Application of Frequency Rectification on Images

Supervisor: Chenliang Zhou (chenliang.zhou@cst.cam.ac.uk)

Background

FrePolad (**f**requency-**re**ctified **po**int **la**tent **d**iffusion) [1] is proposed as a **point cloud** generation pipeline integrating a variational autoencoder (VAE) with a denoising diffusion probabilistic model (DDPM) for the latent distribution. FrePolad simultaneously achieves high quality, diversity, and flexibility in point cloud cardinality for generation tasks while maintaining high computational efficiency. The improvement in generation quality and diversity is achieved through (1) a novel **frequency rectification** via **spherical harmonics** designed to retain high-frequency content while learning the point cloud distribution; and (2) a latent DDPM to learn the regularized yet complex latent distribution. While FrePolad has demonstrated great success in **3**D point cloud generation, it raises an intriguing question: can these **frequency rectification techniques** be successfully applied to **2D images** for tasks such as image reconstruction, generation, or manipulation?

Goal

The goal of this project is to investigate how **frequency rectification**, a concept successfully applied in FrePolad for point clouds, can enhance the capabilities of **2D image reconstruction, generation, or manipulation**. The successful implementation of these ideas could lead to more detailed, diverse, and visually compelling image generation models.

Tasks

1. **Background Review**:
    - **FrePolad [1]**: Gain a deep understanding of the key concepts behind FrePolad, particularly how frequency rectification and spherical harmonics are applied to point clouds. There is no need to run the FrePolad code, but a thorough conceptual understanding is essential.
    - **Fourier Transforms & Spherical Harmonics**: Familiarize yourself with **Fourier transforms** and **spherical harmonics** (spherical harmonics, used in FrePolad, are essentially Fourier transforms applied on the spherical domain), understanding their roles in data processing, especially how these techniques enable the extraction of high-frequency information from complex datasets.
    - **Literature Review**: Conduct a comprehensive review of key works in 2D image reconstruction, generation, and manipulation, with particular focus on those utilizing frequency information (e.g., [2-3]). Analyze how frequency-based methods are currently applied and think about how we can innovate to extract and utilize frequency data in new ways to improve these tasks.
2. **Identify and understand a baseline model**: Select a suitable 2D image

generation, reconstruction, or manipulation model to serve as the baseline for applying frequency rectification. Ensure that the chosen model fits well within the project's scope. Study the structure and methodology of the baseline model, ensuring that you fully understand its core components, data representation, and how it processes image data. Be prepared to adapt and modify the model for the integration of frequency rectification. Run the model's codebase and ensure that it operates correctly. This step is crucial for understanding its behavior and identifying areas where frequency rectification can be integrated.

3. **Dataset and evaluation**: You will likely reuse the datasets and evaluation metrics from the original baseline model. However, depending on the chosen model, modifications may be needed to accommodate the application of frequency rectification techniques.
4. **(Main Task) Design and develop the pipeline**: Develop a pipeline that integrates frequency rectification into the baseline model. Keep in mind that the approach used in FrePolad for point clouds—leveraging spherical harmonics—may not directly apply to 2D images due to the difference in data structures. As a result, you may need to devise new methods for extracting and utilizing frequency information from images. You might want to approach the task in the following two ways:
    - **How to extract?:** Consider innovative ways to extract frequency information from 2D images. For example, you could explore how the grid-like structure of 2D images lends itself to efficient frequency extraction via Fourier transforms or other frequency-based methods.
    - **How to Use?:** Decide how the extracted frequency data will be applied in the generation or reconstruction pipeline. Will it be used for regularization, enhancing detail generation, or providing spatial coherence in the output? Ensure that the frequency information is smoothly integrated into the pipeline without causing significant performance degradation.

Possible Extensions

- **Multi-scale frequency representation**: Explore the incorporation of multi-scale frequency features. By enabling the pipeline to analyze frequency information at different levels of detail, you can improve the generation or reconstruction of images with varying spatial resolutions. This approach could be particularly useful in generating complex scenes or hierarchical image structures.
- **Hybrid domain transformations**: Beyond traditional Fourier transforms, consider hybrid approaches such as combining **wavelet transforms** or **Discrete Cosine Transforms (DCT)** with Fourier-based methods to capture both local and global frequency information. This may provide richer details in texture or shape during reconstruction or generation tasks.
- **Applications to editing and manipulation**: Investigate how frequency rectification can be used for more targeted manipulation or editing tasks. For

example, it may improve the quality of super-resolution, denoising, or inpainting tasks by focusing on reconstructing lost high-frequency details.

References

[1] https://chenliang-zhou.github.io/FrePolad/
[2] https://arxiv.org/abs/2305.02541
[3] https://openaccess.thecvf.com/content/ICCV2021/papers/Jiang_Focal_Frequency_Loss_for_Image_Reconstruction_and_Synthesis_ICCV_2021_paper.pdf