment reveals two key findings that align with human perception. Firstly, the difference between solely the character and solely the background regions differs from that of the original character, indicating that the pipeline relies on both regions to make decisions, rather than solely relying on either region. Secondly, the mean difference between the original and solely background data is smaller than the mean difference between the original and solely foreground data ($|\Delta sb - \Delta s| < |\Delta sf - \Delta s|$), suggesting that the background region has a more pronounced impact on the metrics.

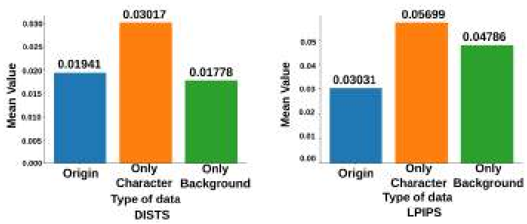

Figure 7. The mean differences in origin, only foreground, and only background data type, as measured by the DISTS and LPIPS metrics.

Table 1. Parameters of different simulated degradations

Type	Parameter	Values
Cut out	Number of holes	[1, 2, 3, 4, 5]
Gaussian noise	Variance of normal dist	[0.015, 0.02, 0.025, 0.03, 0.035]
Impulse noise	Percent of impulsed points	[0.1, 0.15, 0.2, 0.25, 0.3]
Uniform noise	Sampling range of uniform dist	[0.015, 0.02, 0.025, 0.03, 0.035]

4.3. Impact of Location, Strength Level of Different Types of Degradation on Metrics

The inevitable digital noise from scanning sensors will contribute to substantial variations in data points. As shown in [17, 18], these variations can be summarized into 25 corruption patterns and categorized into four groups based on common types of corruption: weather, noise, density, and transformation. For the 3D woodblock digitization process, we focus on two types: noise and density. Specifically, we address four types of degradation in this research: uniform noise, impulse noise, Gaussian noise, and cut-out. In this section, we analyze the effects of various degradation types on metric scores for each region of the woodblock characters. The procedure in this section is the same as described in subsection 4.2. However, instead of removing the foreground or background region, we introduce different types of degradation to either the foreground region, background region, or the entire woodblock character. The level of degradation for each type of noise is controlled by modifying parameter values, which are described in Table 1.

The results of our experiments are presented in Figure 8. We observed that the DISTS metric exhibits higher sensitivity than the LPIPS metric to different levels of noise, particularly when degradations are added to the background region. This finding aligns with the observations made in subsection 4.2, where these metrics demonstrated a greater emphasis on the background. Importantly, the behavior of these metrics varies under different types of degradation. For instance, in the case of impulse noise, the DISTS metric shows higher sensitivity

