

Modern Techniques for Medical Data Imaging on Consumer-Grade Hardware

Matej Lorinc

compiled: February 5, 2026

ii

f

Preface

Draft

Taking a step back, modern Western medicine can be thought of as the center of mass for all interactions and exchanges of data between researchers and clinicians. As progress is made, this center shifts and so does the paradigm. Clinicians gather data from patients with diseases, subjecting them to treatments synthesized by researchers. Researchers, motivated by the chorus of clinicians signaling the need for cures, or by the insufficient effectiveness of existing treatments, shift their focus and resources toward creating new, improving, or maintaining these protocols.

Looking through a magnifying mirror into this web of data collection and analysis and fixing our focus to a single instance of doctor-patient interaction we can see a certain trend. It seems that the contact between the provider of treatment and the recipient has become a metaphor for measurement. We extract blood or tissue samples, we measure temperature, and analyze the chemical composition of various bodily fluids. As each new iteration of treatment aspires to take action as early as possible, the ability to recognize signals that predict the imminent coming of a disease has become of greatest importance to modeling and executing treatment effectively.

However, humans have long hit their limits in their ability to probe large sets of data while simultaneously performing calculations on them. We have turned to machines that collect, analyze, and output data only some of which we may fully comprehend. It does not appear that this disconnection between the actual corpus of the patient and the hands and eyes of the doctor will slow down. On the contrary, even for medical fields like orthopedics, the physical touch of the doctor and the subsequent communication about the patient's pains have been, to a large degree, shifted to non-invasive scans and visualizations. One effectively no longer needs to move a knee joint to determine which knee hurts. The need to actually see a tumor to diagnose a cancer

Draft

iv

slowly fades in the face of blood analysis.

This thesis builds upon a crucial subtlety within this trend. The acceleration of data processing does not translate to an obsolescence of human perception. While the tactile link may be fading, the visual link is staying and through technology is becoming stronger. Specific domains, namely surgery, such as craniofacial surgery, orthopedics, and a behemoth that is medical education, refuse to distance themselves from the act of looking.

In these fields, "looking with eyes" is not a metaphor, it is a mechanical necessity. A surgeon cannot cut based on a spreadsheet of probabilities. They must see and internalize the spatial relationship of the animal anatomy. An orthopedic specialist must see the fracture within the context of the muscle and tissue. A medical student cannot learn the complexities of the human body through raw data points alone. They require a visual model to bridge the gap between theory and reality.

And as such, as we distance ourselves from the physical patient, the *medical scan* becomes the proxy for the patient's body. The burden of reality presentation shifts from the flesh to the pixel. If the doctor is to rely on the screen rather than the touch, the image on that screen must not just be a data visualization. It should aim to be a faithful optical reconstruction of reality.

Introduction

Motivation and Overview

The visualization of animal, particularly human, anatomy through non-invasive means is one of the most significant achievements in modern medicine.¹ **Medical scans** allow clinicians to diagnose pathology and enable them to gather knowledge about the internal state of a human body without making a single incision. However, as hardware improved, it begun producing increasingly more complex, high resolution datasets. The challenge has shifted from the question of how we capture data to how do we display it effectively.

The oldest technique for imaging the internal human body is projectional radiography, commonly known as X-ray. To image a common pathology, such as a broken bone, the subject is placed between a source of electromagnetic radiation and a detector. As radiation passes through the body, it is attenuated by various **tissues**, decreasing on intensity. This creates a non-uniform intensity distribution on the detecting plane, which allows clinicians to distinguish internal structures and organs.

A different imaging approach, sonography, uses high-frequency sound waves. A transducer emits high-frequency sound waves through contact with the skin into the body, where they are either scattered or reflected. The reflected waves are detected and the scanner determines how long did the echo take to return and how strong the echo was. These two variables determine the spatial location and brightness of the corresponding pixels in the resulting visual.

While these techniques are foundational, ultimately, they result in 2D representations of 3D structures. This thesis focuses on scans that produce spatially complete² **volumetric data**. The availability of 3D spatial information allows users to project data

¹ Or rather we should say one of the most significant accidental gifts of physics to modern medicine. X-rays were discovered by Röntgen in 1895 and found their immediate application in medicine.

² Such that it covers the whole volume uniformly typically as a 3d grid of voxels.

onto a 2D monitor at any orientation or scale. ALternatively, users can view the data stereographically through virtual reality system. This unlocks the full potential of human visual perception and, together with real-time interactivity, creates an ideal environment for both **education** and **clinical analysis**.

