Inverting Image Signal Processing Pipeline with Diffusion Models
Xinman Liu, Xuanchi Ren, Ziyi Wu
University of Toronto

Motivation
- Raw images are valuable not only for various image editing tasks but also for computer vision tasks such as image denoising. However, they are memory-intensive.
- Diffusion models are good at conditional generation tasks such as image super-resolution and image inpainting.
- Take advantage of the incredible generative power of diffusion models to solve raw image reconstruction task as a general image-to-image translation task.

New Technique
Contributions:
- The first attempt in RAW image reconstruction via diffusion models.
- Study the quality-speed trade-off with different sampling algorithms because diffusion models are slow and memory-consuming in the test-time generation process.

Methodology:
- (a) shows the iterative process RAW image reconstruction procedure, where a U-net performs denoising conditioned on the RGB image.
- (b) illustrates the architecture of U-net with skip-connections, which takes in the concatenation of a noisy RAW image \( X_t \) and an RGB image \( Y \), and predict the noise \( \epsilon_t \) which is then used to obtain \( X_{t-1} \).

Related Work
Raw Image Reconstruction
- [Source Image]
- [Reconstructed Image]

Diffusion Models
- Forward add Gaussian noise
- Backward train a model to denoise

References
- [1] The authors of the paper reference previous works in the field of image processing and diffusion models, providing a comprehensive overview of related research.

Experimental Results
- Table 1: Quantitative evaluation between our method and baselines.
  | Method | D3FIM | Canon EOS 5D
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>U-Net</td>
<td>40.98</td>
<td>34.82</td>
</tr>
<tr>
<td>DPM-Solver</td>
<td>40.98</td>
<td>34.82</td>
</tr>
<tr>
<td>Ours</td>
<td>40.98</td>
<td>34.82</td>
</tr>
</tbody>
</table>

- Table 2: Ablation study on model design on NIKON D7000 subset.
  | Method | D3FIM | Canon EOS 5D
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>U-Net</td>
<td>40.98</td>
<td>34.82</td>
</tr>
<tr>
<td>DPM-Solver</td>
<td>40.98</td>
<td>34.82</td>
</tr>
<tr>
<td>Ours</td>
<td>40.98</td>
<td>34.82</td>
</tr>
</tbody>
</table>

- Table 3: Ablation study on sampling strategy of diffusion models. All speed is measured on 6 NVIDIA VC10 GPUs with a batch size of 12 patches.
<table>
<thead>
<tr>
<th>Method</th>
<th>D3FIM</th>
<th>Canon EOS 5D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ours</td>
<td>40.98</td>
<td>34.82</td>
</tr>
<tr>
<td>DPM-Solver</td>
<td>40.98</td>
<td>34.82</td>
</tr>
</tbody>
</table>

Fig. 4: Visualization of our result on NIKON D7000 dataset. Our diffusion-based method is capable of synthesizing plausible RAW images.

Fig. 5: Qualitative comparison between different sampling strategies. Though DPM consumes much more time than other sampling strategies, it achieves the best RAW reconstruction quality.