UNIVERSITÀ DEGLI STUDI DI UDINE

POLYTECHNIC DEPARTMENT OF ENGINEERING AND ARCHITECTURE

PHD IN INDUSTRIAL AND INFORMATION ENGINEERING

Abdul Salam Rasmi ASRAF ALI

Maxillofacial CBCT Reconstruction in the Face of Challenging Patient Motions

Supervision

Prof. Andrea FUSIELLO

February, 2024

Abstract

Maxillofacial Cone Beam Computed Tomography (CBCT) plays a key role in dentistry for diagnostic purposes and treatment planning. However, the challenge of artifacts induced by patient movement during scanning persists. While manufacturers have incorporated mechanical fixations in cone beam CT systems to prevent patient movement, these fixations are not entirely rigid and may still allow some degree of motion. To address this issue, researchers have explored motion compensation algorithms, yet these solutions are often time-consuming and only partially address the problem.

This thesis presents a novel motion compensation approach based on the presumption that a motion-free subset exists within the scanned projections. The approach integrates a motion detection framework to identify and extract this motion-free subset. Subsequently, a short-scan reconstruction is executed using this subset, serving as a reference for the compensation algorithm. This eliminates the necessity for multiple full-scan reconstructions at this stage. To enhance the quality of the reference reconstruction, a regularization technique is incorporated. The motion compensation process involves estimating motion parameters through a regularized 3D-2D image registration. The outcomes of this methodology highlight its effectiveness in compensating for even substantial and prolonged motions. This approach not only simplifies the compensation process but also showcases its robust performance in the face of diverse motion challenges.

Acknowledgments

I thank **Allah**, the almighty and the most merciful, for granting me wisdom, courage, and strength to be who I am.

I am very much thankful to my **Parents** and my whole **family** who were always been there for me, praying for me and never letting me down.

I would like to express my sincere gratitude to my supervisor, **Prof. Andrea FUSIELLO** and my advisors **Cristina SARTI** and **Claudio LANDI** (See Through s.r.l., Italy) for supporting me throughout my Ph.D.

I would also like to thank my mentor, **Prof. David FOFI** (University of Burgundy, France) for all his support during my research stay in France.

I'm pleased to acknowledge **Michele ANTONELLI, Ivan TOMBA, Lorenzo ARICI, Andrea DELMIGLIO, Luca FRACASSETTI** and all colleagues from See Through s.r.l. for providing the necessary technical support.

Finally, I wish to thank everyone else who supported me directly or indirectly during my Ph.D.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Cone Beam Computed Tomography (CBCT) is a type of imaging technology that uses a cone-shaped X-ray beam to capture and reconstruct detailed three-dimensional (3D) images of hard tissues, like bones and teeth. CBCT imaging has emerged as an invaluable tool in modern dentistry, providing high-resolution 3D images of the oral and maxillofacial regions. These images are crucial for accurate diagnosis, treatment planning, and assessment of various dental and craniofacial conditions. However, like any imaging modality, CBCT is susceptible to adverse effects due to various factors. Among these factors, patient motion is one of the major causes of artifacts in oral and maxillofacial CBCT imaging. The longer scan duration of CBCT imaging largely increases the likelihood of patient movement. Many efforts have been made to prevent patient motion during CBCT acquisition. Patients are immobilized using head supports, chin rests, and a fixed bite block. However, head fixations may not prevent all potential motions (Hanzelka et al., 2013), since as small as $3mm$ of motion displacement can significantly affect the image quality (Spin-Neto et al., 2018). To address this challenge, the concept of motion compensation has gained significance in the field of maxillofacial CBCT imaging. Motion compensation minimizes the necessity for scan repetition, subsequently reducing the overall radiation exposure to the patient.

Motion compensation encompasses a variety of techniques and strategies designed to mitigate or correct the artifacts caused by patient motion during the scanning procedure. Existing solutions for motion compensation often rely on complex algorithms, and some of these approaches involve the iterative reconstruction of the entire set of projections. In fact, this iterative nature can be computationally intensive and time-consuming, posing practical challenges for real-time or near-real-time applications. Notable works in the literature, such as those by Maur et al. (2019), Sun et al. (2021), and Birklein et al. (2023), showcase the complexity associated with these algorithms.

A prevalent limitation of many existing motion compensation algorithms is their effectiveness primarily against short-duration motions (Sun et al., 2021). Prolonged or complex motions may pose challenges that current methods struggle to address adequately. The complexity of patient movements demands robust compensation strategies that can accommodate a spectrum of motion scenarios. Moreover, inconsistencies stemming from data truncation and scatter are common challenges faced in CBCT imaging. These issues can compromise the accuracy of motion compensation algorithms, particularly when dealing with incomplete or scattered data. As a result, there is a critical need for innovative approaches that not only address motion compensation but also account for the specific challenges posed by the unique characteristics of CBCT data, ensuring reliable and artifact-free reconstructions. In this thesis, we introduce one such approach for motion compensation, specifically designed for maxillofacial CBCT imaging with detailed methodology and evaluation.

1.2 Challenges and Constraints

There are a few challenges related to motion compensation in maxillofacial CBCT imaging.

1. **Low dose data:** CBCT holds an advantage over traditional CT in terms of subjecting patients to significantly lower radiation doses. Nevertheless, this presents a challenge for motion compensation, as the lower dose leads to reduced contrast resolution in the acquired projections. This reduction in spatial resolution poses a difficulty for the compensation algorithm in detecting edges of the hard tissues, especially in projections captured in the anteroposterior view.

2. **Time Complexity:** Currently, there are algorithms capable of rapidly performing CBCT reconstruction in a matter of seconds. Therefore, it is anticipated that the compensation algorithm should align with the speed of the reconstruction algorithms to meet the demands of clinical applications.

3. **Patient Pose:** The majority of maxillofacial CBCT devices available on the market are designed for patients to be standing, increasing the likelihood of patient movements.

4. **High-Resolution Reconstructions:** The reconstructions of maxillofacial CBCT demand high resolution, given their critical role in treatment planning for procedures like implantology and oral surgery. Consequently, even minor motions can have a substantial impact on the quality of the reconstruction.

Considering these challenges, we have included a few simplified assumptions aligned with the nature of CBCT devices.

1. **Visualization of Hard Tissues:** Maxillofacial CBCT devices capture greater detail in hard tissues compared to soft tissues. As a result, movements of soft

tissues, such as swallowing, have minimal impact on the diagnostic quality of the images. Therefore, in the context of motion compensation, our focus is on addressing rigid motions like nodding and tilting, excluding soft tissue movements.

2. **Head fixation:** Common accessories for maxillofacial CBCT devices include mechanical head fixations like fixed bite blocks and chin rests. These fixations can effectively prevent jaw motions; hence, jaw motions are not taken into consideration.

1.3 Research Objectives

This research aims to develop methods that effectively compensate for motion artifacts, enhancing the overall image quality and improving the diagnostic capabilities of maxillofacial CBCT reconstructions.

We have defined several criteria that must be met in the course of this research:

- The compensation method must demonstrate speed and robustness.
- It should effectively address prolonged and persistent motion scenarios.
- The method should be applicable for both full-scans ($360°$) and short-scans ($\approx 180°$).

1.4 Thesis Structure

This thesis is organized into several chapters, systematically addressing the research objectives. Chapter 2 offers an extensive literature review covering the background and principles of maxillofacial CBCT imaging, the influence of motion artifacts on image quality, and existing motion compensation techniques. Chapter 3 provides details about the device and algorithms utilized in this research.

Chapter 4 outlines the methodology, presents experimental results, and evaluates performance with both quantitative and qualitative assessments for the proposed motion compensation technique. Chapter 5 outlines the research methodology, encompassing data collection, preprocessing, and the proposed algorithm for motion detection, along with corresponding experimental results. Chapter 6 introduces an alternative strategy for motion compensation without relying on motion detection.

Chapter 7 delves into motion compensation implemented for short-scan reconstructions. Finally, in Chapter 8, the thesis concludes with a summary of research outcomes and their implications, provides insights into the significance of the findings, and suggests potential future directions.

1.5 Publications

Part of the materials contained in this thesis have been/yet to be published in the following works:

- Asraf Ali, Abdul Salam Rasmi, Cristina Sarti, and Claudio Landi. "Compensation for Patient Movements in CBCT Imaging for Dental Applications." *International Conference on Image Analysis and Processing*. Cham: Springer Nature Switzerland, 2023.

- Asraf Ali, Abdul Salam Rasmi, Claudio Landi, Cristina Sarti, and Andrea Fusiello. "Non-iterative Compensation for Patient Motion in Dental CBCT Imaging." *2023 IEEE Nuclear Science Symposium, Medical Imaging Conference and International Symposium on Room-Temperature Semiconductor Detectors (NSS MIC RTSD)*. IEEE, 2023.

- Asraf Ali, Abdul Salam Rasmi, Cristina Sarti, Claudio Landi, and Andrea Fusiello. "Motion Compensation in Short-scan CBCT Reconstructions for Dental Applications." *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. 2024.

Chapter 2

Background

2.1 Tomographic Imaging

Tomographic imaging refers to a diagnostic imaging technique that produces cross-sectional images of an object or living organism by assembling data from multiple two-dimensional images. This method enables the visualization of internal structures in detail, providing valuable insights into the composition and characteristics of the imaged subject. The imaging process can be performed using different methods and technologies, including fluorescence imaging (Zeng et al., 2022), x-ray imaging (Gupta et al., 2021), wavefield measurements (Mansour et al., 2021), and electromagnetic inverse problems (Bhadra et al., 2022). Tomographic imaging techniques, such as computed tomography (CT) and magnetic resonance imaging (MRI), have become indispensable tools in various fields, including medicine, biology, and materials science.

In medical applications, tomographic imaging allows clinicians to obtain detailed images of anatomical structures, aiding in the diagnosis and treatment of diseases. CT scans, for instance, use X-rays to create detailed cross-sectional images of the body, offering valuable information for identifying abnormalities, tumors, or fractures (Bushberg and Boone, 2011). Magnetic resonance imaging, on the other hand, utilizes strong magnetic fields and radio waves to generate detailed images without ionizing radiation, making it particularly suitable for soft tissue imaging (Haacke, 1999). The evolution of tomographic imaging technologies has led to advancements such as cone beam computed tomography (CBCT), which provides three-dimensional images with reduced radiation exposure, making it well-suited for applications in dentistry and orthopedics (Patel et al., 2015).

2.2 Computed Tomography

Computed tomography (CT) is an imaging modality that has greatly improved diagnostic and interventional radiology (Lell and Kachelrieß, 2023). It has undergone significant advancements in terms of scan speed, spatial and soft tissue resolution, and dose reduction. CT technology has also seen the integration of photon-counting detectors and the use of artificial intelligence in patient positioning, protocol adjustment, and image reconstruction (Ibad et al., 2023). In the field of musculoskeletal imaging, CT techniques such as 4-dimensional CT, cone-beam CT, and dual-energy CT have emerged as valuable tools for diagnosing and characterizing various conditions, including trauma, gout, and pathological biomechanical states (Saad and Moore, 2022). CT scans offer higher contrast resolution and improved tissue differentiation compared to standard radiography, making them useful in the diagnosis of medical emergencies (Gallastegui, 2022).

2.2.1 Working Principle of CT

A CT scan operates on the principle of utilizing X-rays to create detailed cross-sectional images of the human body. The process begins with the emission of a fan-shaped X-ray beam from an X-ray tube. As the X-rays traverse the body, different tissues absorb varying amounts of radiation, resulting in a contrast of densities. On the opposite side of the body, a detector array collects the transmitted X-rays. Multiple measurements are taken from various angles around the body, and this raw data is then processed by a computer using sophisticated algorithms. The computer reconstructs the collected data into a series of two-dimensional cross-sectional images, known as slices, through a process called tomographic reconstruction. The resulting three-dimensional images provide detailed anatomical information, enabling healthcare professionals to diagnose and assess various medical conditions (Hsieh, 2003).

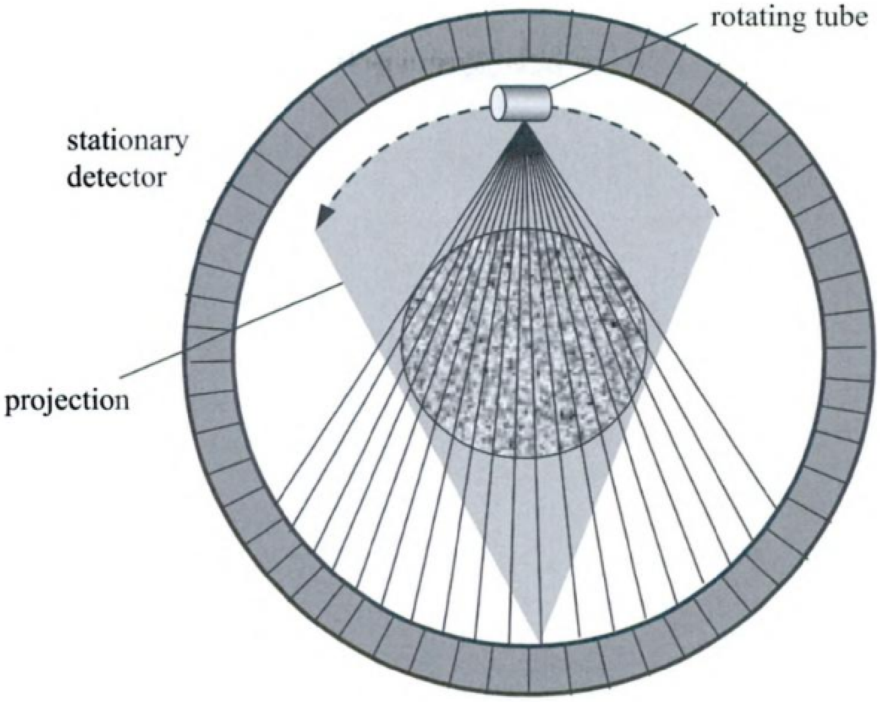


Figure 2.1: Geometry of a CT scanner (Hsieh, 2003)

2.2.2 Artifacts in CT

In CT imaging, artifacts are unwanted discrepancies or distortions in the reconstructed images that can arise from various sources, including patient-related factors, equipment issues, or image reconstruction algorithms. Understanding these artifacts is crucial for accurate diagnosis and interpretation of CT scans. The following are some of the common artifacts:

- **Beam Hardening Artifacts:** Beam hardening is a significant artifact in CT caused by the polychromatic nature of X-ray sources. It results in a nonlinear relationship between attenuation and material thickness, leading to image degradation. Various methods have been proposed to correct beam hardening artifacts in CT. One approach is to enforce consistency conditions on projection pairs, such as the Grangeat consistency condition (GCC), Smith consistency condition (SCC), and fan-beam consistency condition (FBCC) (Abdurahman et al., 2022). Another method involves using energy-integrating detectors (EID) or photon-counting detectors (PCD) subsystems, with EID-CT images showing fewer beam hardening artifacts compared to PCD-CT images (Holmes et al., 2022). Additionally, a beam-hardening correction method for lab-based CT has been developed by modifying existing iterative tomographic reconstruction algorithms, resulting in improved self-consistency of projected attenuation and quantitative tomograms (Yang et al., 2020). Furthermore, the correction of beam hardening can be achieved through filtration and linearization methods, with linearization offering better correction without changing signal-to-noise ratio (Ahmed et al., 2020). Finally, the attenuation of polychromatic beams can be calculated and corrected using the polynomial fitting, transforming the data into equivalent monochromatic data (Ahmed et al., 2019).