This research focuses on methods that map the body onto a Euclidean space of data points. Each point contains sufficient information to determine the **optical properties** of that physical location.

In Computed Tomography (CT), the output data is created by the measurement of X-ray attenuation from multiple angles. The mathematical reconstruction of this data relies on the Radon transform ³, an integral transform which reconstructs a function from its line integrals. This results in a scalar field of attenuation coefficients that must then be interpreted by the rendering algorithm.

Magnetic Resonance Imaging (MRI) uses the principles principles of nuclear magnetic resonance. Here, the output data is derived from the relaxation times of hydrogen nuclei in a magnetic field. The reconstruction of these spatial frequencies into an image is performed via the Fourier transform.

Both techniques produce volumetric grids of datapoints. This thesis focuses on translating these datasets into a physically based optical model that simulates light behavior. This requires establishing a correlation between each data point and the tissue it represents. Modern approaches may utilize transformer-based neural networks for this task. Tissues can be classified using the global context of the datapoint position, geometric flow, or state-of-the-art pretrained segmentation models⁴.

Once the biological tissue represented by a datapoint is identified, we can proceed to model its specific optical properties ⁵ and subsequently simulate light interactions within the dataset, rendering the results on screen.

³ An integral transform $f \mapsto Rf$ where f is some function defined on a plane and Rf is a function defined on a space of line in the plane whose values are defined as the line integral of that function over that line. The reconstruction of the original object is then achieved by taking the inverse of this transform.

⁴ Liu et al., "MedSAM3: Delving into Segment Anything with Medical Concepts"

⁵ Jacques, "Optical properties of biological tissues: a review"

Goals

The primary objective of this thesis is to develop a **rendering framework**⁶ that transforms the raw volumetric data into a visual representation that is both intuitive for non-professionals⁷ and functionally valuable for professionals.

We move beyond traditional method of **transfer functions** towards photorealistic **path tracing** and **splatting**. The thesis aims to explore a theoretical imaging solution that would provide clinicians and educators with a tool that would improve the comprehension of spatial relations and increase the diagnostic quality and effectiveness. The goal here is to demonstrate that (optical) **anatomical realism** does not have to be sacrificed for **performance**. Building on the potential of the architecture, we aim to enable a degree of **interactivity** previously achievable only by lower-fidelity models. This includes dynamic lightning where users can manipulate light sources in real-time allowing for better understanding of depth and surface texture, linear and non-linear transformations that is, the ability to move, rotate and scale specific anatomical structures and temporal interactivity, where scans containing additional temporal dimension could be used to animate organs such as the heart or lungs.

For surgeons and medical students, the benefit of photorealistic visualization lies in the depth of perception and the ability of broad material differentiation. Traditional rendering of medical scans often results in 'flat' or 'plastic' appearances that lack both the crude and subtle anatomical details⁸. By simulating the complex behavior of light, such as subsurface scattering and realistic shadowing, our goal is to produce visualizations that mimic the look and optical feel of a physical human body dissection. This level of visual fidelity allows for better understanding of the relative distance and orientation of overlapping structures. This backend framework, together with an UI that would envelop the internals of the rendering system, could provide students with a 'virtual cadaver', aiding in the transition from reading textbook theory to seeing the operating theater as close to in person as the technology allows.

A core goal of this framework is to allow the user to manipulate the digital volume. Rather than viewing a static image, users will be able to interact with the anatomy through:

⁶ Or at least a prototype of rendering framework. Features such as full user interface are outside of the core research scope of this work.

⁷ An experienced radiologist might see a fracture where unskilled eye sees only a gray overlapping shapes.

⁸ Since these methods discard or do not explore the secondary complex interactions of light. These interactions comprise of the majority of 'realism' that is achieved by path tracing methods. The sum total of these interactions create a crude texture that is difficult to achieve by targeted texture design.

- Segmentation by using neural network-assisted classification to isolate specific organs or tissue layers.
- Volumetric editing by enabling the 'peeling back' of tissue layers, creating slices through the volume or scaling specific structures to reveal underlying pathology.
- Variable transparency by adjusting the opacity of structures (like the skin or musculature) to make them semi-transparent, allowing for a 'see-through' effect that visualizes the internal relations within the body.

