Supplementary Material: Deep Halftoning with Reversible Binary Pattern

Menghan Xia[1,3] Wenbo Hu[1] Xueting Liu[2] Tien-Tsin Wong[1,3*]
[1] The Chinese University of Hong Kong [2] Caritas Institute of Higher Education
[3] Guangdong-Hong Kong-Macao Joint Laboratory of Human-Machine Intelligence-Synergy Systems, SIAT, Chinese Academy of Science
{mhxia, wbhu, ttwong}@cse.cuhk.edu.hk tliu@cihe.edu.hk

1. Network Architecture

Figure 1 illustrates the detailed network architecture, which is employed in both our dithering network and restoration network.

Pretrained Network for Guidance Loss. As mentioned in the training details, we propose to initially warm up the halftoning network with a guidance loss that is formulated through a pretrained inverse halftoning network. In our experiment, we employ a network with the same architecture as shown in Figure 1 and train it end-to-end. The training dataset and training details exactly follow the settings of [**?**] but the input halftone is the one obtained by applying Ostromoukhov method [**?**] on each grayscale sample. Once the training is converged, we take this pre-trained network, denoted as $\mathbf{F}$, to compute the guidance loss as defined in Eq.12 of the paper.

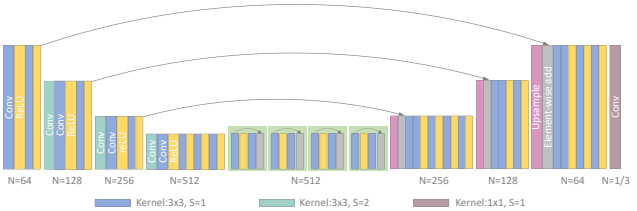


Figure 1. Network architecture. "N" denotes the channel number of feature maps in each layer, where for the last layer, $N = 1$ in the dithering network and $N = 3$ in the restoration network. "S" denotes the striding size of convolution layers. The upsample layer means $\times 2$ upscaling the feature maps by nears neighbor strategy.

2. Visual Comparison on Halftone

We compare our invertible halftone with Ostromoukhov method [**?**] and structure-aware halftoning [**?**] on several typical examples, as illustrated in Figure 2, Figure 3, Figure 4, and Figure 5. For quantitative reference, the PSNR and SSIM values are annotated with each resultant image. It shows that our invertible halftone achieves comparable visual quality with the state-of-the-art halftoning methods.

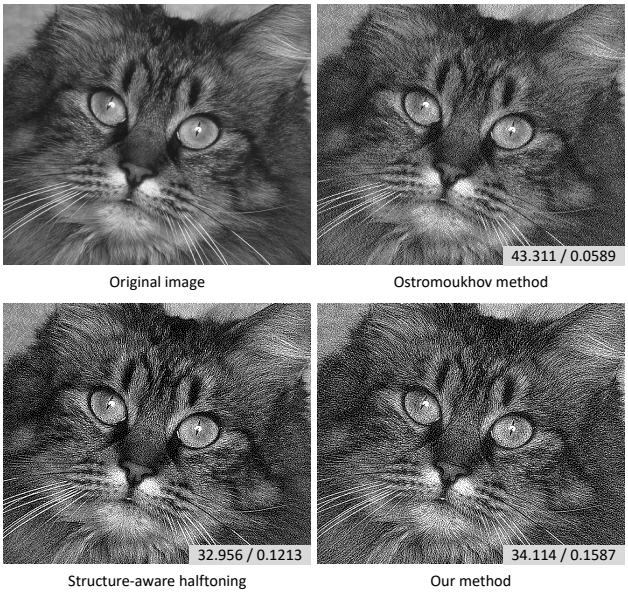


Figure 2. "Cat" example. All images are of resolution 660×560

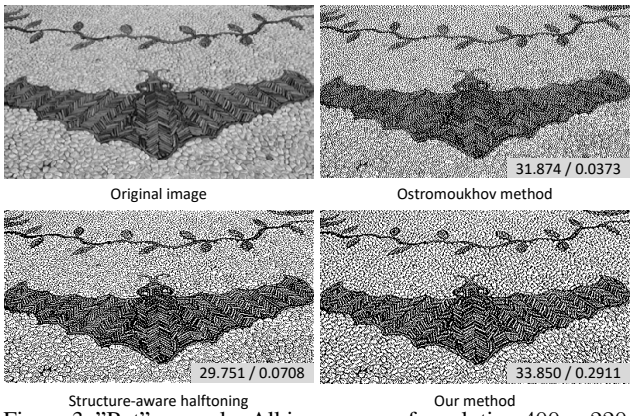


Figure 3. "Bat" example. All images are of resolution 400×220.