- **Metal Artifacts:** Metal artifacts in CT imaging are a significant challenge as they degrade image quality and hinder accurate diagnosis. Previous methods for reducing metal artifacts either require prior knowledge of the location of metal implants or have modeling deviations, limiting the ability to obtain high-quality CT images (Lee et al., 2021; Osipov et al., 2023). However, recent studies have proposed novel approaches to address this issue. One study explored the clinical potential of a photon-counting detector (PCD) CT system in reducing metal artifacts in head CT scans (Sharma et al., 2021). Su et al. (2023) proposed RetinexFlow which is a novel end-to-end image domain model. It formulates the problem as a combination of decomposition and completion tasks and is based on Retinex theory and conditional normalizing flow. The model incorporates a feature decomposition encoder to decompose the metal implant component and extract the inherent feature, followed by a feature-to-image flow module to complete the metal artifact-free CT image through invertible transformations. These advancements in metal artifact reduction techniques have the potential to improve image quality and enhance the diagnosis and treatment of patients.

- **Ring Artifacts:** Ring artifacts in CT are a common issue caused by the nonuniformity of detector pixels. These artifacts compromise image quality and result in nonuniform bias. Several methods have been proposed to address this problem. An et al. (2020) introduced a fast ring artifacts removal method that corrects the inconsistency of all detector pixels simultaneously without additional scans, preserving image details without extra computational cost. Šalplachta et al. (2021) developed a complex and robust approach that differentiates two types of ring artifacts and addresses them separately in the sinogram domain, achieving high efficiency of ring removal and preservation of spatial resolution. Murata and Ogawa (2020) proposed a compressed sensing method that corrects ring artifacts in material images with less degradation of image quality, showing better results compared to previous methods. Hein et al. (2022) used deep learning techniques to train a UNet to remove streak artifacts in the sinogram domain, successfully producing ring-corrected virtual monoenergetic images. Another study proposed a ring artifact correction method based on a residual neural network (ResNet) that effectively suppresses ring artifacts in X-ray tomography (Fu et al., 2023).

- **Motion Artifacts:** Motion artifacts in CT refer to distortions or blurring in CT images caused by patient motion during the scanning process. These artifacts can degrade the quality and diagnostic performance of CT images. Various methods have been proposed to address this issue. Gupta et al. (2023) developed a reconstruction method called DiFiR-CT that generates artifact-free images without explicit motion estimation. Ko et al. (2021) proposed a real-time technique that utilizes a deep residual network with an attention module to reduce motion artifacts in CT images. Deng et al. (2021) used a generative adversarial network (GAN) to correct motion artifacts in coronary CT angiography (CCTA) images. The rapid progress in CT hardware has also enabled such fast scans that the beating heart can now be imaged without artifacts, unaffected by respiratory motion (Zhang et al., 2022).

2.3 Cone Beam Computed Tomography (CBCT)

Cone Beam Computed Tomography (CBCT) stands as a revolutionary advancement in medical imaging, offering a three-dimensional perspective that has significantly enhanced diagnostic capabilities in various clinical specialties. Rooted in the principles of computed tomography, CBCT employs a cone-shaped X-ray beam and a specialized detector to produce detailed and accurate 3D images of anatomical structures (Buchanan, 2023). This imaging modality finds extensive applications in dentistry, maxillofacial surgery, orthodontics, and ENT disciplines, providing clinicians with invaluable insights into the intricacies of the head and neck region.

2.3.1 CT vs CBCT

2.3.1.1 Advantages of CBCT over CT

1. Lower Radiation Dose: CBCT uses a cone-shaped X-ray beam, resulting in a lower radiation dose compared to CT (Scarfe et al., 2006).

2. Improved Spatial Resolution: CBCT offers higher spatial resolution, allowing for more detailed visualization of small anatomical structures (Scarfe et al., 2006).

3. Cost-Effectiveness: CBCT systems are generally more affordable and compact compared to CT systems (Scarfe et al., 2006).

4. Ease of Use: CBCT systems are designed specifically for dentistry and maxillofacial imaging, providing specialized software and tools (Scarfe et al., 2006).

2.3.1.2 Disadvantages of CBCT compared to CT

1. Lower Volumetric Coverage: CBCT has a limited field of view (FOV) compared to CT (Scarfe et al., 2006).

2. Lower Soft Tissue Contrast: CBCT may have lower soft tissue contrast compared to CT (Scarfe et al., 2006).

3. Limited Dynamic Range: CBCT may have a limited dynamic range, resulting in potential image artifacts or overexposure (Scarfe et al., 2006).

4. Reduced Multiplanar Reconstruction: CBCT has limited capability for multiplanar reconstruction compared to CT (Scarfe et al., 2006).

2.4 Maxillofacial CBCT Imaging

Maxillofacial Cone Beam Computed Tomography (CBCT) is a specialized imaging technique that provides high-resolution three-dimensional images of the maxillofacial region (Nickenig et al., 2006; Scarfe and Farman, 2008). It has gained significant importance in various fields, including dentistry, maxillofacial surgery, and orthodontics. CBCT imaging offers numerous advantages over conventional radiography, such as improved visualization of anatomical structures, accurate diagnosis, and treatment planning (Kumar et al., 2007; Ludlow and Ivanovic, 2008).

The maxillofacial region comprises the bones of the face, skull, and jaws, as well as the associated soft tissues. Maxillofacial CBCT better visualizes the hard tissues than the soft tissues and plays a crucial role in the evaluation and management of various maxillofacial conditions, including dental pathologies, trauma, orthognathic surgery, and implant placement (Loubele et al., 2006; Motro et al., 2019). It enables detailed visualization of complex anatomical structures, such as the temporomandibular joint, maxillary sinuses, and dental roots, aiding in the precise diagnosis and treatment of patients.

2.4.1 Acquisition Process

During the acquisition process, the patient is positioned in a standing, seated, or supine position, depending on the specific CBCT machine. The patient's head is placed in a stable position using head fixation devices to minimize movement artifacts. The CBCT unit rotates around the patient's head, capturing a series of X-ray projections at various angles (Figure 2.3a). These projections are acquired in a single rotation, typically spanning $180°$ to $360°$. The X-ray beam and detector are synchronized to capture the necessary data (Hartshorne, 2018).

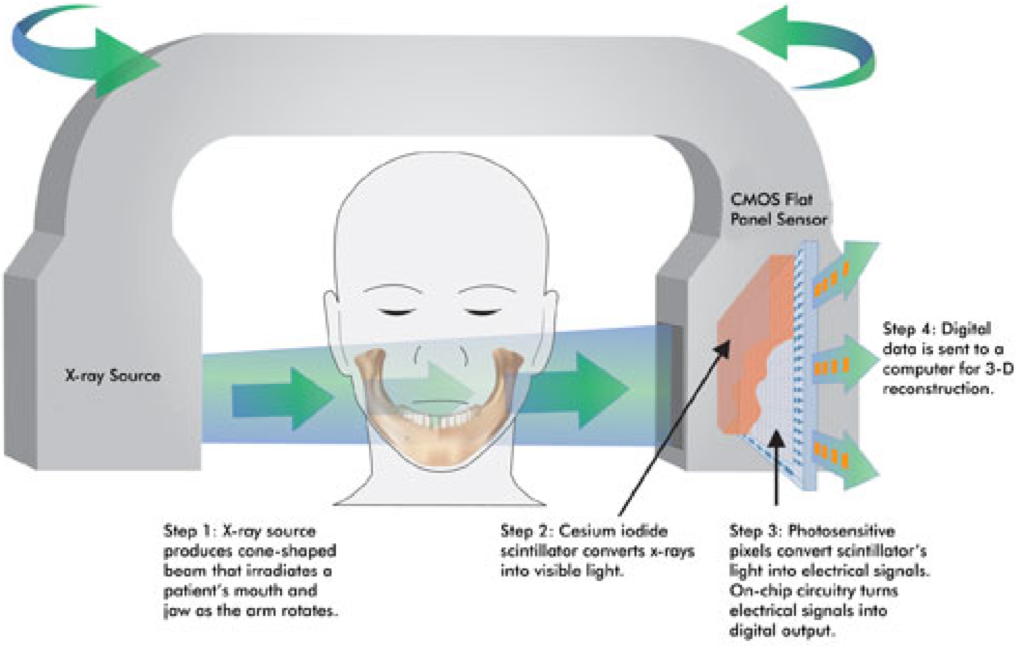
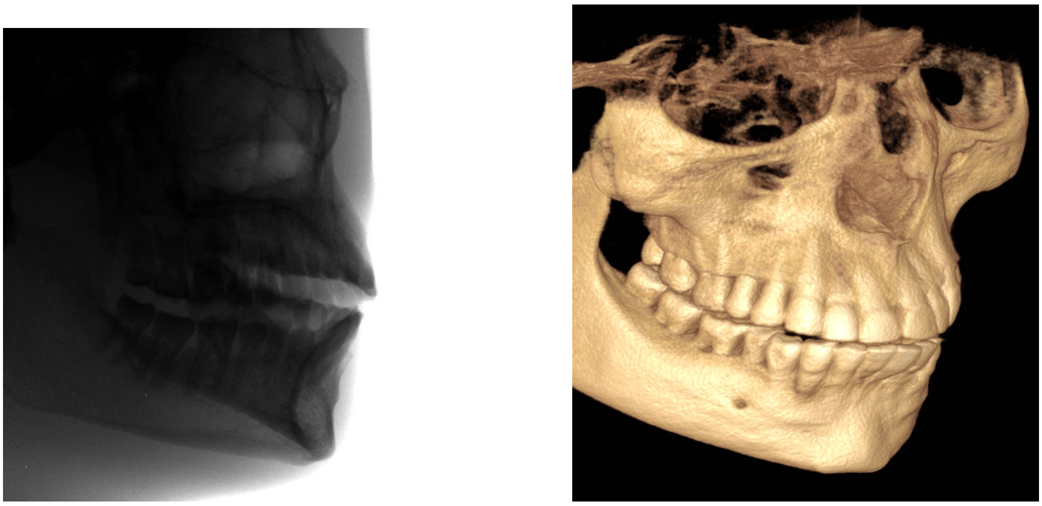


Figure 2.2: Illustration of Data Acquisition (Hartshorne, 2018)

2.4.2 Image Reconstruction Techniques

Once the projections are acquired, specialized software is used to reconstruct the CBCT volume. Image reconstruction techniques, such as Filtered Back Projection (FBP) (Feldkamp et al., 1984) or iterative reconstruction algorithms (Andersen and Kak, 1984), are employed to convert the acquired projection data into a 3D volume. FBP is a commonly used technique that applies a series of mathematical operations to convert the acquired data into voxel values, resulting in a 3D volume (Figure 2.3b). Iterative reconstruction algorithms, on the other hand, iteratively refine the initial estimate of the 3D volume to improve image quality and reduce artifacts.



(a) Projection (b) Reconstructed 3D Volume

Figure 2.3: CBCT Image Reconstruction

2.4.3 Common Applications in Dentistry and Maxillofacial Surgery

Maxillofacial CBCT imaging finds extensive applications in dentistry and maxillofacial surgery. Some common applications include:

1. Orthodontics: CBCT imaging allows for accurate assessment of dental and skeletal relationships, aiding in treatment planning for orthodontic procedures, such as tooth extractions, orthognathic surgery, and placement of orthodontic implants (Botticelli et al., 2011).

2. Implant Dentistry: CBCT scans provide detailed information about bone quantity, quality, and anatomical structures, enabling precise implant placement planning. It helps identify the ideal implant sites, assess bone density, and evaluate proximity to vital structures like nerves and sinuses (Tyndall et al., 2012).

3. Endodontics: CBCT imaging aids in diagnosing and managing complex endodontic cases. It enables visualization of root canal morphology, identifi-

cation of root fractures, detection of periapical lesions, and assessment of treatment outcomes (Patel et al., 2009).

4. Maxillofacial Trauma: CBCT scans provide detailed 3D information about facial fractures, allowing accurate evaluation and surgical planning for maxillofacial trauma cases.

5. Temporomandibular Joint (TMJ) Disorders: CBCT imaging helps in assessing TMJ pathologies, including condylar position, internal derangements, and joint space abnormalities. It aids in treatment planning for TMJ disorders and orthognathic surgeries involving the TMJ.

6. Oral and Maxillofacial Pathology: CBCT scans assist in the evaluation of various pathologies, such as cysts, tumors, and impacted teeth. It provides detailed anatomical information for surgical planning and evaluation of treatment outcomes (Scarfe et al., 2006).

2.4.4 Artifacts in Maxillofacial CBCT

Similar to traditional CT imaging, cone-beam computed tomography (CBCT) is vulnerable to various artifacts that may compromise the clarity and quality of images. Some common artifacts are scatter artifacts, ring artifacts, cone-beam artifacts, and motion artifacts (Sinha et al., 2016). While numerous techniques have been devised to successfully mitigate the impact of most of these artifacts, addressing motion artifacts remains a challenge. Given the prolonged acquisition time associated with maxillofacial CBCT, the occurrence of motion during scanning is unavoidable.

2.5 Motion Artifacts in Maxillofacial CBCT

Unlike traditional computed tomography (CT), CBCT systems have acquisition time ranging from 5.4 to 40 seconds (Nemtoi et al., 2013), which is often long enough for significant patient motion to occur. It is estimated that approximately 21–42% of the in vivo examinations exhibit motion artifacts (Spin-Neto et al., 2015). Patients move during the scan for various reasons such as fear of the tube/detector movement (Yıldızer Keriş, 2017), Parkinson's disease (Donaldson et al., 2013), or simply because it is difficult for someone to stay idle for quite a long time, especially children (Spin-Neto et al., 2015).

2.5.1 Impact of motion artifacts on image quality

The presence of motion artifacts can have several implications for image quality and diagnostic accuracy. Motion artifacts introduce blurring, streaking, and ghosting artifacts, leading to reduced spatial resolution and image distortion. This degradation in image quality can hinder the accurate visualization of anatomical structures and the detection of pathologies. Moreover, motion artifacts can introduce false or misleading information, potentially leading to diagnostic errors (Kau et al., 2010).