The nuances of these goals, however, lie in the tradeoff between what is achievable if high-end computing clusters are available and what can be done with a consumer-grade hardware found⁹ in clinical institutions and teaching faculties. Achieving path-traced realism in real-time requires clever data structures and pruning algorithms to minimize redundant calculations and the waste of available compute power by rendering elements outside of line of sight. Optimizing how the splatting engine traverses the volumetric grid to ensure that only visible and relevant data points contribute to the final image. Implementing efficient ways to handle high-resolution medical datasets without exceeding the VRAM limits of standard GPUs. Balancing the power-hungry but desired stochastic character of path tracing with the rasterization speed of Gaussian splatting to maintain high framerate experience without inducing sickness in the user¹⁰.

Finally, the development of this framework will not occur in a communication vacuum. A significant goal of this research is to ensure the output meets the standards of the medical community. We will be gathering qualitative critiques from clinical professionals to assess the diagnostic utility and accuracy of the rendered images and educational faculties to evaluate the efficacy of the tool as a pedagogical resource for medical training.

This iterative process of consultation and improvement aims to ensure that our rendering optimizations translate into tangible benefits for the clinical field.

⁹ This metric will be determined by an indirect survey.

¹⁰ A common problem in VR systems with low framerates.

Methodology and Related Work

As the field of medical visualization has moved from simple cross-sectional slices to complex, interactive 3D environments, achieving the current state of the art requires a clever combination of classical volume rendering, high-end cinematic techniques ¹¹, and the results of recent explosion of interest in differentiable rendering.

Historically, the visualization of volumetric grids has relied on two primary methods: **isosurface reconstruction** and **direct volume rendering** (DVR).

DVR uses ray marching method to simulate a ray attenuation from arbitrary viewing angle. The rays are extended from the viewing plane according to the desired type of projection (orthographic, perspective) and for each datapoint hit, the resulting dimming of the ray is calculated. Using transfer function to map scalar values to color and opacity, DVR allows for a semi-transparent view of internal structures. This method can be combined with precalculated gradient to simulate basic shading on the normals. While versatile, standard DVR often lacks realistic lighting, resulting in images that struggle to convey depth and material properties properly, acting more like a x-ray with arbitrary user-specific rotation and scale.

Iso-surface reconstruction, most notably the **marching cubes** algorithm implementation, extracts a polygonal mesh from a volume at a specific intensity threshold. While efficient for hardware-accelerated rendering, it discards the internal 'fuzziness' of the data and tends to create choppy, smoothed out plastic-like surfaces.

The current industry gold standard for photorealistic medical visualization is **Cinematic Rendering** first introduced to mainstream by Siemens Healthineers ¹². In contrast to simplistic lighting models of DVR this model employs **monte carlo path tracing** to simulate the complex physics of light transport, including global illumination, soft shadows, and ambient occlusion. While Cinematic Rendering results in state-of-the-art visual quality, it is computationally expensive. It continues to struggle to maintain high framerate interactivity on consumer-grade hardware, often requiring specialized workstations or offline rendering.

In the race for of real-time performance, various **novel view synthesis**¹³ (NVS) approaches building on the advances in dif-

¹¹ Such as those used in movie making and computer game development.

¹² Niedermayr et al., "Application of 3D Gaussian splatting for cinematic anatomy on consumer class devices"

¹³ Essentially outsourcing the problem of calculating high volumes of linear algebra appearing in ray optics mechanics to linear algebra of machine learning.

¹⁴ A rendering pipeline that is differentiable with respect to the parameters of the rendered scene. This allows for a gradient flow from a loss calculated using the ground truth and some scene with flexible enough components back to the scene, adjusting the components to represent the ground truth faithfully.

¹⁵ Mildenhall et al., "NERF: Representing scenes as neural radiance fields for view synthesis"

¹⁶ Kerbl et al., "3D Gaussian Splatting for Real-Time Radiance Field Rendering"

¹⁷ Gao et al., Jiang et al., "Relightable 3D Gaussians: Realistic Point Cloud Relighting with BRDF Decomposition and Ray Tracing", "Gaussian-Shader: 3D Gaussian Splatting with Shading Functions for Reflective Surfaces"

¹⁸ Bidirectional reflectance distribution function describes the way light behaves once it hits the surface of given texture. It is a core part of physically based rendering mathematics.

¹⁹ Liang et al., "GS-IR: 3D Gaussian splatting for inverse rendering"

²⁰ Yao et al., "Reflective gaussian splatting"

²¹ Bai et al., "GaRe: Relightable 3D Gaussian Splatting for Outdoor Scenes from Unconstrained Photo Collections"

ferentiable rendering¹⁴ have been explored. Neural radiance fields¹⁵ (NeRFs) introduced a way to represent scenes within neural networks. However, NeRFs often suffer from long training times and slow inference speeds, making them not suitable for real-life performance and the interactive needs of a clinical environment.