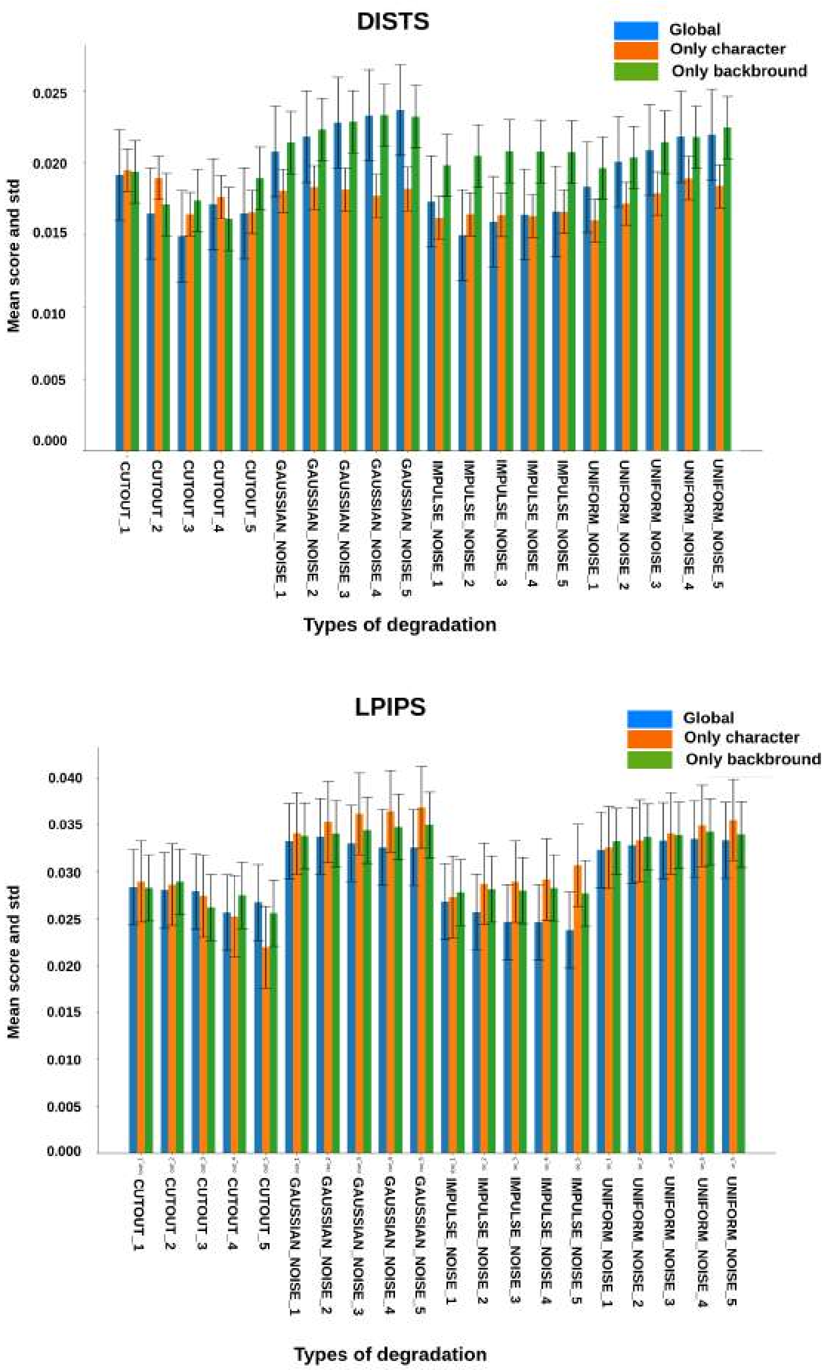

Figure 8. The impacts of each degradation with different levels in each component of woodblock into the score of two metrics.

to degradation in the background, while the LPIPS metric is more influenced by degradation in the foreground.

4.4. Feature Inversion

This section investigates whether the features extracted from the DISTS or LPIPS metric can effectively capture the visual appearance of woodblock character depth images as perceived by the human visual system. Similar approaches in prior works [7, 19, 20], we leverage image synthesis to address this question. We optimize a total uniform noise image to match the target image by leveraging the feature set from either DISTS or LPIPS. If these features encompass a sufficiently comprehensive set of statistics, the synthesized image should be perceptually indistinguishable from the original, particularly in terms of preattentive judgments [21]. Specifically, the features of DISTS or LPIPS are used to calculate the image synthesis by addressing the equation detailed in Equation 4:

$$\begin{aligned} D^* &= \arg\min_{D_2} D\left(D_1, D_2\right) \\ &= \arg\min_{D_2} \sum_{i,j} \left(F_{i,j}(D_1) - F_{i,j}(D_2)\right)^2, \end{aligned} \tag{4}$$

where D_1 is the target texture image, D^* is the synthesized woodblock image obtained by gradient descent optimization from a random initialization and $F_{i,j}(D_1)$ and $F_{i,j}(D_2)$ are the of feature map of channel i in the stage j from DISTS or LPIPS. Figure 9 showcases the results of this synthesis procedure. The hyperparameters for feature map extraction, including the channels and stages, follow the default settings as specified in the respective papers on DISTS and LPIPS. It is apparent from the figure that the synthesized images are perceptually indistinguishable from the original images, based on preattentive judgments. This outcome indicates that the set of features extracted from the DISTS or LPIPS metric is sufficient to capture the visual appearance of woodblock character depth images.

Figure 9. The output of image inversion of each method, the left column is ground-truth, the mid column is the results of DISTS metric and the right is the results of LPIPS metric.

5. Conclusion and Future Work

We introduce an objective visual quality evaluation pipeline for the 3D woodblock character that maps a 3D character woodblock to a 2D image using orthogonal projection, which incorporates advanced deep perceptual metrics such as DISTS and LPIPS. We found that the objective pipeline is sufficient to capture the visual appearance of woodblock character depth by using learned perceptual metrics. Additionally, our analysis reveals that our pipeline assigns greater weight to background regions than to foreground regions and that the behavior of these metrics varies significantly for different types of noise.

In this study, we utilized the HMI-TRIPLE-WB dataset for our experiments. To get rid of bias existing in the dataset, such as characters being carved by the same artisan and the limited number of samples, we plan to expand and diversify the dataset. For future work, we will 3D scan additional woodblocks to increase the number and variety of characters, and include more varied types of degradation and damage. This expansion will enhance the diversity and representativeness of the dataset, thereby reducing potential biases and improving the robustness of our experiments.

Building upon our work, we have integrated the proposed metric pipeline into a 3D reconstruction framework. This enables the complete reconstruction of lost character woodblocks. The evaluation pipeline is not only applicable to Vietnamese woodblocks but also paves the way for preserving similar ancient woodblocks throughout East Asia. This demonstrates the significant potential of our pipeline for broader cultural heritage preservation efforts.

Acknowledgment

This work has been supported by Korea-Vietnam joint research project number NT/KR/21/17 - "AI-based Application for 3D-Model Restoration of Vietnamese and Korean Cultural Heritage".

This work of Seung-won Jung was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT Future Planning, under Grant NRF-2020K1A3A1A05103460.