2.5.2 Prevention of Motion

In maxillofacial CBCT, a crucial component for ensuring accurate and artifact-free scans is the head restraint system. Engineered to stabilize and immobilize the patient's head during imaging, this system minimizes unintended movement that could otherwise introduce motion artifacts and compromise the quality of the images. The head restraint employs carefully designed supports, including forehead rests, fixed bite blocks, and chin rests, to secure the patient's head in a stable position. This immobilization is essential to maintain the precise alignment required for acquiring clear and detailed images of the maxillofacial region.

Studies, such as the one by Nardi et al. (2017), indicate a notable reduction in scan repetitions attributed to the effectiveness of head restraints. However, it's important to note that these devices exhibit a degree of flexibility to accommodate patient comfort, and as a result, complete elimination of all motion is not achievable. Consequently, reliance on head fixation devices may not entirely prevent potential motions, given that even minimal displacements as small as $3mm$ can significantly impact image quality (Spin-Neto et al., 2018).

2.6 Motion Compensation Techniques

Numerous correction methods are available for general CBCT applications, but only a limited number of algorithms have been specifically designed to address patient movements during maxillofacial CBCT imaging. In this review, we will comprehensively examine these algorithms to provide an in-depth understanding. For clarity, these methods are systematically categorized into four distinct groups, as outlined below.

2.6.1 Optimization of Data Fidelity Terms

Wicklein et al. (2013) developed an online motion and misalignment correction method for medical flat-detector CT. The method aims to correct misalignment artifacts and motion artifacts during the reconstruction process without the need for a calibration phantom or predefined scan protocols. The method utilizes image features such as entropy and gray-level histogram-based methods to quantify misalignment and optimize the geometry parameters of the acquisition system.

Sisniega et al. (2017) presented a statistical optimization algorithm to estimate a motion trajectory that optimizes an objective function consisting of an image sharpness criterion augmented by a regularization term that encourages smooth motion trajectories. The method supports application to multiple locally rigid regions and does not require a priori motion models, external trackers, or fiducials (Sisniega et al., 2016). The algorithm involves an initial coarse estimation of gross motion followed by estimation of fine-scale displacements using high-resolution reconstructions. It has been shown to significantly reduce motion-induced artifacts and blurring across a broad range of motion amplitudes, from 0.5 mm to 10 mm, and improve the delineation of tissue boundaries and trabecular structures in extremity CBCT (Sisniega et al., 2019). The motion compensation algorithm has also been found to improve the visualization of bone and soft tissue structures in extremity CBCT for cases exhibiting patient motion.

2.6.2 Auto-calibration based Approaches

CBCT auto-calibration by contour registration is a strategy for estimating projection matrices using object contours in projection images. It involves formulating mathematical conditions for projection parameter consistency based on contour detection and tracking in 2D (Maur et al., 2016). The approach utilizes iterative reconstruction and registration to calculate geometrical projection parameters from unknown patient geometry, resulting in a robust correction for various types of patient motion (Maur et al., 2018). The method is implemented as a global optimization problem and includes a greedy optimizer for efficient computation (Maur et al., 2019). It is robust towards inaccurate initialization, truncation, noise, and other typical artifacts of CBCT reconstruction. However, a notable constraint of this method lies in the challenging task of detecting the contour, with no explicit solution presented.

2.6.3 3D-2D Registration based Approaches

3D-2D Registration approaches are quite popular for motion compensation in maxillofacial CBCT imaging. Ouadah et al. (2017) introduced a fiducial-free approach employing 3D-2D image registration to address motion artifacts in CBCT. Through the registration of each projection to a pre-existing 3D image, the method corrects residual errors in geometric calibration parameters, leading to a significant enhancement in image quality demonstrated in phantom experiments and clinical images. Niebler et al. (2019) introduced a marker-free iterative motion correction algorithm designed for local tomography, operating on projection images. Utilizing a fast GPU-accelerated 3D reconstruction algorithm and assuming rigid motion, the method minimizes a pixel-wise cost function between X-ray images and parameterized projections, effectively correcting motion.

Vijayan et al. (2022) implemented a method for precise target localization in fluoroscopically-guided pulmonary interventions using cone-beam CT (CBCT) and 3D-2D registration. The algorithm involves three steps: CBCT-to-fluoroscopy rigid registration, local rigid registration of lung-thresholded CBCT to fluoroscopy within a region of interest (ROI), and aggregation of local registrations for global deformation estimation. With soft-tissue thresholding and contrast enhancement, the proposed method reduces target registration error from 10.3 mm to 3.8 mm, making it a potential tool for enhancing accuracy in pulmonary interventions by updating target overlays in fluoroscopy.

Sun et al. (2021) developed a retrospective motion compensation method for cone-beam CT imaging, adapting a previous technique used in helical head CT imaging. The proposed motion estimation/motion compensation (ME/MC) method enhances reconstructed image quality and effectively eliminates most artifacts, as demonstrated in simulations, phantom studies, and patient evaluations. The primary issue with these methods stems from the utilization of iterative reconstruction algorithms, leading to time-consuming processes and imposing a substantial computational cost.

2.6.4 Deep Learning Approaches

Deep Learning has evolved as an important tool in medical image processing. Several works have been published related to motion compensation in cone-beam CT using deep learning techniques. Birklein et al. (2023) introduced a combined motion estimation and motion correction algorithm based on 2D projection images from a single scan. This algorithm uses a deep neural network to segment the cranium and mandible in the projection images, and an iterative three-step algorithm to model the two articulated motions as independent rigid motions. Another approach involves using deep convolutional neural networks (CNN) as pre- and post-processing steps during CBCT reconstruction (Amirian et al., 2023). The neural networks are trained end-to-end using motion-corrupted and motion-free CBCT data to reduce motion-induced artifacts and improve image quality.

Huang et al. (2022) used a reference-free similarity metric called DL-VIF, which leverages deep CNNs to quantify motion-induced image quality degradation and distortion of anatomical structures in CBCT. Additionally, they proposed a novel optimization strategy called Multi-Stage Adaptive Spine Autofocus (MASA) for the compensation of complex deformable motion in abdominal CBCT (Huang et al., 2023). MASA uses a multi-stage adaptive sampling strategy of the motion trajectory, sampled with Hermite spline basis with variable amplitude and knot temporal positioning. Although these algorithms demonstrate promising outcomes on their specific datasets, the broader effectiveness, particularly in the context of maxillofacial motion compensation, remains an open question.

2.7 Motion Detection

Although several works have been published related to the compensation for patient motion, only a few focused on detecting the motion artifacts in CBCT Imaging. Ens et al. (2010) proposed an automatic detection of patient motion in the CBCT projection images using different similarity measures like Structural Similarity Index Measure and Mutual Information. The drawback of this work is that the motion is detected by comparing two consecutive projections at a time which is time-consuming. Also, this method might not be robust to changes in the quality of projections which heavily depends on radiation dose. Sun et al. (2020) presented a motion artifact detection algorithm based on the Convolutional Neural Network (MADA-CNN) implemented with transfer learning and ensemble modeling. While their method has achieved an AUC-ROC value of 0.966 for motion detection in slices from the reconstructed volumes, the effects on full 3D volumes have not been investigated. Welch et al. (2020) and Arrowsmith et al. (2021) used a 3D Convolutional Neural Network (CNN) for detecting dental artifacts directly in the reconstructed CBCT volumes which is much faster than detecting such artifacts in projection images but training a 3D CNN is computationally expensive. Furthermore, artifacts caused by different reasons need different compensation techniques. Grouping all artifacts together can make it difficult to determine the best approach for each individual case.

Chapter 3

Materials and Methods

3.1 Device

The imaging data were obtained using a head CBCT scanner located at See Through s.r.l., Brusaporto, Italy (Figure 3.1). This apparatus features a gantry that moves along a 360° trajectory around the patient's head. The patient is positioned in a way that the head aligns with the isocenter of the device and is securely fixed using a head fixation mechanism. As the gantry moves, X-rays traverse through the head, getting attenuated and subsequently captured by a flat panel detector (Figure 3.2). The acquisition process involves acquiring 720 projection images over a span of 24 seconds. The geometric configuration of the device and the positioning of the patient are known and are employed in the subsequent reconstruction procedure. The resulting reconstructed field of view is sufficiently extensive to encompass the entire dental arch (Figure 3.3).

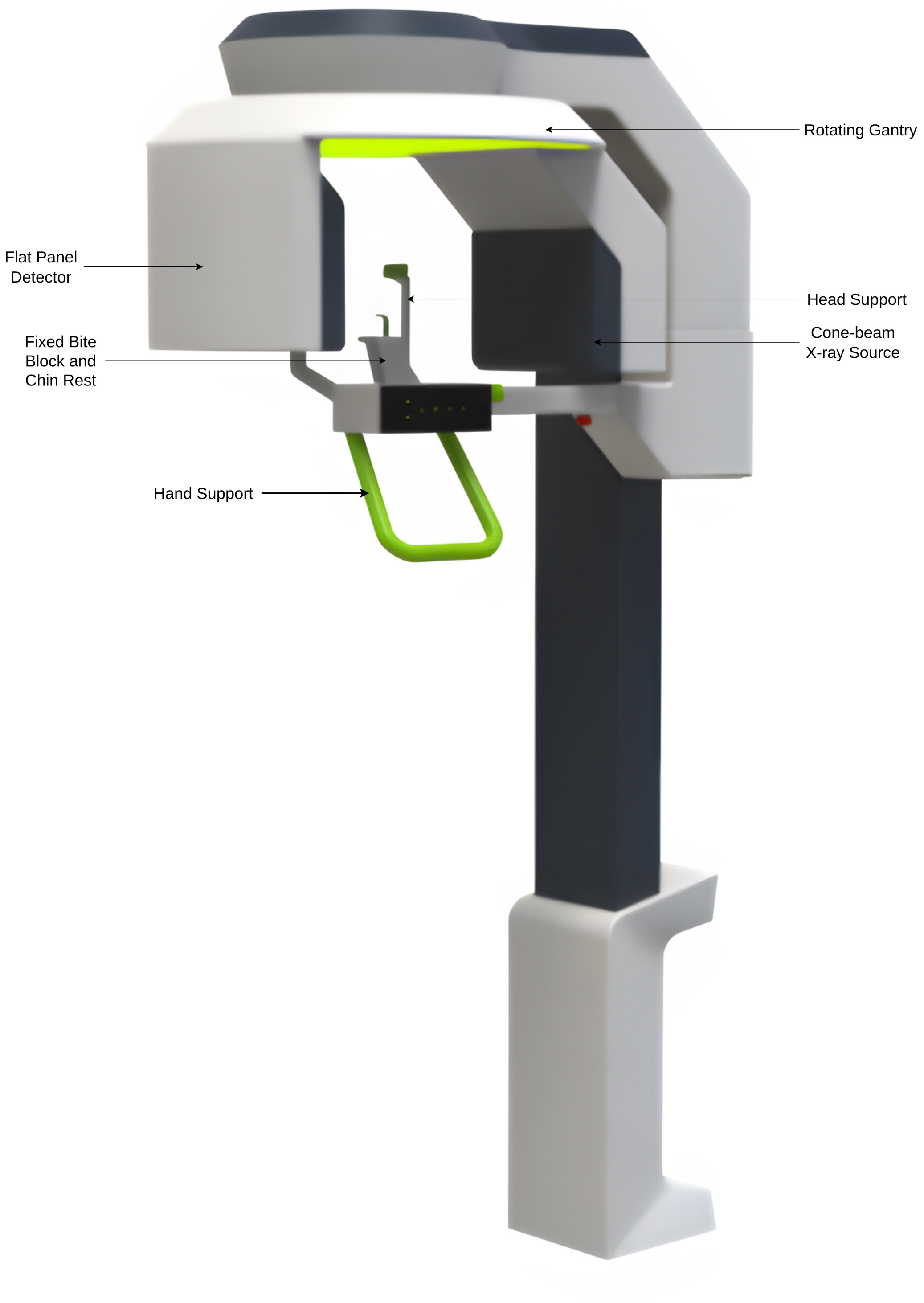


Figure 3.1: Maxillofacial CBCT device from See Through s.r.l.®

Figure 3.2: Example of a flat panel detector

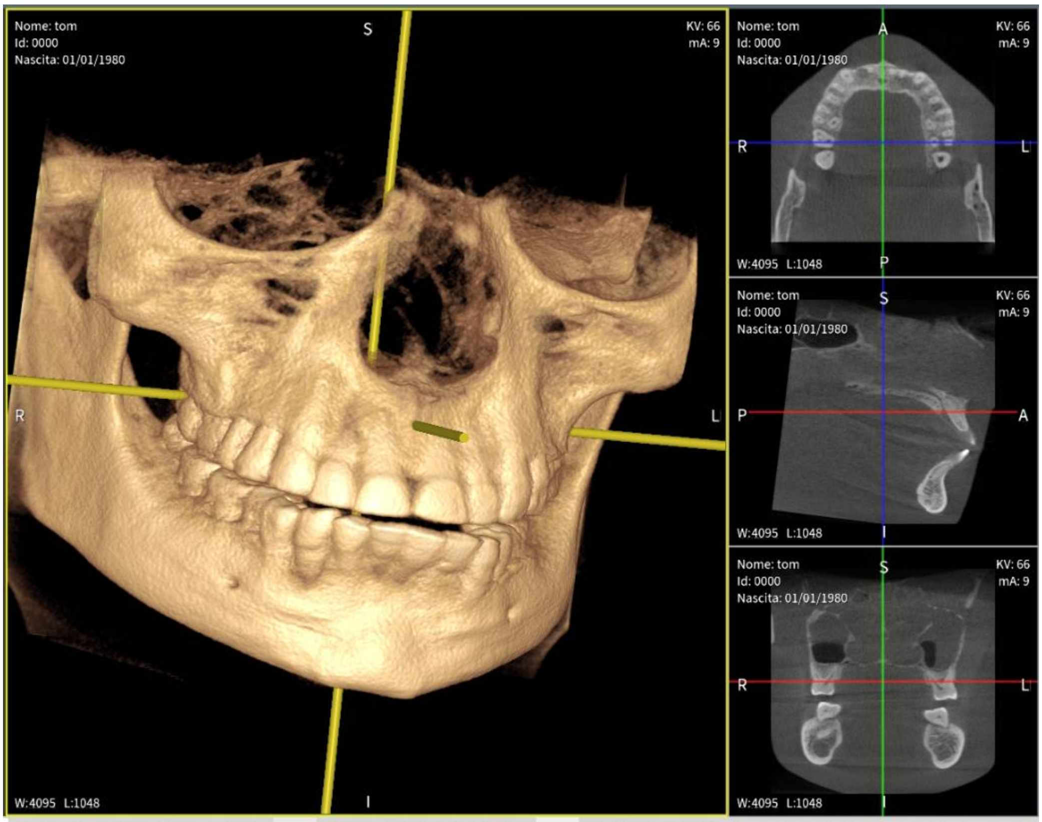


Figure 3.3: Reconstruction of the full dental arch

3.1.1 Patient Immobilization Accessories

- **Head Restraints:** Mechanical fixations, such as head restraints or chin and forehead supports, minimize patient movement during the scan, reducing motion artifacts.