*Corresponding author.

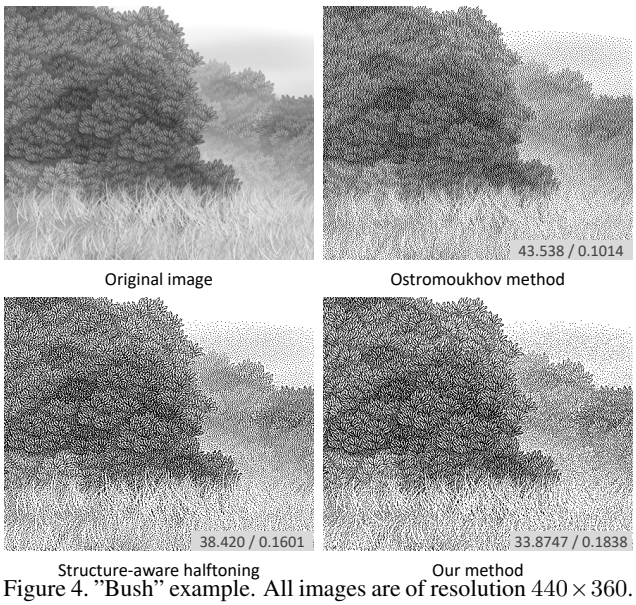


Figure 4. "Bush" example. All images are of resolution 440×360.

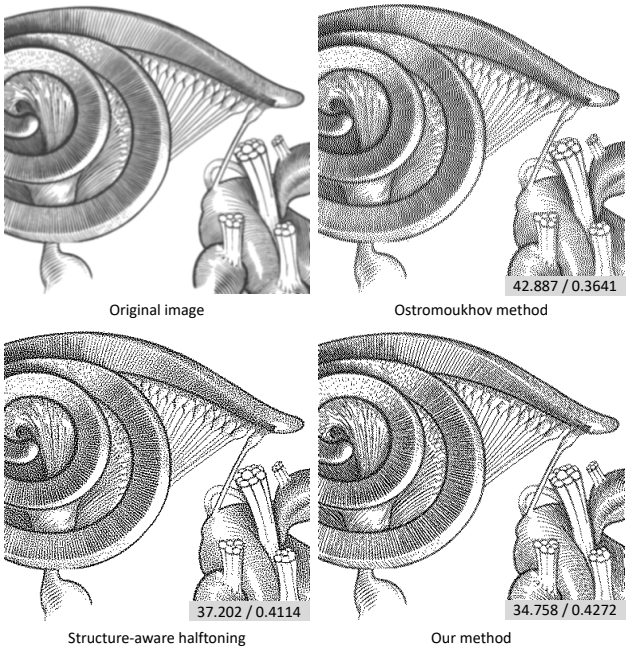


Figure 5. "Bush" example. All images are of resolution 380×360.

3. Visual Comparison with Inverse Halftoning

To show the significant superiority over inverse halftoning algorithms, we perform comparative results on several challenging examples, as illustrated in Figure 6. Particularly, PRL-Net [**?**], as the state-of-the-art inverse halftoning method on error diffusion, is employed for comparison. The input to PRL-Net is error diffusion, while the input to our model (decoder) is the invertible halftone. Due to the encoded information, our method can restore the grayscale images with very high fidelity, even for those very fine details. On the contrary, the PRL-Net can only recover the global appearance but has no way to get back those fine details from the error diffusion.

4. Robustness to Real-World Scanning

To validate our restoration ability in real world, we conduct an experiment to restore the color image from the printed hardcopy. Specifically, we first printed the generated invertible halftone on a regular A4 paper with a black frame. Then we scan the printed paper using a regular office scanner (default scanning: 300 dpi.). The scanned image is then cropped and rescaled to the original resolution. Finally, we binarize the scanned image by simple thresholding (0.5 as threshold) and feed this binary image to our restoration network.

Figure 7 illustrates three examples of restoring from printed halftones. Despite of various potential noise contamination during the scanning process, the scanned bitonal image contains binary pixel error less than 1% when compared to the original invertible halftone, thanks to its bitonal nature. For quantitative evaluation, we apply impulse noise (black/white dot) to the halftones and restore the color images from these noisy halftones. Over our testing dataset, the statistic result of noise ratio $p\%$ vs. restoration accuracy as (p / PSNR): *0/28.13, 1/27.72, 5/25.81, 10/23.70, 15/21.99, 20/20.59*.