References

[1] T. D. Ngo, T. M. Tran, T. C. Ma, X. H. Nguyen, T. H. Le, Royal Printing Woodblocks of Nguyen Dynasty: 3D Reconstruction for Digital Preservation, International Association for Printing Woodblocks (IAPW) 2019, pp. 333–348Https://doi.org/10.1145/3532854.

[2] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, H. C. Wolf, Parametric Correspondence and Chamfer Matching: Two New Techniques for Image Matching, in: Proceedings: Image Understanding Workshop, Science Applications, Inc, 1977, pp. 21–27, https://dl.acm.org/doi/10.5555/1622943.1622971.

[3] A. A. Taha, A. Hanbury, An Efficient Algorithm for Calculating the Exact Hausdorff Distance, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 37, No. 11, 2015, pp. 2153–2163, http://dx.doi.org/10.1109/TPAMI.2015.2408351.

[4] E. Alexiou, T. Ebrahimi, On the Performance of Metrics to Predict Quality in Point Cloud Representations, in: Applications of Digital Image Processing XL, Vol. 10396, SPIE, 2017, pp. 282–297, https://doi.org/10.1117/12.2275142.

[5] G. Meynet, J. Digne, G. Lavoué, PC-MSDM: A Quality Metric for 3D Point Clouds, in: 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), IEEE, 2019, pp. 1–3, http://dx.doi.org/10.1109/QoMEX.2019.8743313.

[6] H. Su, Z. Duanmu, W. Liu, Q. Liu, Z. Wang, Perceptual Quality Assessment of 3D Point Clouds, in: 2019 IEEE International Conference on Image Processing (ICIP), IEEE, 2019, pp. 3182–3186, http://dx.doi.org/10.1109/ICIP.2019.8803298.

[7] K. Ding, K. Ma, S. Wang, E. P. Simoncelli, Image Quality Assessment: Unifying Structure and Texture Similarity, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 44, No. 5, 2020, pp. 2567–2581, http://dx.doi.org/10.1109/TPAMI.2020.3045810.

[8] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, O. Wang, The Unreasonable Effectiveness of Deep Features as a Perceptual Metric, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 586–595, http://dx.doi.org/10.1109/CVPR.2018.00068.

[9] F. I. Apollonio, et al., Classification Schemes and Model Validation of 3D Digital Reconstruction Process, in: Proceedings of the 20th International Conference on Cultural Heritage and New Technologies, 2015.

[10] K. Ding, K. Ma, S. Wang, E. P. Simoncelli, Comparison of Full-Reference Image Quality Models for Optimization of Image Processing Systems, International Journal of Computer Vision, Vol. 129, 2021, pp. 1258–1281, http://dx.doi.org/10.1007/s11263-020-01419-7.

[11] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image Quality Assessment: From Error Visibility to Structural Similarity, IEEE Transactions on Image Processing, Vol. 13, No. 4, 2004, pp. 600–612, http://dx.doi.org/10.1109/TIP.2003.819861.

[12] H. R. Sheikh, A. C. Bovik, Image Information and Visual Quality, IEEE Transactions on Image Processing, Vol. 15, No. 2, 2006, pp. 430–444, http://dx.doi.org/10.1109/TIP.2005.859378.

[13] N. Su, Ve Quy Trinh Khac In Moc Ban Truyen Thong O Viet Nam, Di San Van Hoa 2016, pp. 1–5.

[14] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv Preprint arXiv:1409.1556 (2014).

[15] M. Kettunen, E. Härkönen, J. Lehtinen, E-LPIPS: Robust Perceptual Image Similarity Via Random Transformation Ensembles, arXiv Preprint arXiv:1906.03973 (2019).

[16] S. H. Park, Y. S. Moon, N. I. Cho, Perception-Oriented Single Image Super-Resolution Using Optimal Objective Estimation, in: Proceedings of

the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 1725–1735, http://dx.doi.org/10.1109/CVPR52729.2023.00172.

[17] S. Li, Z. Wang, F. Juefei-Xu, Q. Guo, X. Li, L. Ma, Common Corruption Robustness of Point Cloud Detectors: Benchmark and Enhancement, arXiv Preprint arXiv:2210.05896 (2022).

[18] J. Sun, Q. Zhang, B. Kailkhura, Z. Yu, C. Xiao, Z. M. Mao, Benchmarking Robustness of 3D Point Cloud Recognition Against Common Corruptions, arXiv Preprint arXiv:2201.12296 (2022).

[19] D. J. Heeger, J. R. Bergen, Pyramid-Based Texture Analysis/Synthesis, in: Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, 1995, pp. 229–238, http://dx.doi.org/10.1145/218380.218446.

[20] L. Gatys, A. S. Ecker, M. Bethge, Texture Synthesis Using Convolutional Neural Networks, Advances in Neural Information Processing Systems, Vol. 28, , https://dl.acm.org/doi/10.5555/2969239.2969269 (2015).

[21] J. Malik, P. Perona, Preattentive Texture Discrimination with Early Vision Mechanisms, JOSA A, Vol. 7, No. 5, 1990, pp. 923–932, http://dx.doi.org/10.1364/JOSAA.7.000923.