- **Bite Blocks:** Customized bite blocks standardize the relationship between the upper and lower jaws, ensuring consistent positioning for accurate imaging.

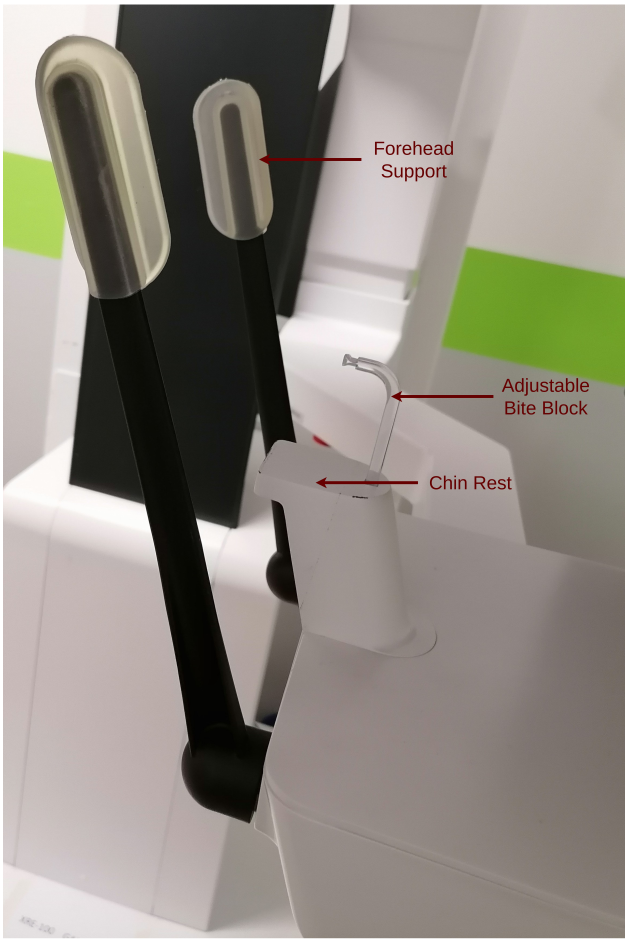


Figure 3.4: Prototype of Head Restraints used in the CBCT device at See Through s.r.l.

3.2 Data

In this research, three head phantoms were used: $Jerry$, a small phantom; Tom, slightly larger in size compared to $Jerry$; and $Silvestro$, the largest among the three (Figure 3.5). The $Silvestro$ phantom contains metal objects such as crowns and fillings, introducing non-motion-related artifacts like beam hardening-induced over or under-estimation of attenuation and photon starvation-induced noise. These phantoms are original skulls obtained from deceased patients, enveloped in a plastic resin that replicates the characteristics of soft tissues in terms of X-ray attenuation. The use of phantoms is particularly advantageous for conducting maxillofacial studies with cone-beam computed tomography (CBCT) as they facilitate experiments in a setup resembling real-world clinical scenarios. Given that CBCT excels in visualizing hard tissues (bones) over soft tissues, utilizing a phantom ensures that experimental outcomes closely resemble real-world conditions.

(a) $Silvestro$ (b) Tom (c) $Jerry$

Figure 3.5: Head phantoms used in this research

3.3 FDK Reconstruction Algorithm

The Feldkamp-Davis-Kress (FDK) reconstruction algorithm (Feldkamp et al., 1984) is a widely used technique in CBCT imaging for reconstructing three-dimensional images from two-dimensional projection data. The algorithm follows a Filtered Back Projection approach, combining two essential steps. Firstly, the acquired projection data, which represents X-ray attenuation through the object from various angles, undergoes a filtering process. This step involves applying a filter in the frequency domain to emphasize high-frequency components and suppress low-frequency noise. The filtering aims to enhance image contrast and reduce artifacts in the reconstructed image. An example comparison between reconstruction using simple back projection vs filtered back projection is shown in Figures 3.6 and 3.7.

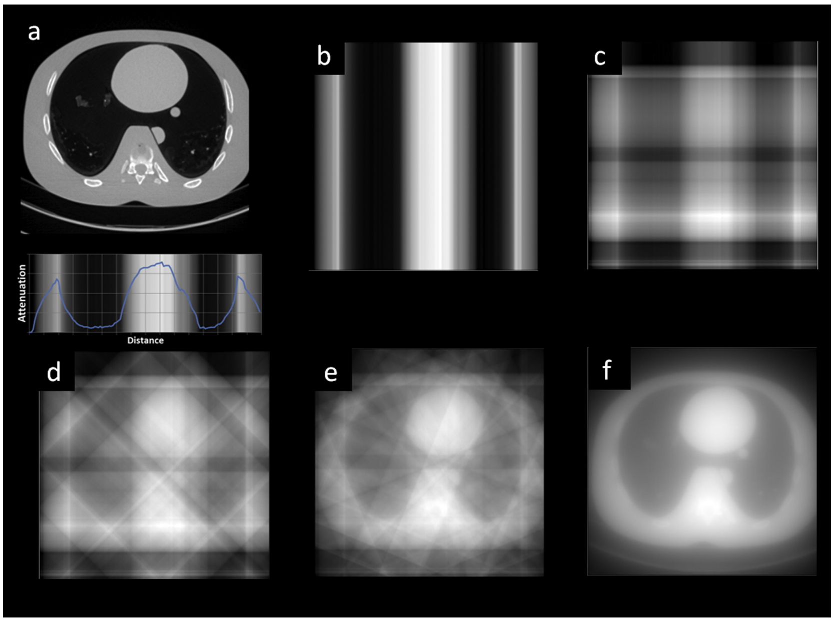


Figure 3.6: Simple Back projection (a) an anthropomorphic chest phantom. (b) back projection from a single projection image, (c) from two perpendicular projection images, (d) from four projection images, (e) from eight projection images, and (f) from 720 projection images. The simple back projection process inherently produces blurred images (Schofield et al., 2020).

Subsequently, the filtered data is back-projected to reconstruct the 3D volume. Back projection involves the accumulation of information from multiple views, where each voxel in the final reconstruction is influenced by the corresponding pixels in all projection images. This process is repeated for each view angle. The FDK algorithm is known for its computational efficiency and has become a standard method for reconstructing high-quality images in cone-beam CT. Its application is vital in various fields providing detailed and accurate representations of anatomical structures for diagnostic and treatment planning purposes. In this research, we employed the reconstruction library provided by See Through s.r.l. which is implemented based on the FDK algorithm.

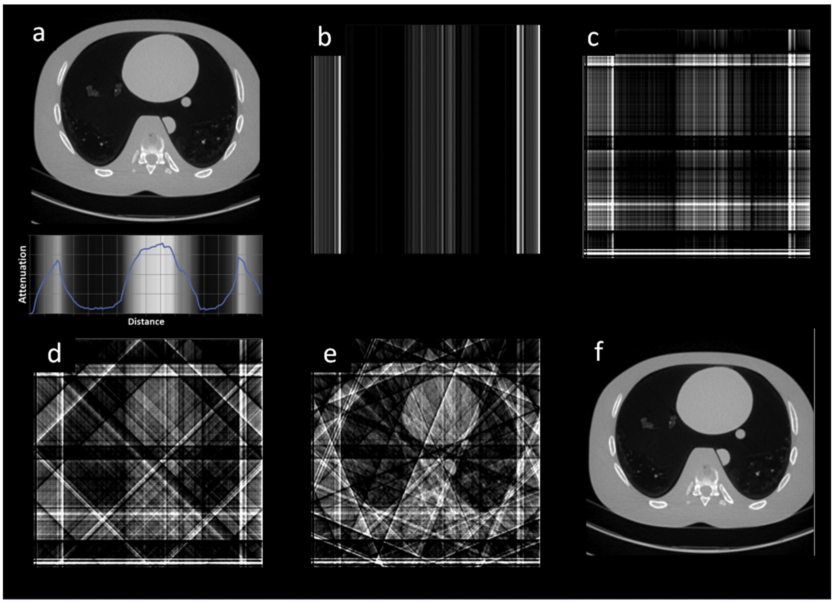


Figure 3.7: Filtered back projection (a) an anthropomorphic chest phantom. (b) back projection from a single projection image, (c) from two perpendicular projection images, (d) from four projection images, (e) from eight projection images, and (f) from 720 projection images. The simple back projection process inherently produces blurred images (Schofield et al., 2020).

3.4 Reference Systems and Projection Matrices

The device is positioned in such a way that we can see the patient's neck, as shown in the figure below. Initially, the X-ray source is situated to the left of the patient, with the detector positioned to the right. Throughout the acquisition process, the C-Arm rotates counterclockwise around the patient, as indicated by the arrow, along its axis of rotation (represented by the dashed line in the image).

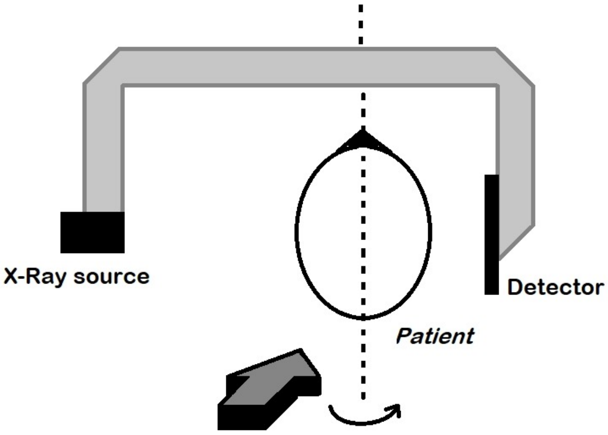


Figure 3.8: Positioning of the Patient's head

3.4.1 Preliminary Definitions

The following are the terminologies used to define the positioning of the patient in the device (Figure 3.9).

- **Axis of rotation:** This is the line around which the overhead gantry rotates during X-ray image acquisition.

- **Focal Point:** This is the point where X-rays are emitted.

- **Sensor:** It is the device that collects the X-rays emitted by the focus rays and allows the generation of X-ray images.
- **Principal point:** It is the orthogonal projection of the focus rays onto the plane that contains the sensor.
- **Principal axis:** This is the line that joins the focus rays to the main point.
- **Sensor Median:** This is the vertical line that passes through the center of the sensor.

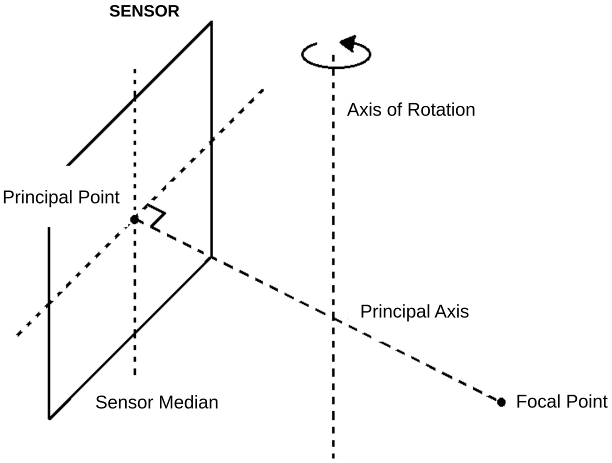


Figure 3.9: Terminologies used to define the patient positioning

The focal point and the sensor are connected to the overhead assembly, sharing their respective reference systems. As the overhead assembly moves, both components (and thus their reference systems) follow a circular path around the axis of rotation. To achieve optimal image quality, precise alignment of the focal point, sensor centerline, and axis of rotation is crucial. This entails ensuring that once the machine is aligned, the focal point coincides with a point on the sensor centerline (or as close as possible), and the sensor centerline forms a straight line parallel to the axis of rotation. Ultimately, the primary axis should intersect with the overhead's axis of rotation.

3.4.2 Definition of Reference Systems

Several different reference systems are involved in the process of acquiring and reconstructing a volume as described below.

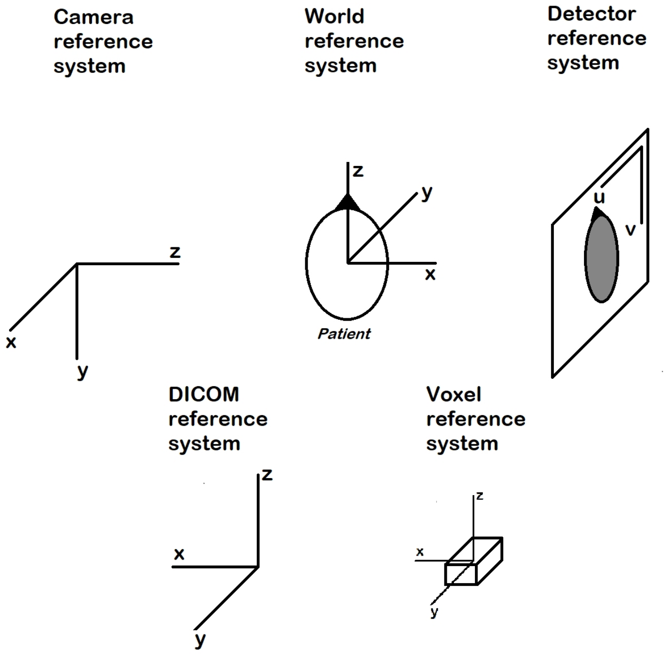


Figure 3.10: Different Reference Systems

1. **World Reference System**: The primary reference system for reconstructing and displaying the patient's volume. Its orientation is defined as:
    (a) The X-axis extends from the patient's left to right.
    (b) The Y-axis extends from the nape of the neck to the patient's face.
    (c) The Z-axis extends from bottom to top.

Originating from a point along the line parallel to the C-Arm's axis of rotation passing through the bite, this system sits approximately 1 cm above the lower part of the detector.

2. **Camera Reference System**: This system, akin to a camera, translates the 3D points of space captured by the X-ray source. Positioned behind the X-ray source to visualize the detector, its orientation is:

   (a) The X-axis runs from left to right.

   (b) The Y-axis proceeds from top to bottom.

   (c) The Z-axis runs from the left to the right of the patient.

   Originating from the point of X-ray emission.

3. **Detector Reference System**: Used to express the pixel coordinates of image points, this reference system is integrated with the detector and oriented as follows:

   (a) The U-axis (X) extends from left to right.

   (b) The V-axis (Y) extends from top to bottom.

   Originating from the pixel at the top left of the detector.

