Novel View Synthesis for Medical X-Ray Images using Progressive Conditional Diffusion Models

Xie Chun
University of Tsukuba

1. Project Purpose

Radiography, a prevalent technique for visualizing internal human anatomy, employs high-energy X-rays to penetrate the body. The residual radiation energy is captured on a flat detector. Since various organs absorb X-rays differently, the measured energy is translated into a two-dimensional image, known as a radiograph or X-ray image, which discloses the body's internal configuration and offers crucial diagnostic data. X-ray images are both rapid and cost-effective; however, the use of high-energy radiation may have detrimental health implications.

Typically, in procedures like chest radiography, only a single frontal view is obtained per session. Although physicians can instinctively interpret the spatial arrangement of organs on a 2D radiograph in a three-dimensional context, this intuition is inherently subjective and can vary in precision.

The development of an X-ray image view synthesis algorithm could be beneficial, providing additional insight into a patient's internal anatomy. Furthermore, this could facilitate other applications, including sparse-view CT reconstruction from X-ray images, bridging the gap between CT and X-ray and enhancing the utility of radiographic imaging.

Diffusion models have shown extremely high performance in image generation tasks. It is also reported that with proper fine-tuning, a pre-trained model can generate realistic medical X-ray images from given text prompts. We hereby consider that with a carefully designed conditioning approach, it is possible to generate novel view X-ray images from a source image and a target viewport.

2. Results

This work introduces a newly designed diffusion network pipeline capable of generating consistent novel-view medical X-ray images based on a source image and relative camera pose. The model is generalized and robust, having been trained on a large-scale dataset of synthetic X-ray images. A comprehensive evaluation of the proposed model is conducted, exploring the compatibility of various image feature extraction approaches, including semantic and visual features, with the pipeline. Ultimately, this research establishes a strong baseline for future studies and advancements in the field.

3. Roles of the MCRP and its significance

MCRP provided the necessary computing resources to prepare our data, and to train and evaluate our model.

4. Future plan

We will improve the consistency between views and reconstruct 3D CT models using the generated multi-view X-Ray images.

5. Publications and conference presentations

(1) Journal papers

(2) Presentations

MICCAI 2025, Under review.

(3) Others

Supercomputer	Use	Allocated resources*		
		Initial resources	Transferred resources**	Additional resources
Cygnus	Yes	36000		
Pegasus	Yes	16000		
Wisteria/BDEC-01	No			
	*in units of node-hour product ** If the budget transfer was performed, fill in here, such as "+2000" and "−1000".			