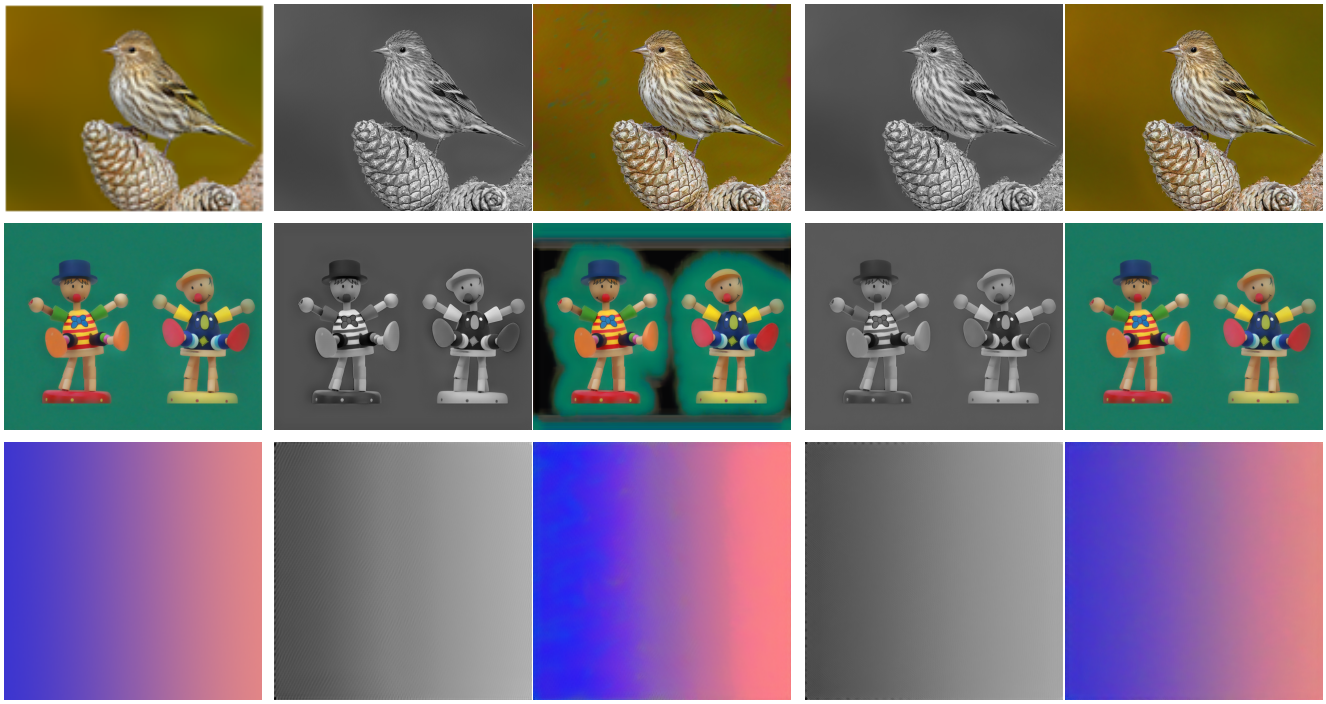
The decent robustness to outside disturbance implies that our method has a good potential in real-world printing applications. For instance, publishers may employ the invertible halftone in their product, and users would be able to enjoy the original color version of bitonal prints via downloading a decoder APP from the publisher.

5. Application of Noise Incentive Block

As we described in the paper, convolution degradation potentially affect the CNN performance on those tasks that expect non-flat output from flat input. In Figure 8 and Figure 9, we provide more qualitative results to inspect the benefits of the noise incentive block (NIB) to color image encoding [**?**] and image synthesis from semantic layout [**?**].

Figure 6. Restoration results of challenging cases. All images are of resolution 256×256.

Figure 7. Restoration from printed-and-scanned halftones. All images have the same resolution of 256×256.



Color input | Encoded grayscale and restored color by standard model | Encoded grayscale and restored color by NIB-equipped model

Figure 8. Applying NIB to color image encoding. In those flat regions, the color restoration failure is because the texture patterns are unable to be encoded on the grayscales due to the convolution degradation. Readers are recommended to zoom in the encoded grayscales to spot texture patterns.

Input | Ground truth | SPADE | SPADE+NIB

Figure 9. Applying NIB to semantic image synthesis. The red arrows point to regions with blurriness or artifacts.