4. **DICOM Reference System**: Expresses an object's position relative to the patient, with orientation defined as:

   (a) The X-axis runs from the right to the left of the patient.

   (b) The Y-axis proceeds from the patient's face to the nape of the neck.

   (c) The Z-axis proceeds from bottom to top.

5. **Voxel Reference System**: Used to express each voxel's position within the area of interest, integral with the DICOM reference system.

3.4.3 Projection Matrices

The role of projection matrices in cone-beam computed tomography (CBCT) imaging is pivotal, as they facilitate the transformation of a three-dimensional object into a series of two-dimensional projection images. These matrices essentially encode the geometric information necessary for reconstructing the object's internal structure from its projections. In CBCT, a source emits X-rays that pass through the object of interest and are detected by an array of detectors. Each projection image represents the attenuation of X-rays along various paths through the object, resulting in a two-dimensional projection of the object's internal structure from a specific viewpoint. Projection matrices are responsible for mapping the coordinates of points within the object onto the two-dimensional detector array. They encompass parameters such as the position and orientation of the X-ray source relative to the object, as well as the geometry of the detector array. By mathematically describing the projection geometry, these matrices enable the conversion of spatial information from the object's three-dimensional volume to the two-dimensional space of the detector. The Projection matrices can be expressed as follows.

$$\mathbf{P} \;=\; \mathbf{K} \cdot [\mathbf{R}|\mathbf{t}] \tag{3.1}$$

where,

- $\mathbf{K} = \begin{pmatrix} f & 0 & u_0 & 0 \\ 0 & f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$, the intrinsic matrix.
    - f is the focal length.
    - u_0 and v_0 are the coordinates of the principal point on the detector.
- $[\mathbf{R}|\mathbf{t}]$ is the 4×4 roto-translation matrix describing the orientation of the x-ray source coordinate system with respect to the world coordinate system.

3.5 Motion Simulation

Motion-affected data can be acquired in two different ways. Some researchers used robot programming and external movement devices to make the phantom move realistically while scanning (Spin-Neto et al., 2018). Other research groups modeled the motion of the phantom with a digital simulation (Kim et al., 2016; Maur et al., 2019) which we adopted in this study. As explained in Section 3.4.3, CBCT scanners typically describe the geometry of the acquisition using a set of projection matrices (say $\mathbf{P}$) of the same type as the pinhole camera model (Hartley and Zisserman, 2004) that encode the relative positions of the x-ray source, detector, and scanned object for each of the acquired projections. These matrices are necessary for reconstructing the volume of the object using either the Filtered Back Projection (FBP) technique or iterative techniques.

$$\mathbf{P}_i \ = \ \mathbf{K} \cdot \mathbf{G}_i, \quad i = 1...N \tag{3.2}$$

Here, N is the number of projections, $\mathbf{K}$ is the intrinsic matrix describing detector parameters such as the orientation of the detector coordinates system with respect to the x-ray source (camera) coordinate system and the pixel size, $\mathbf{G}_i = [\mathbf{R}_i|\mathbf{t}_i]$ is the 4×4 roto-translation matrix for projection i. These projection matrices can also be conveniently used to simulate motion by altering their parameters. Since any jaw movement is usually prevented by a fixed bite block and a chin rest, we consider only rigid motions here. These motions can be described as a combination of rotations and translations. Let $\mathbf{G}'_i = [\mathbf{R}'_i|\mathbf{t}'_i]$ be the matrix describing the patient motion at scan position i. Since the patient's head is positioned on a chin rest, pure translations are very unlikely to occur. Therefore, we simulate motions by perturbing only rotation parameters. Subsequently, to replicate motion, we replace the projection matrix $\mathbf{P}_i$ with $\mathbf{P}'_i$.

$$\mathbf{P}'_i = \underbrace{\mathbf{K} \cdot \mathbf{G}_i}_{\mathbf{P}_i} \cdot \mathbf{G}'_i \tag{3.3}$$

Since our goal here is to simulate only rigid motions, this methodology enables the simulation of a more extensive range of motion patterns compared to a limited set of clinical scenarios. Also, clinical data would need to be accompanied by ground-truth motion measurement, which would require a complex setting, difficult to implement in a clinical environment. This would likely prevent the collection of large amounts of data. To simulate realistic motion, we rely upon the motion studies conducted by Spin-Neto et al. (2013, 2016, 2018, 2015). Various rigid motions can occur during cone-beam computed tomography (CBCT) scanning. In this study, we simulated a total of four distinct motion types: nodding, tilting, lateral rotation, and tremor. Tremor motion can be seen mostly in elderly people with Parkinson's disease. An illustration depicting these motions is presented in Figure 3.11.

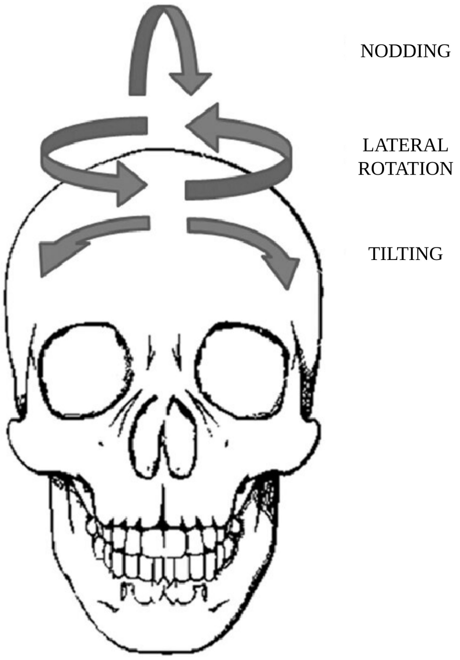
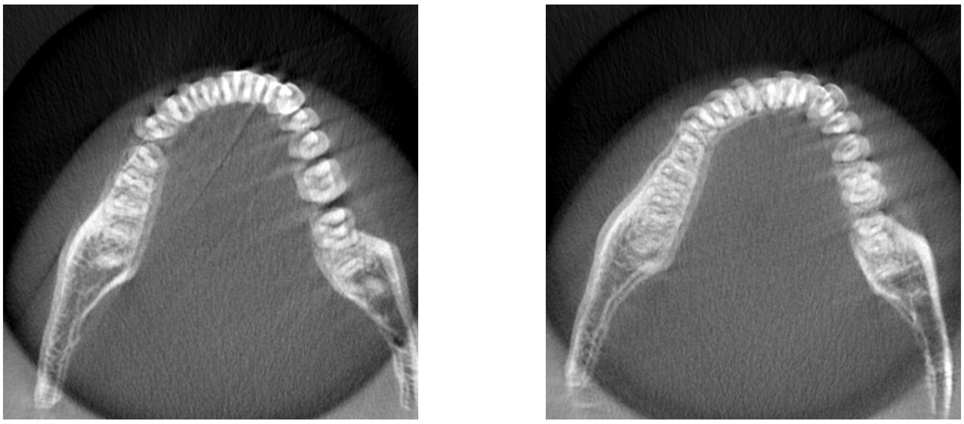


Figure 3.11: Three types of movement: nodding, tilting, and lateral rotation (Nardi et al., 2016).

Moreover, we took into account three distinct motion patterns: 1. the patient moves and remains in the final position for the remaining duration of the scan (**sf**), 2. the patient moves and returns to the initial position after a certain duration (**ri**), 3. the patient moves and returns to a specific state other than the initial position (**rd**). All these considerations provide us with the capability to simulate motion realistically, mirroring scenarios that can occur during real-life scanning. We validated our simulation technique by comparing it with actual motion-affected data obtained by physically moving the phantom during scanning (Figure 3.12).



(a) Real Motion (b) Simulated Motion

Figure 3.12: Comparison between artifacts from real and simulated motion for lateral rotation.

A sample pseudocode for simulating motion, considering three distinct motion patterns, is outlined in Algorithm 1. The algorithm is adaptable for all three motions by simply substituting the axis of rotation (‘**x**’ for nodding; ‘**y**’ for tilting; ‘**z**’ for lateral rotation). Figure 3.13 illustrates the patient coordinate system utilized in maxillofacial CBCT imaging.

```
Input   : Read projection matrices, $P(3, 4, nFrames)$
Get motion parameters: $motionDegree$, $motionDuration$, $motionStart$ (in seconds), $motionEnd$ (in
  seconds);
Set $motionPattern$ (1 - 'sf', 2 - 'ri', 3 - 'rd');
numMotionFrames = motionDuration × fps;
step = motionDegree/(numMotionFrames);
$\theta$ = [0, step, 2*step, 3*step,...(numMotionFrames)*step];
Initialize $j$ to $0$;
for $i$ from $0$ to $nFrames - 1$ do
    switch $motionPattern$ do
        case $1$ do
            if $i$ is within the range of $motionStart$ to $motionEnd$ then
                Calculate $R_i^{'}$ using $\theta[j]$;
                Create $G_i^{'}$ with $R_i^{'}$;
                $P_i^{'}$ = $P_i \cdot G_i^{'}$;
                Increment $j$;
            end
        end
        case $2$ do
            if $i$ is within the range of $motionStart$ to $motionStart + \frac{motionDuration}{2}$ then
                Calculate $R_i^{'}$ using $\theta[2j]$;
                Create $G_i^{'}$ with $R_i^{'}$;
                $P_i^{'}$ = $P_i \cdot G_i^{'}$;
                Increment $j$;
            end
            else if $i$ is within the range of $motionStart + \frac{motionDuration}{2}$ to $motionEnd - 1$ then
                Calculate $R_i^{'}$ using $\theta[2j]$;
                Create $G_i^{'}$ with $R_i^{'}$;
                $P_i^{'}$ = $P_i \cdot G_i^{'}$;
                Decrement $j$;
            end
        end
        case $3$ do
            if $i$ is within the range of $motionStart$ to $\frac{motionEnd}{2}$ then
                Calculate $R_i^{'}$ using $\theta[2j]$;
                Create $G_i^{'}$ with $R_i^{'}$;
                $P_i^{'}$ = $P_i \cdot G_i^{'}$;
                Increment $j$;
            end
            else if $i$ is within the range of $\frac{motionEnd}{2}$ to $motionEnd$ then
                Calculate $R_i^{'}$ using $\theta[j]$;
                Create $G_i^{'}$ with $R_i^{'}$;
                $P_i^{'}$ = $P_i \cdot G_i^{'}$;
                Decrement $j$;
            end
        end
    end
end
Output: Updated projection matrices
```

Algorithm 1: Realistic Motion Simulation

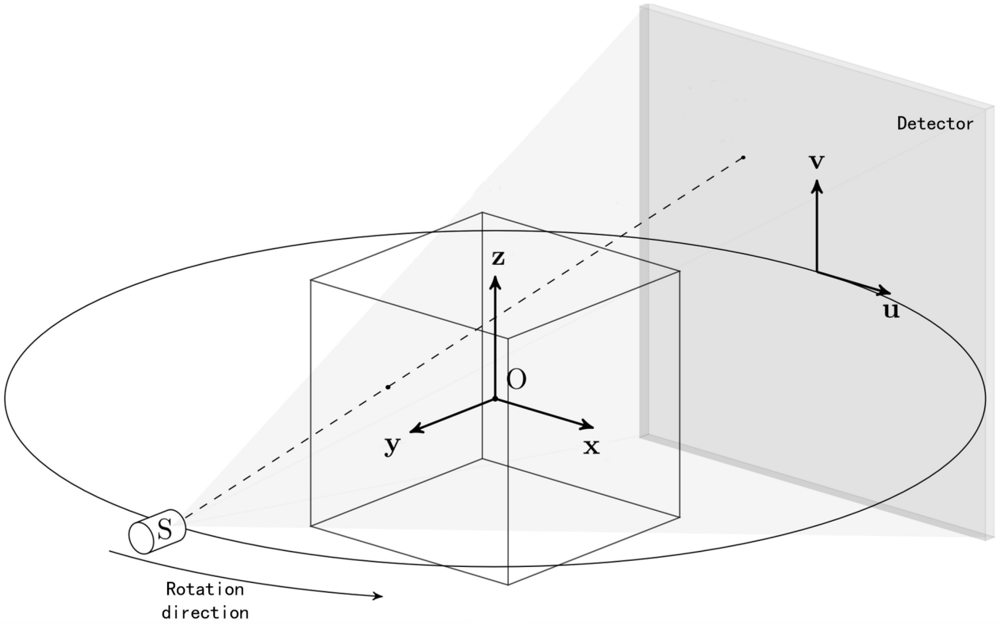


Figure 3.13: Maxillofacial CBCT coordinate system. **x,y,z** - 3D patient coordinates; **u,v** - 2D detector (image plane) coordinates; S - X-ray source. (Zhang et al., 2023)

Chapter 4

Non-iterative Motion Compensation

4.1 Introduction

Motion compensation in maxillofacial CBCT imaging uses a set of techniques and strategies employed to mitigate or correct the impact of the patient motion that occurred during the scanning process. Existing solutions for motion compensation often involve complex algorithms that require multiple reconstructions of the complete set of projections, making the process time-consuming (Maur et al., 2019; Sun et al., 2021; Birklein et al., 2023). Another prevalent limitation is that these algorithms can compensate only for short-duration motions. Additionally, these approaches may suffer from inconsistencies resulting from data truncation and scatter, common issues encountered in CBCT imaging (Würfl et al., 2019). In this chapter, we proposed a novel method to address these drawbacks. The presented approach is a non-iterative compensation method employing the Feldkamp-Davis-Kress (FDK) algorithm (Feldkamp et al., 1984) for reference reconstruction. Our method aims to effectively address substantial and prolonged rigid motions, which notably deteriorate the quality of reconstructions. Importantly, this is achieved without necessitating multiple full-scan reconstructions during the compensation stage.

4.1.1 Objectives and Constraints

The main research objective focused in this chapter is to build a fast, simple yet robust algorithm to compensate for rigid patient motions that significantly degrade the reconstruction quality. In this regard, we identify two major goals,

- Minimize the number of reconstructions in the compensation stage.
- Compensate for large and long-duration rigid motions.

4.2 Outline of the Chapter

Section 4.3 describes the proposed methodology. The implementation details are described in Section 4.4, and the results of the experiments are discussed in Section 4.5. Finally, a short conclusion with a few insights has been provided in Section 4.6.

4.3 Methodology

The proposed method for motion-compensated CBCT reconstruction is a non-iterative approach that eliminates the need for repeated reconstructions during the compensation stage. It relies on the assumption that among all projections a subset of motion-free projections exists large enough to reconstruct a motion-free short-scan volume. Our experiments indicated that projections corresponding to half of the scan duration (i.e. $180°$) are adequate for reconstructing a good-quality reference volume (refer to Section 4.4.2). This relies on experimental studies of Spin-Neto et al. (2018) showing that motion affecting the reconstruction quality is confined to a smaller region. The pipeline is illustrated in Figure 4.1. It consists of two main steps: the first is to detect and extract a motion-free subset of projections, and the next is to use these projections to create a short-scan reference reconstruction that can be used to compensate for artifacts in the motion-affected projections.