The introduction of **3D Gaussian Splatting** (3DGS)¹⁶ has marked a significant shift in the field, offering the high fidelity of neural methods with the speed of traditional point-based rasterization. Current research in various splatting methods is rapidly expanding the capabilities of this approach. Recent advances have introduced relightable Gaussians¹⁷, allowing for dynamic lighting changes by incorporating BRDF-like¹⁸ properties into each splat. Advances in shading functions¹⁹ and reflective properties²⁰ are moving splatting toward true **physically based rendering** (PBR). New methods are addressing the challenge of shadows and complex occlusions within splatted scenes²¹, which is vital for understanding the spatial depth of anatomical structures. The question of optimization and visual quality²² of Gaussian splatting is being investigated by a large number of teams thanks to the Meta's push for the deployment of this technology in virtual avatar communication systems²³. Leveraging these advancements in Gaussian Splatting and combining them with path-tracing principles, this thesis aims to deliver a framework that matches the visual quality of Siemens' Cinematic Rendering while maintaining the real-time performance necessary for modern medical practice.

3D Gaussian Splatting

Virtual Reality

The motivating drive behind the proposal of this thesis is supported by the rapid improvement and increasing accessibility of computing hardware capable of such feats. Over the last several years, the sector has moved beyond large, costly solutions requiring flagship hardware toward self-contained, ergonomic systems. Virtual Reality (VR) dominates the potential for immersive experience of medical training²⁴. It provides a fully controlled environment for risk-free procedural rehearsals with a active proprietary software provider sector. There has also been a significant surge caused by a push for in the development of augmented reality (AR) and mixed reality (MR) devices that allow user to mix the light coming from the real environment with rendered objects. [TODO]

Modern, eyeglass-like wearables, such as smart glasses, are increasingly viewed as the primary global development target for augmented reality. These systems enable surgeons to overlay volumetric data directly onto the surgical field or allowing students to collaborate around a shared holographic specimen in a physical classroom, or potentially, from home. As the industry moves toward miniaturized, lightweight form factors with higher pixel densities and wider fields of view, the demand for rendering frameworks that can deliver photorealistic, low-latency results increases. This thesis addresses this need by optimizing the path-tracing and splatting pipeline specifically for the performance profiles of these emerging miniaturized wearable platforms²⁵.

While 2D monitors are sufficient for standard diagnostics, VR provides a 1:1 spatial mapping that aligns with how humans naturally perceive the physical world. This is particularly important in two key areas.

Traditional medical training often relies on textbooks, 2D atlases, and limited access to cadaver specimens. VR overcomes these problems by providing a 'digital cadaver' that is infinitely reusable and free from biological decay. By allowing students to 'walk through' a heart or navigate the vascular system at an exaggerated scale, VR enables for a deeper understanding of complex spatial relationships that are often lost in cross-sectional imaging. Students can perform repeated manipulations such as isolating a specific nerve or simulating a surgical pathway, allowing them to learn from mistakes without patient risk.

²² Yu et al., "Mip-Splatting: Alias-free 3D Gaussian Splatting"

²³ Saito et al., "Relightable Gaussian Codec Avatars"

²⁴ Mergen, Graf, and Meyerheim, "Reviewing the current state of virtual reality integration in medical education - a scoping review"

²⁵ We will research the capabilities of such systems already in-market and in development.

For surgeons, VR acts as a bridge between the digital scan and the operating theater. Surgeons can use the proposed rendering framework to interact with patient-specific anatomy, testing different surgical approaches by manipulating and 'hiding' tissue layers to find the most efficient route. Our framework could offer a multi-user environment where a team (e.g., a radiologist and a surgeon) can simultaneously view the same model in a shared virtual space, regardless of their physical location.

Progress so Far

The chosen topic covers both medical and computer science. The initial phase involved a study of the medical literature to understand the broad language of anatomy and diagnostics based on medical scans. A review of current standards in medical imaging (e.g., DICOM, CT/MRI segmentation) was conducted where the limitations of traditional 2D slice viewing were discovered. Parallel to the medical study, a review of state-of-the-art computer science literature was conducted. Investigating the mathematical framework of 3D Gaussians scene reconstruction and studying how gradient-based optimization can be used to refine medical models. We have studied the shading programming language HLSL and rendering process of DVR and isosurface reconstruction on concrete datasets using the Unity²⁶ engine.

