Leveraging The Availability Of Larger FOV Images In Panoramas For Image Inpainting

Yu-Hsuan Lin*, Chih-Yu Wang*, Chi-Han Peng*
*Institute of Artificial Intelligence, National Yang Ming Chiao Tung University, Taiwan
roerca1222@gmail.com, wanglear6.ai12@nycu.edu.tw, pengchihan@nycu.edu.tw

***Abstract*—Image inpainting methods synthesize missing regions ("masks") within images. Modern approaches, powered by neural networks, generate content far beyond simple color extrapolation. A common observation is that performance degrades as the mask-to-image ratio increases. This motivates a strategy for panoramic inpainting: reducing the mask-to-image ratio by temporarily using an alternative view with a larger field-of-view (FOV) angle. In this paper, we analyze the impact of varying FOVs on the performance of two recent inpainting methods [1], [2] across different datasets. Our results reveal that moderate FOV enlargement improves performance, while excessive enlargement can instead degrade it.**

***Index Terms*—Image Inpainting, Panorama, Neural Network**

I. Introduction

Image inpainting methods synthesize missing regions ("masks") of an input image using generative neural networks [1]–[3]. A widely accepted notion is that inpainting performance is inversely proportional to the mask-to-image ratio — smaller masks generally yield better results [1]. While this principle is often used to categorize task difficulty, it has rarely been exploited to improve algorithm design, as both the mask and input image sizes are typically fixed.

We observe, however, that when working with *panoramas*, it is possible to temporarily suppress the mask-to-image ratio to enhance inpainting results. In a standard panorama inpainting pipeline — cropping a perspective view, performing inpainting, and reprojecting back — we can extract an alternative view with the same orientation but a larger field-of-view (FOV) angle. The masked region remains unchanged, but its relative size in the view is reduced.

In this paper, we implement such a pipeline and analyze the effect of different FOVs on recent inpainting methods [1], [2]. Our experiments show that slight FOV enlargement consistently improves performance, both quantitatively and qualitatively, while excessive enlargement can be detrimental. These findings suggest a simple yet effective strategy for enhancing panoramic image inpainting.

II. Related Work

A. Image Inpainting

Image inpainting, which allows users to remove unwanted objects from photos, has become a core feature in modern tools like Adobe Photoshop's "Generative Fill" [4] and Google's Magic Eraser. For a comprehensive overview, we refer readers to the recent survey [5] and curated resource [6].

We focus on two recent state-of-the-art methods. LAMA [1], proposed by Suvorov et al. (2021), uses Fourier convolutions within a ResNet framework to enhance large-mask inpainting by expanding the receptive field early in the network. MAT [2], introduced by Li et al. (2022), adopts a transformer-based architecture combined with a style manipulation module to better control the synthesized outputs.

Unlike previous works, our study proposes improving inpainting results by slightly enlarging the field of view (FOV), a strategy enabled when operating on panoramas, which provides the network access to more global structural information.

B. Panorama

Panoramas are images with very large or even omnidirectional FOVs, traditionally created by stitching multiple perspective images. Modern panoramic datasets, however, are typically captured directly using 360-degree cameras. Crucially, panoramas can be cropped into perspective views with arbitrary FOV angles, which we leverage to temporarily reduce the mask-to-image ratio and improve inpainting results.

We conduct experiments on two datasets: Matterport3D [7], a large-scale indoor RGB-D panorama dataset with 10,800 images, and CVRG-Pano [8], an outdoor street view panorama dataset.

III. Experiment Design

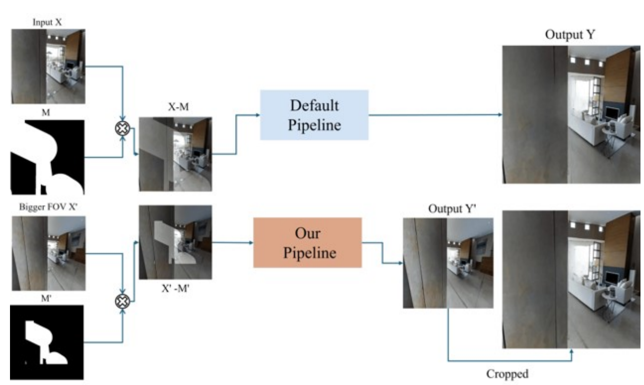

Fig. 1: Top(Blue): Standard image inpainting pipeline for panoramas. Bottom(Red): Our method reduces the mask-to-image ratio by creating a larger FOV image, only cropping the central part for the original view.

979-8-3315-9880-8/25/$31.00 ©2025 IEEE
DOI 10.1109/NICOInt67466.2025.00024

Our experimental pipeline is illustrated in Figure 1. Briefly, for each panorama, we randomly sample multiple perspective views. For every sampled view, we perform two types of inpainting: one on the original image and one on an alternative view with the same camera orientation but a larger FOV. The same mask is applied to both views; however, it appears proportionally smaller in the larger-FOV image. After inpainting, the results are projected back to the panorama.