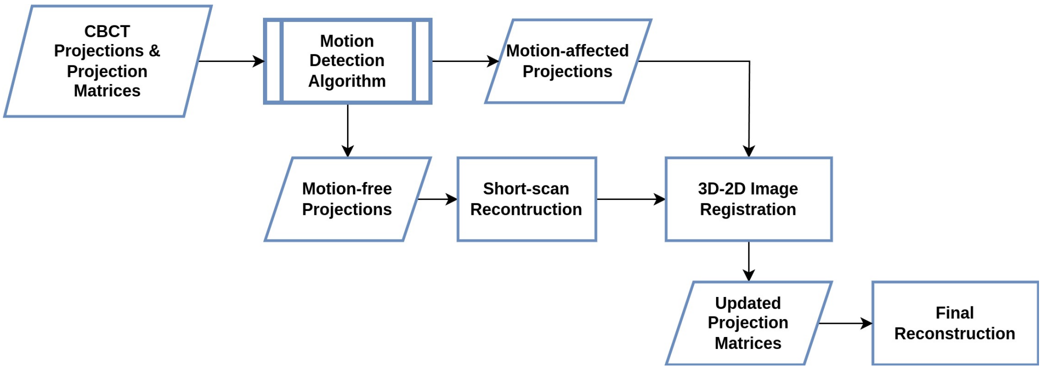


Figure 4.1: Pipeline of the Proposed Method

Our approach to detect motion artifacts in the reconstructed volumes is discussed in Chapter 5. This approach not only detects motion artifacts but also extracts a motion-free subset of projection that can be used for reference reconstruction. The focus of the current chapter is on the second step (motion compensation).

4.3.1 Motion Compensation

The process begins with an initial short-scan reference reconstruction generated from a motion-free subset of the acquired projections as explained in Section 4.3.2. Each motion-affected projection is then registered to a synthetic projection, which is generated by reprojecting the reference reconstruction. The compensation algorithm is depicted in Figure 4.2.

The algorithm performs registration of the original projections with the synthetic forward projections by optimizing the similarity score, which is obtained using the cost function explained in Section 4.3.3. The optimization process adjusts the motion parameters until convergence is achieved, essentially retrieving the perturbation matrix $\mathbf{G}_i^{'}$ mentioned in Equation 3.3. A significant advantage of this method is that the same reference reconstruction is used for all projections, allowing the optimizer to run in parallel for multiple projections. A final high-quality reconstruction is performed using the estimated motion parameters.

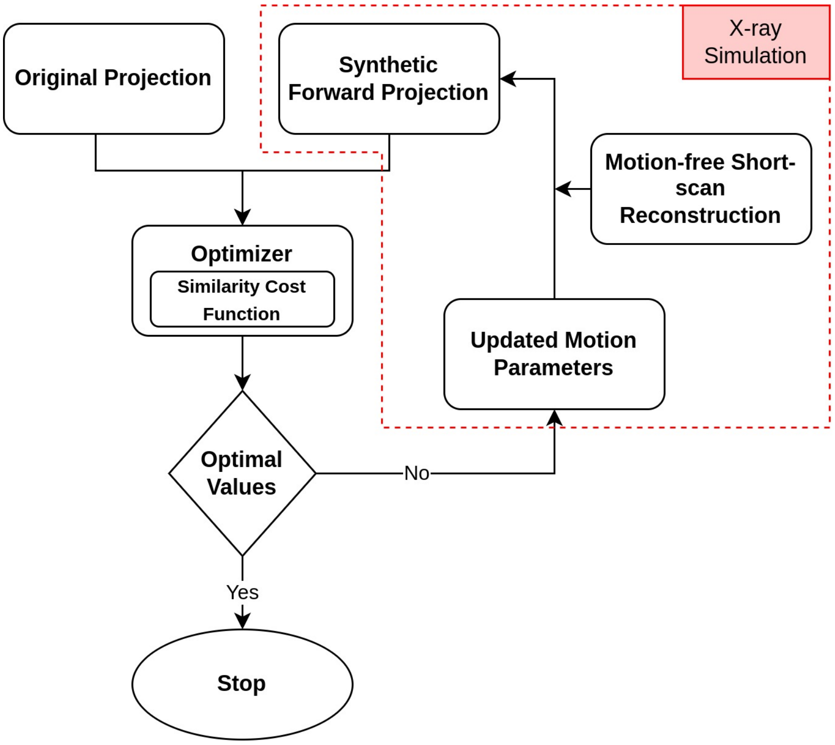
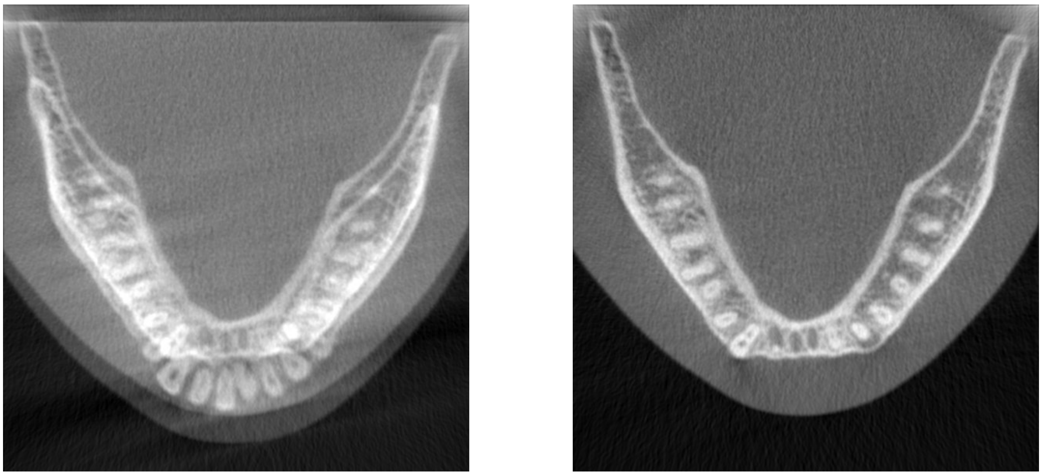


Figure 4.2: Motion Compensation Algorithm

4.3.2 Reference Reconstruction

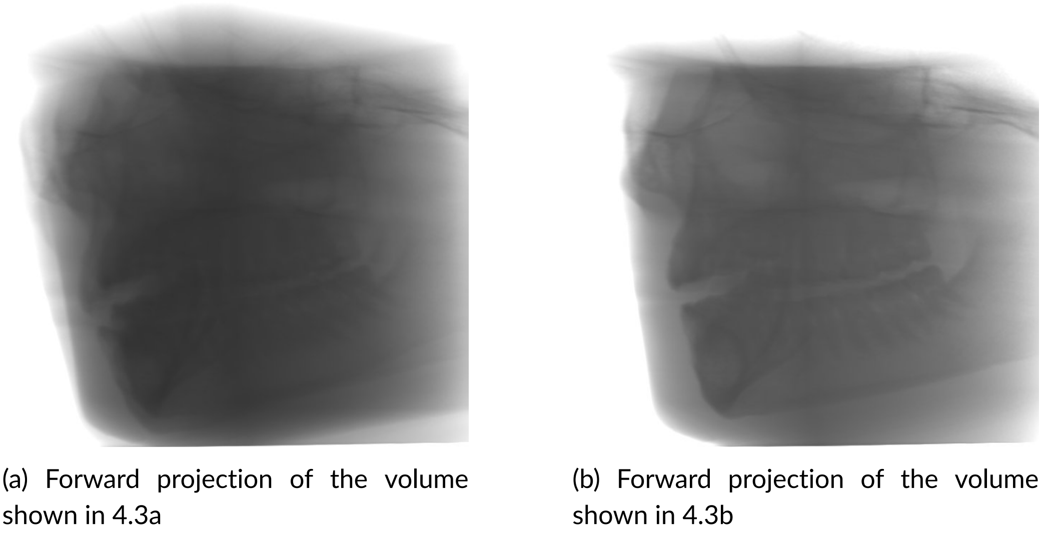
By utilizing the motion-free sub-set of the acquired original projections, the FDK algorithm (Feldkamp et al., 1984) generates a short-scan reference reconstruction (Figure 4.3b). A comparison of the full-scan initial reconstruction and the short-scan reconstruction from a motion-free subset of projections is shown in Figure 4.3. The advantage of having a short-scan reference reconstruction is that it is free from motion artifacts, hence the synthetic projections generated by forward projecting this reconstruction are consistent with the acquisition geometry and ensure robust registration with the original motion-affected projections.



(a) 360° full-scan Reconstruction with motion artifacts

(b) Short-scan reconstruction from a motion-free subset

Figure 4.3: Reference reconstruction.



(a) Forward projection of the volume shown in 4.3a

(b) Forward projection of the volume shown in 4.3b

Figure 4.4: Comparison between synthetic forward projections from full-scan reconstruction and motion-free short-scan reconstruction.

4.3.3 Similarity Cost Function

Once we have a clean reference reconstruction we can then create synthetic projections by forward projecting the reference reconstruction using the default projection matrices that correspond to each motion-affected projection. We can then use a gradient-based approach to perform the 3D-2D image registration. To do this, we define a cost function based on Gradient Information. The gradient of an image represents how the pixel intensities change in different directions. It is commonly used in image processing to identify edges and other important features. In our case, the cost function $G(r, a)$ is computed as the sum of the minima of gradient magnitudes computed for corresponding pixels in the synthetic image (r), and the actual image (a) as expressed in the equation below,

$$G(r,a) = \sum_{(\mathbf{x},\mathbf{x}') \in (r \,\cap\, a)} w_{\mathbf{x},\mathbf{x}'} \; \min\left(|\nabla \mathbf{x}|, |\nabla \mathbf{x}'|\right) \tag{4.1}$$

Where,

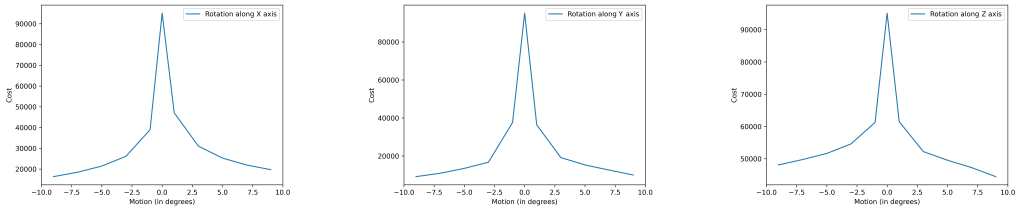
- $\mathbf{x}$ and $\mathbf{x}'$ refer to the corresponding sample points (pixels) in r and a, respectively.
- $|\nabla \mathbf{x}|$ and $|\nabla \mathbf{x}'|$ are the gradient magnitudes of $\mathbf{x}$ and $\mathbf{x}'$, respectively.

The minima operator is applied to exclude strong extraneous gradients (De Silva et al., 2016). This helps to ensure that the alignment considers regions where the gradients are similar.

The weighting function w in Equation 4.1 is the gradient orientation which is computed as the cosine of the angle between each corresponding gradient of r and a expressed as follows:

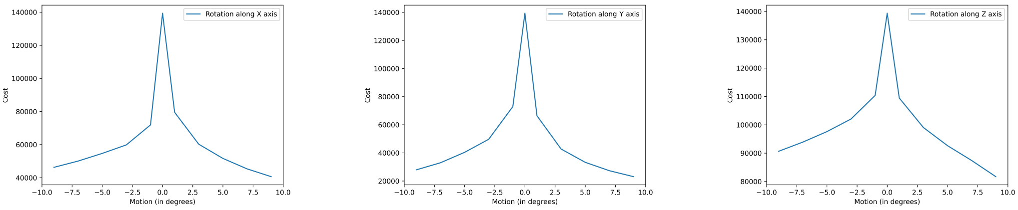
$$w_{\mathbf{x},\mathbf{x}'} = \cos(\theta) = \frac{\nabla \mathbf{x} \cdot \nabla \mathbf{x}'}{|\nabla \mathbf{x}|\,|\nabla \mathbf{x}'|}. \tag{4.2}$$

Figures 4.5, 4.6, 4.7, and 4.8 illustrate how the cost function behaves when registering the reference projection with the motion-affected projections across a range of rotations (-9° to 9°) in different views.



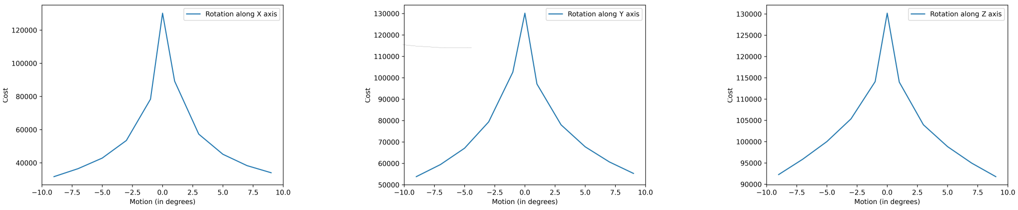
(a) Rotation along X axis (b) Rotation along Y axis (c) Rotation along Z axis

Figure 4.5: Cost Function for matching reference projection and motion-induced forward projections (anterior view)



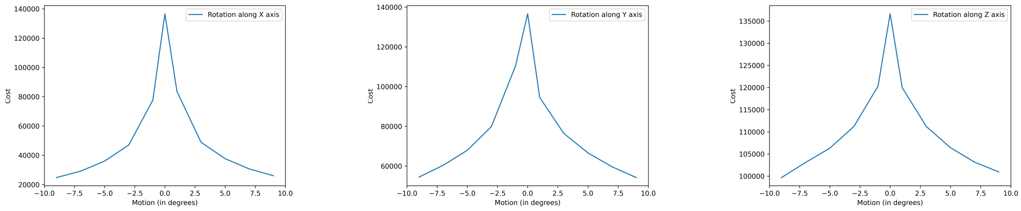
(a) Rotation along X axis (b) Rotation along Y axis (c) Rotation along Z axis

Figure 4.6: Cost Function for matching reference projection and motion-induced forward projections (posterior view)



(a) Rotation along X axis (b) Rotation along Y axis (c) Rotation along Z axis

Figure 4.7: Cost Function for matching reference projection and motion-induced forward projections (lateral view - left)



(a) Rotation along X axis (b) Rotation along Y axis (c) Rotation along Z axis

Figure 4.8: Cost Function for matching reference projection and motion-induced forward projections (lateral view - right)

As depicted in the figures above, the cost function attains its peak value at zero-degree motion for all motion types. This indicates that the cost function reaches its highest point when the reference projection aligns effectively with the synthetic projection.