²⁶ <https://unity.com/>

We continue to review the literature to build an intuition for the possibilities and limitations of the gaussian splatting technology and the direction in which the general research is headed.

The next steps will consist of translating the theory behind gaussian splatting into practice by studying the reference implementation of the papers. This will give us a outlook on the possibilities and problems to try to find solutions to in the next phase.

In parallel, we will research the hardware capabilities of current and next-generation virtual reality systems together with benchmarking the current generation of graphics hardware found on personal computers on rendering path tracing and splatting methods.

Bibliography

- [1] Anglin Liu et al. "MedSAM3: Delving into Segment Anything with Medical Concepts". In: *arXiv (Cornell University)* (Nov. 2025). DOI: 10.48550/arxiv.2511.19046. URL: <https://doi.org/10.48550/arxiv.2511.19046>.
- [2] Steven L Jacques. "Optical properties of biological tissues: a review". In: *Physics in Medicine & Biology* 58.11 (May 2013), R37. DOI: 10.1088/0031-9155/58/11/R37. URL: <https://doi.org/10.1088/0031-9155/58/11/R37>.
- [3] Simon Niedermayr et al. "Application of 3D Gaussian splatting for cinematic anatomy on consumer class devices". In: *arXiv (Cornell University)* (Apr. 2024). DOI: 10.48550/arxiv.2404.11285. URL: <http://arxiv.org/abs/2404.11285>.
- [4] Ben Mildenhall et al. "NERF: Representing scenes as neural radiance fields for view synthesis". In: *arXiv (Cornell University)* (Mar. 2020). DOI: 10.48550/arxiv.2003.08934. URL: <http://arxiv.org/abs/2003.08934>.
- [5] Bernhard Kerbl et al. "3D Gaussian Splatting for Real-Time Radiance Field Rendering". In: *arXiv (Cornell University)* (Aug. 2023). DOI: 10.48550/arxiv.2308.04079. URL: <http://arxiv.org/abs/2308.04079>.
- [6] Jian Gao et al. "Relightable 3D Gaussians: Realistic Point Cloud Relighting with BRDF Decomposition and Ray Tracing". In: *arXiv (Cornell University)* (Nov. 2023). DOI: 10.48550/arxiv.2311.16043. URL: <http://arxiv.org/abs/2311.16043>.
- [7] Yingwenqi Jiang et al. "GaussianShader: 3D Gaussian Splatting with Shading Functions for Reflective Surfaces". In: *arXiv (Cornell University)* (Nov. 2023). DOI: 10.48550/

- arxiv.2311.17977. URL: <http://arxiv.org/abs/2311.17977>.
- [8] Zhihao Liang et al. "GS-IR: 3D Gaussian splatting for inverse rendering". In: *arXiv (Cornell University)* (Nov. 2023). DOI: 10.48550/arxiv.2311.16473. URL: <http://arxiv.org/abs/2311.16473>.
- [9] Yuxuan Yao et al. "Reflective gaussian splatting". In: *arXiv (Cornell University)* (Dec. 2024). DOI: 10.48550/arxiv.2412.19282. URL: <http://arxiv.org/abs/2412.19282>.
- [10] Haiyang Bai et al. "GaRe: Relightable 3D Gaussian Splatting for Outdoor Scenes from Unconstrained Photo Collections". In: *arXiv (Cornell University)* (July 2025). DOI: 10.48550/arxiv.2507.20512. URL: <https://doi.org/10.48550/arxiv.2507.20512>.
- [11] Zehao Yu et al. "Mip-Splatting: Alias-free 3D Gaussian Splatting". In: *Conference on Computer Vision and Pattern Recognition (CVPR)* (2024).
- [12] Shunsuke Saito et al. "Relightable Gaussian Codec Avatars". In: *arXiv (Cornell University)* (Dec. 2023). DOI: 10.48550/arxiv.2312.03704. URL: <http://arxiv.org/abs/2312.03704>.
- [13] Marvin Mergen, Norbert Graf, and Marcel Meyerheim. "Reviewing the current state of virtual reality integration in medical education - a scoping review". In: *BMC Medical Education* 24.1 (July 2024), p. 788. DOI: 10.1186/s12909-024-05777-5. URL: <https://doi.org/10.1186/s12909-024-05777-5>.