For data preparation, we randomly sample 20 perspective views (90° FOV both horizontally and vertically) from each of the 2014 panoramas in the Matterport3D test split [7]. For each view, alternative versions with larger FOVs (up to 150°, in 10° increments) are generated. Masks are created following the protocol from LaMa [1], generating ”thin”, ”medium”, and ”thick” random masks, applied only on the original 90° FOV images. To complement the indoor Matterport3D dataset, we also utilize the outdoor CVRG-Pano dataset [8] consisting of 600 urban street panoramas.

Regarding resolution settings, all panoramas are at 2048×1024 resolution. Accordingly, 90° FOV views are rendered at 512×512 resolution. To ensure fair comparisons, perspective views with larger FOVs are rendered at proportionally higher resolutions (e.g., 120° FOV views at 886×886) such that the central region corresponding to 90° FOV maintains 512×512 resolution. This scaling is consistently applied from 90° to 150° FOV views.

IV. Experimental Results

We evaluated LAMA [1] and MAT [2], two state-of-the-art image inpainting methods, using common metrics: Learned Perceptual Image Patch Similarity (LPIPS) [9], Paired/Unpaired Inception Discriminative Score (PIDS/UIDS) [10], and Fréchet Inception Distance (FID) [11]. LPIPS measures perceptual similarity between images using a pretrained neural network, PIDS/UIDS assess inpainting results based on separability in feature space, and FID computes the Wasserstein distance between distributions of ground truth and synthesized images.

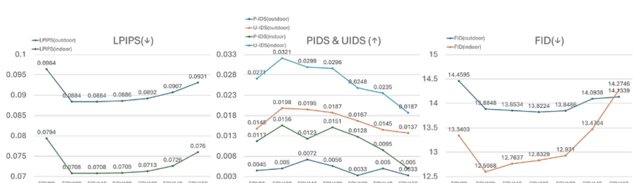

Fig. 2: Comparative results of LAMA on the Matterport3D indoor and CVRG-Pano outdoor datasets.

Quantitative Results. Figure 2 compares the results of LAMA on the Matterport3D indoor and CVRG-Pano outdoor datasets. In both cases, increasing the FOV by 10° generally improves inpainting performance, with similar trends observed in both indoor and outdoor settings. However, further enlarging the FOV beyond this point does not lead to additional improvements.

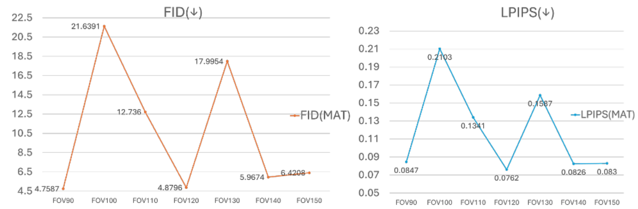

Fig. 3: Quantitative results of MAT on the Matterport3D dataset.

A. MAT Results

MAT [2] was also tested for image inpainting. Qualitatively, the trends observed were similar to LAMA. However, the quantitative results (Figure 3) appeared more inconsistent, likely due to MAT's variability in results even with the same masks and style manipulation parameters.

V. Conclusion

We found that slightly increasing the field of view (FOV), such as from 70° to 80°, improves inpainting performance across two datasets (Matterport3D and CVRG-Pano) and two SOTA methods (LaMa and MAT). However, excessive enlargement degrades quality due to resolution stretching.

Our main contributions are: (1) introducing FOV enlargement as a novel strategy to enhance panorama inpainting, (2) implementing and evaluating a prototype pipeline, and (3) identifying that only modest FOV increases are beneficial. Future work will explore applying this strategy to other image processing tasks.

References

[1] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, N. Kong, H. Goka, K. Park, and V. Lempitsky, “Resolution-robust large mask inpainting with fourier convolutions,” *CoRR*, vol. abs/2109.07161, 2021. [Online]. Available: https://arxiv.org/abs/2109.07161

[2] W. Li, Z. Lin, K. Zhou, L. Qi, Y. Wang, and J. Jia, “Mat: Mask-aware transformer for large hole image inpainting,” 2022.

[3] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro, “Image inpainting for irregular holes using partial convolutions,” 2018.

[4] Adobe, “Next-level generative fill. now in photoshop.” [Online]. Available: https://www.adobe.com/products/photoshop/generative-fill.html

[5] H. Xiang, Q. Zou, M. A. Nawaz, X. Huang, F. Zhang, and H. Yu, “Deep learning for image inpainting: A survey,” *Pattern Recognition*, vol. 134, p. 109046, 2023.

[6] Y. ZENG, “Awesome image inpainting.” [Online]. Available: https://github.com/zengyh1900/Awesome-Image-Inpainting

[7] A. Chang, A. Dai, T. Funkhouser, M. Halber, M. Nießner, M. Savva, S. Song, A. Zeng, and Y. Zhang, “Matterport3d: Learning from rgb-d data in indoor environments,” 2017.

[8] S. Orhan and Y. Bastanlar, “Semantic segmentation of outdoor panoramic images,” *Signal, Image and Video Processing*, 2021.

[9] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, “The unreasonable effectiveness of deep features as a perceptual metric,” 2018.

[10] S. Zhao, J. Cui, Y. Sheng, Y. Dong, X. Liang, E. I. Chang, and Y. Xu, “Large scale image completion via co-modulated generative adversarial networks,” 2021.

[11] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, “Gans trained by a two time-scale update rule converge to a local nash equilibrium,” 2018.