4.4 Implementation

The experiments were conducted on a Windows-10 PC with an Intel(R) Core(TM) i7-2.90 GHz processor, 16GB RAM, and NVIDIA RTX 3070 (8GB VRAM). The volume reconstruction and forward projection libraries used in this study were provided by See Through S.r.l., Italy. Powell's conjugate direction optimizer from the SciPy Library is used for optimizing the 3D-2D registration.

4.4.1 Evaluation Metrics

The evaluation of the resulting reconstruction's quality was done using Root-Mean-Square Error (RMSE) and Structural Similarity Index Measure (SSIM) (Wang et al., 2004). To assess the quality of the final reconstruction, a motion-free reconstruction of the 360° scan was performed as the ground truth (Figures 4.10c, 4.11c, 4.12c, 4.13c).

4.4.1.1 Root Mean Square Error

Root Mean Square Error (RMSE) is a commonly used metric in image processing to quantify the difference or error between two images. It measures how different the pixel values of two images are on average, which is useful for tasks like image compression, denoising, and image reconstruction. RMSE is often used when you have a pair of images: a reference image and a processed (or reconstructed) image, and you want to assess how accurately the processed image reproduces the reference image. The RMSE is calculated using the following formula:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \tag{4.3}$$

Where:

- N is the total number of pixels (or voxels) in the images (assuming they have the same dimensions).
- x_i is the pixel value of the reference image.
- y_i is the corresponding pixel value of the processed (or reconstructed) image.

The RMSE value provides a measure of the average difference between corresponding pixels in the two images. A lower RMSE indicates that the two images are more similar, while a higher RMSE suggests greater dissimilarity. It's worth noting that RMSE is sensitive to outliers and can be influenced by a single pixel with a large error, so it should be used in conjunction with other metrics when assessing image quality.

4.4.1.2 Structural Similarity Index Measure

The Structural Similarity Index (SSIM) is a widely used metric in image processing and computer vision to quantify the similarity between two images. It's designed to capture both structural information (like edges, textures, and patterns) and luminance information (brightness and contrast) to provide a more comprehensive assessment of image similarity compared to simple pixel-wise comparison methods like Mean Squared Error (MSE).

The SSIM index is calculated using the following formula:

$$\text{SSIM}(x, y) = [L(x, y) \cdot C(x, y) \cdot S(x, y)]^{\alpha} \tag{4.4}$$

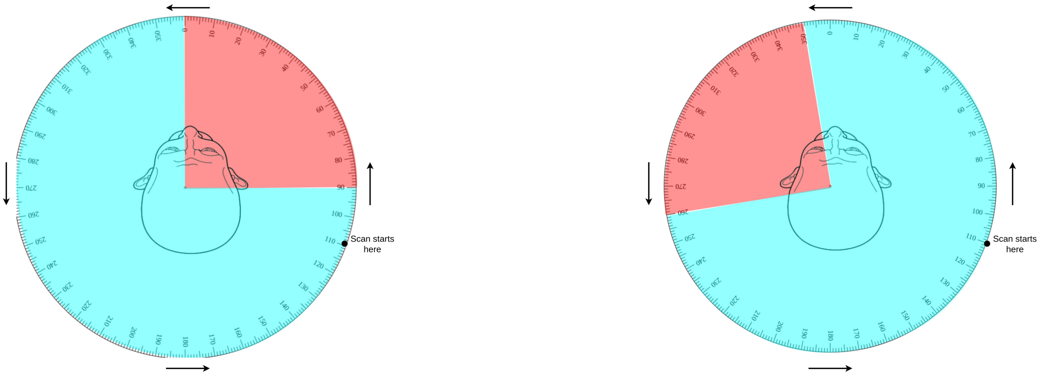
Where:

- x and y are the two input images being compared.
- $L(x, y)$ represents the luminance comparison, measuring the similarity in brightness and contrast.
- $C(x, y)$ is the contrast comparison, measuring the similarity in contrast.
- $S(x, y)$ is the structure comparison, measuring the similarity in the structural content.
- α is a parameter to control the relative importance of luminance, contrast, and structure comparisons. Typically, it's set to a value close to 1.

The SSIM index ranges from -1 to 1, where 1 indicates that the two images are identical in luminance, contrast, and structure, while -1 means they are completely dissimilar. Higher SSIM values generally indicate greater similarity between the images. We utilize a cubic sliding window with dimensions of $11 \times 11 \times 11$ voxels to calculate the average SSIM index over the reconstructed volume. SSIM is a useful metric for tasks like image quality assessment, image compression evaluation, and image restoration, as it provides a more perceptually meaningful measure of image similarity compared to simpler metrics like MSE or PSNR.

4.4.2 Data

Realistic motions were simulated using the strategy described in the section 3.5. Three returning motions (**nodding**, **tilting** and **lateral rotation**) (Table 4.1) were simulated of 6 seconds (one quarter) for a full $360°$ scan of 24 seconds. Because of the mechanical fixations, e.g. head support, it is realistic to assume that the patient returns to the original position after the movement, so that it is possible to find at least one motion-free subset of $180°$, which is sufficient for a good reference reconstruction. An illustration showing the motion-affected (marked in red) and motion-free (marked in blue) region of scanned projections for the data used in this experiment is presented in Figure 4.9.



(a) Nodding (b) Tilting

(c) Lateral Rotation (d) Abrupt motion

Figure 4.9: Illustration showing the motion-affected subset (red) and motion-free subset (blue) of projections

To test the robustness of the method, a **non-returning abrupt motion** was also simulated, where the patient remains in one pose for part (almost half) of the scan, and then suddenly moves to another pose and remains still for the rest of the scan. This created double contours in the reconstruction (Figure 4.13), which is one of the most difficult cases to compensate for.

4.5 Experimental Results

4.5.1 Quantitative Results

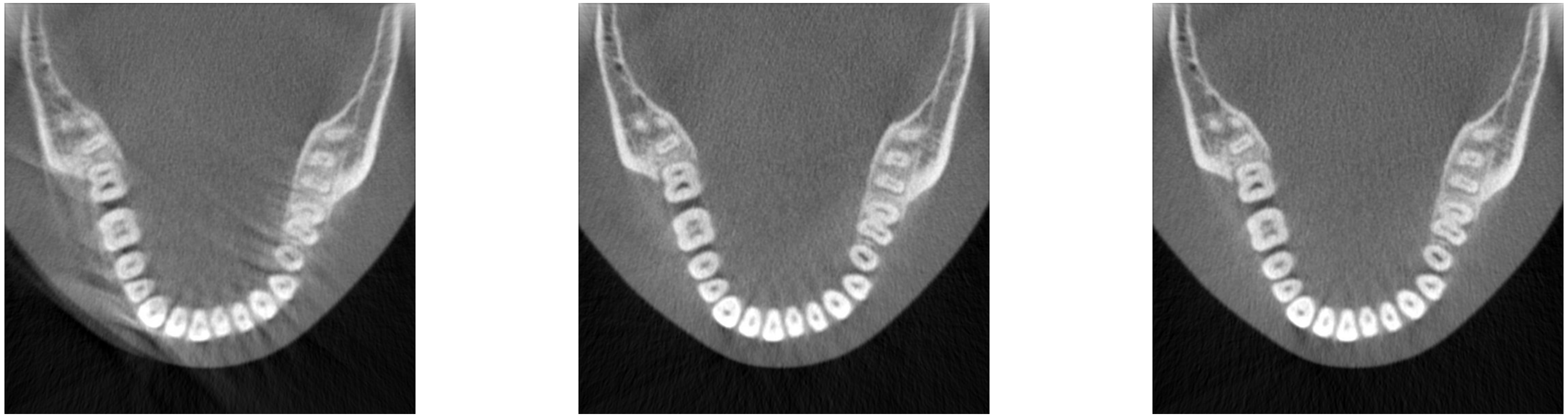
Table 4.1 provides a quantitative comparison of motion-affected and compensated reconstructions for the three types of motions as evaluated by the SSIM and RMSE metrics. For nodding, tilting, and abrupt motions, there was a significant improvement in SSIM value and a considerable decrease in RMSE value for motion compensated reconstruction. For lateral rotation, there was only a small improvement in SSIM value and a moderate decrease in RMSE value.

Table 4.1: SSIM and RMSE values for different motions [MA = motion-affected, MC = motion-compensated]

	SSIM (↑)			RMSE ($\times 10^{-2}$)(↓)		
Motion type	MA	MC	Diff %	MA	MC	Diff %
Nodding	0.876	0.948	+7.23	2.38	1.25	-1.13
Tilting	0.857	0.947	+9.05	3.00	1.47	-1.53
Lateral rotation	0.850	0.878	+2.74	3.16	2.23	-0.93
Abrupt motion	0.697	0.925	+22.80	8.40	1.43	-6.97

4.5.2 Qualitative Results

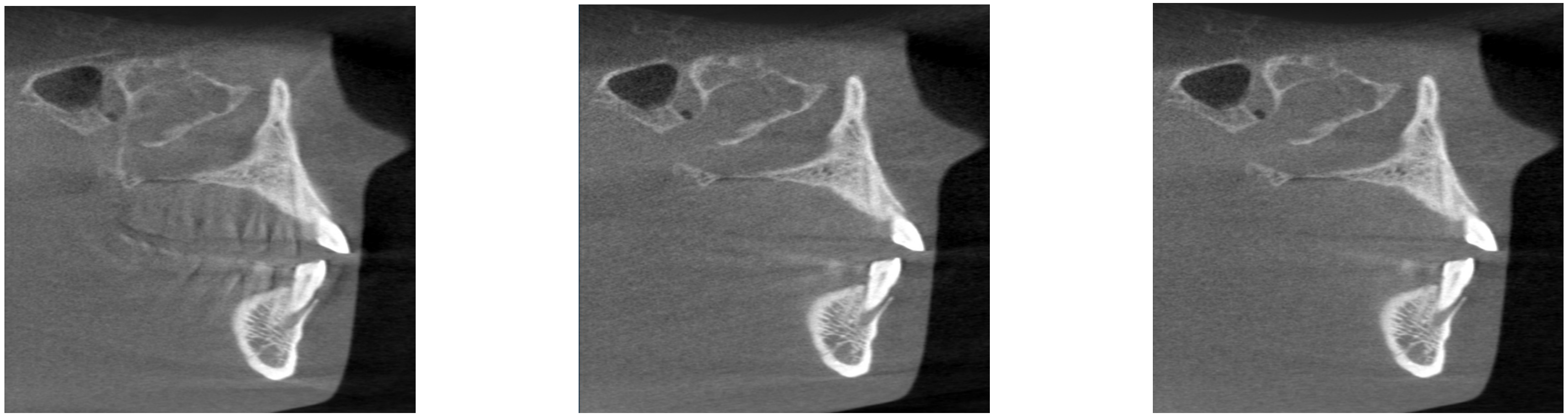
The qualitative comparison between the ground truth, motion-affected, and motion-compensated reconstructions are shown in Figures 4.10, 4.11, 4.12, and 4.13. The proposed method was effective in compensating for motion-induced artifacts in nodding, tilting, and non-returning abrupt motion cases, as almost all artifacts were successfully compensated. In the case of lateral rotation, some minor artifacts were still present, but the crucial diagnostic structures such as teeth were unaffected.



(a) Motion-affected

(b) Compensated

(c) Ground truth



(d) Motion-affected

(e) Compensated

(f) Ground truth

Figure 4.10: Results for nodding motion

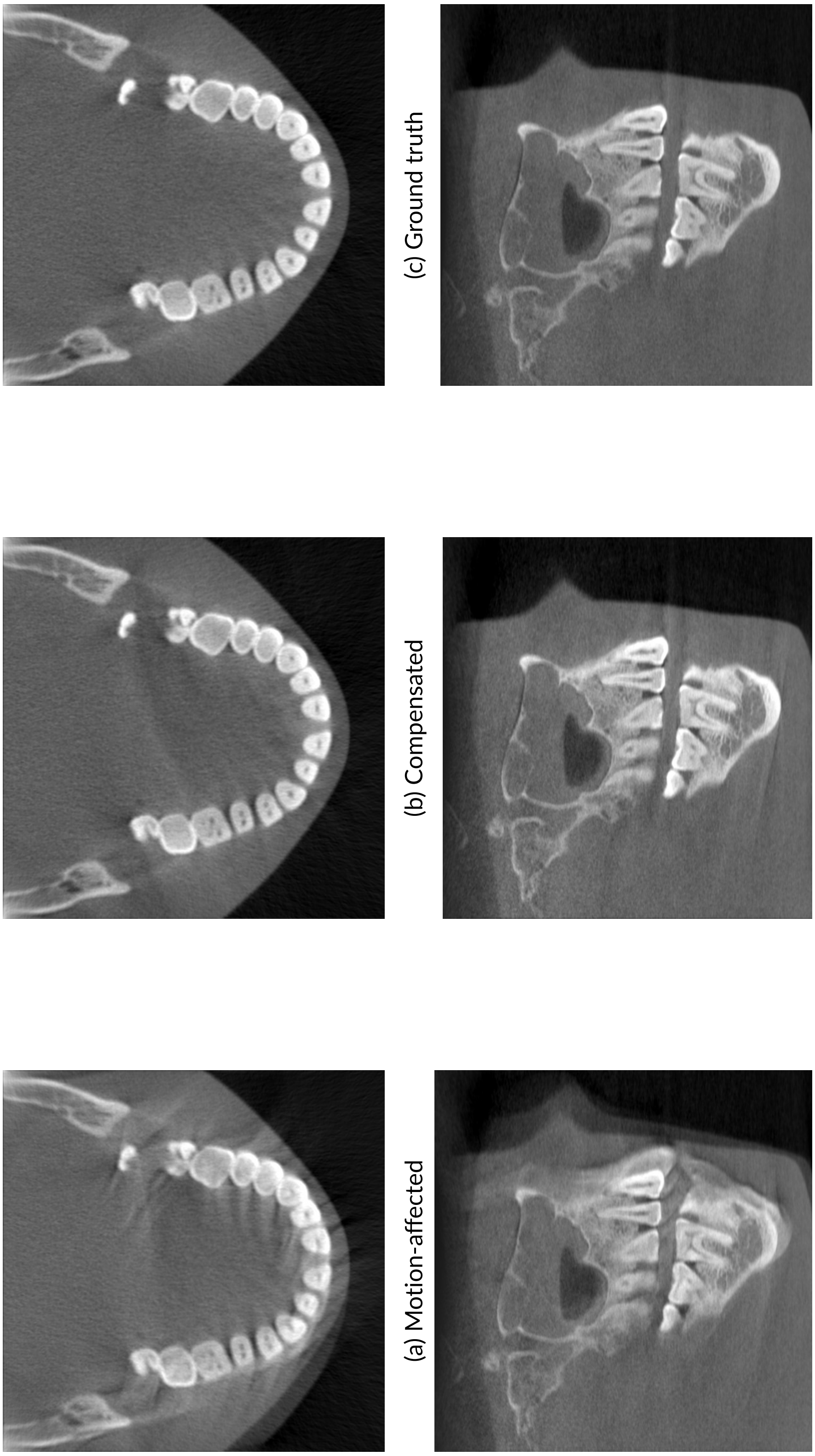
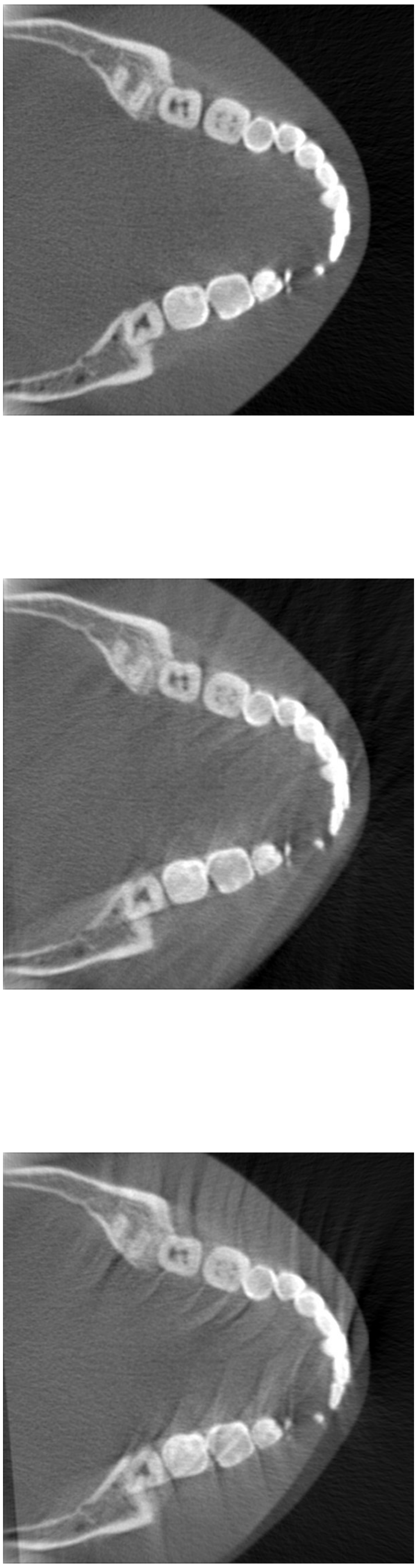

(a) Motion-affected
(b) Compensated
(c) Ground truth
(d) Motion-affected
(e) Compensated
(f) Ground truth

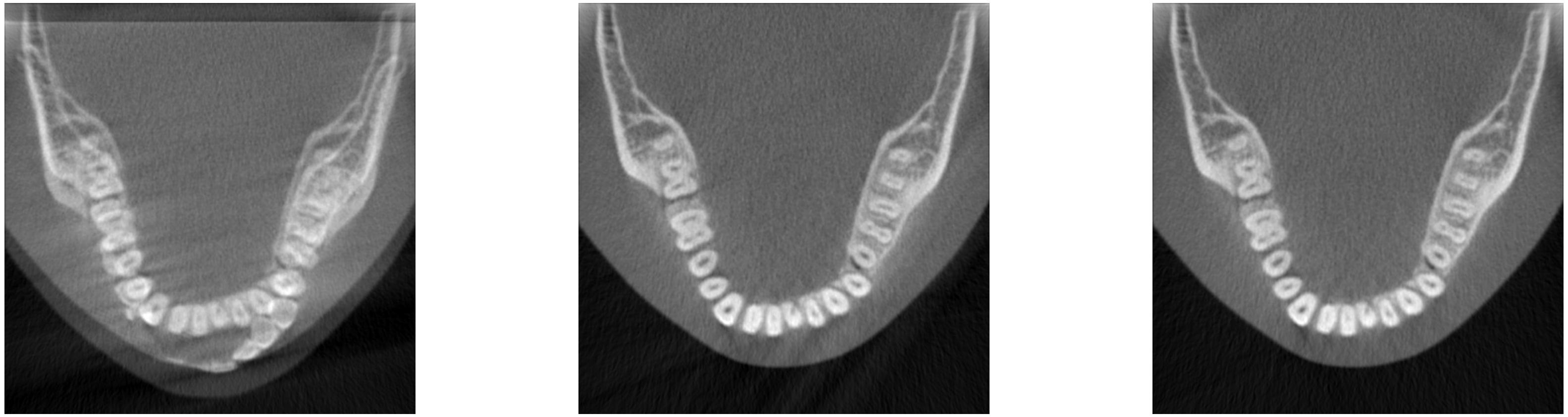
Figure 4.11: Results for tilting motion



(a) Motion-affected (b) Compensated (c) Ground truth

(d) Motion-affected (e) Compensated (f) Ground truth

Figure 4.12: Results for lateral rotation



(a) Motion-affected (b) Compensated (c) Ground-truth

(d) Motion-affected (e) Compensated (f) Ground-truth

Figure 4.13: Results for non-returning abrupt motion

4.5.3 Discussion on Complex Cases

The occurrence of multiple movements in one full-scan is quite common in children (Spin-Neto et al., 2016), which can make it less likely to find a set of projections that is motion-free when scanning at a 180° angle. To simulate this situation, we introduced multiple motions during a full 360° scan, specifically, a non-returning nodding motion lasting 6 seconds and a returning tilting motion lasting 2 seconds as illustrated in Figure 5.12. These motions were simulated in a way that there is no possibility of obtaining a 180° motion-free subset of projections. From this data, we extracted a subset that exhibited fewer motion artifacts and used it for reference volume reconstruction. Since this subset still contained some motion artifacts, we performed 3D-2D registration for all the projections. The result is shown in Figure 4.15. This finding suggests that the algorithm is capable of effectively mitigating motion, even in the case of more complex motion types, leaving only minor artifacts. To further enhance reconstruction quality, we implemented a regularization technique in the reference reconstruction, which is detailed in Chapter 6.

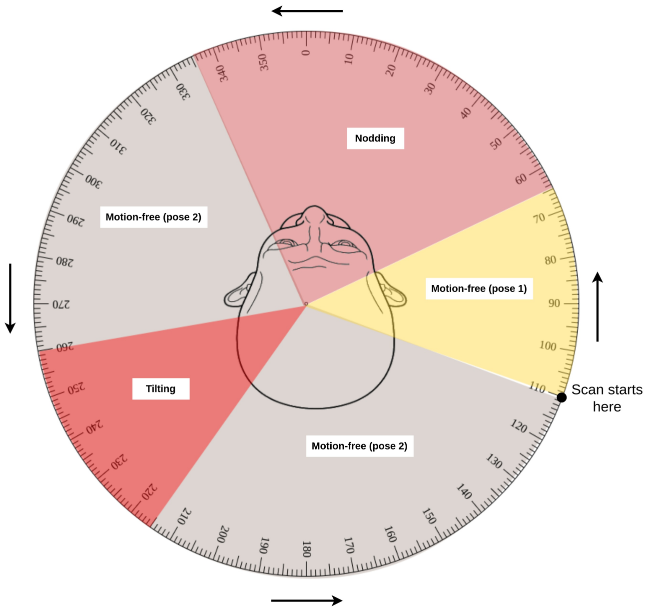
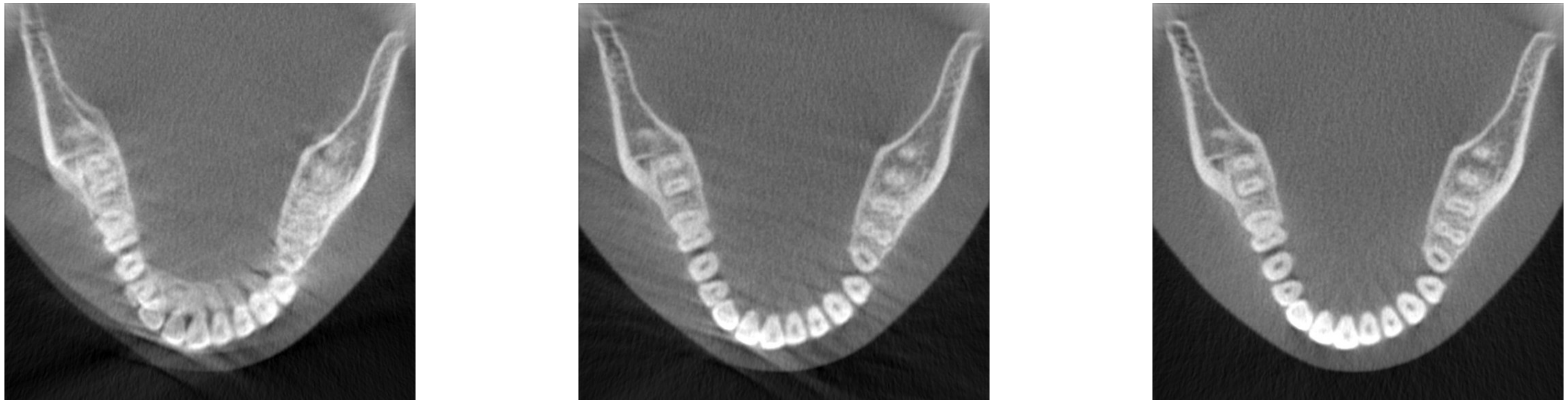


Figure 4.14: Illustration showing multiple motion



(a) Motion-affected (b) Compensated (c) Ground-truth

(d) Motion-affected (e) Compensated (f) Ground-truth

Figure 4.15: Results for multiple motions (nodding and tilting)

4.6 Conclusion

Our approach utilizes a non-iterative technique to effectively compensate for artifacts caused by rigid patient motions. By incorporating a motion detection strategy, our method can provide high-quality reconstruction without the need for multiple full-scan reconstructions during the compensation stage for most motion types. However, for scans containing multiple patient movements, additional constraints need to be applied to the proposed method to enhance the quality of reconstructions. Further testing involving clinical data is necessary to assess the reliability of the proposed method across various scenarios.

Chapter 5

Motion Artifacts Detection

5.1 Introduction

In Chapter 4, we demonstrated the substantial improvement achievable through a motion detection strategy in enhancing the robustness of the compensation algorithm. The primary focus of the current chapter is on extracting a subset of motion-free projections essential for reconstructing a clean short-scan reference volume. To achieve this objective, we generate four short-scan volumes from selected projections, covering the entire 360° view. Employing a deep learning-based framework, we classify these volumes into positive (containing motion artifacts) or negative (free from motion artifacts). Deep learning encompasses powerful algorithms capable of automatically learning and recognizing complex patterns within large datasets. Specifically, in the context of CBCT scans, these algorithms can be trained to identify motion artifacts by capturing unique patterns such as double contours and streaks, distinguishing them from artifacts induced by other factors like metal-related distortions.

5.2 Objective

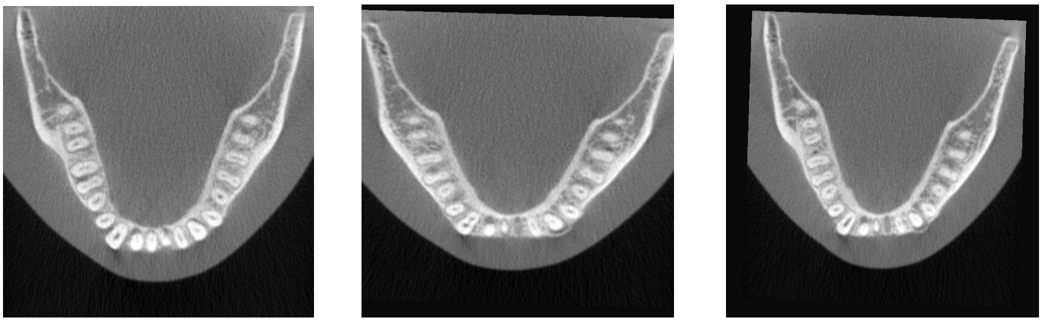
The focus of this chapter is to extract a part of the scanned CBCT projections, free from motion artifacts, to reconstruct a short-scan volume that can be used for reconstructing a reference volume as explained in 4.

5.3 Outline of the Chapter

The shortage of data is addressed in Section 5.4 explaining the augmentation and motion simulation strategy performed in this study. Section 5.5 describes the proposed methodology of motion detection and the implementation details are described in Section 5.6. The detailed experimental results are discussed in Section 5.7. Finally, a short conclusion with a few insights is provided in Section 5.8.

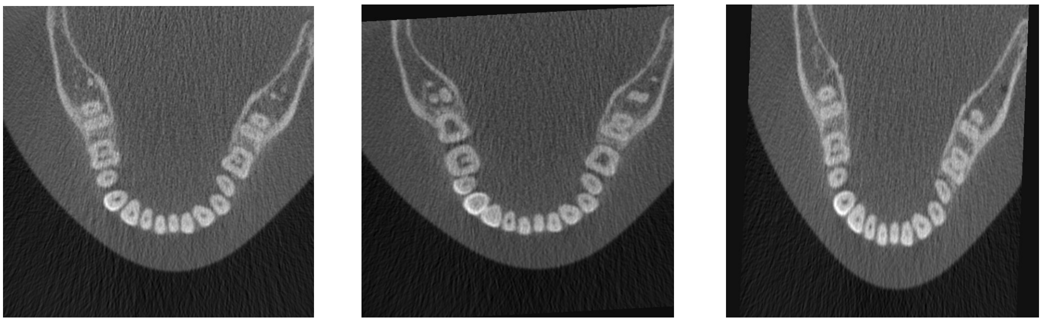
5.4 Data Augmentation

The issue of data imbalance is a common problem in machine learning-based medical imaging approaches. In the case of motion detection, there is often a lack of positive data (with motion artifacts) and comparatively more negative data (without motion artifacts) (Sun et al., 2020; Welch et al., 2020). In contrast, we have more positive data as the motion is artificially simulated but the negative data is reconstructed from only three head phantoms. To address this imbalance issue, we used data augmentation to increase the diversity of the negative data set. In most people, the dental arch (mandible and maxilla) is similar in shape with a slight variation in the size (Braun et al., 1998; Kairalla et al., 2014). Thus we can generate more virtual phantoms than the three available ones by scaling with a factor between 0.8 and 1.2 to each of the three phantoms, and by forward projecting these virtually augmented phantoms, we can generate more negative cases of motion-free scans. Additionally, random affine transforms with rotations up to 5° and translations up to $5\,mm$ along X, Y, and Z axes are also applied. This helps to create a more balanced dataset to improve the performance of the model. Few samples of augmented volumes are shown in Figures 5.1, 5.2, 5.3.



(a) Original Phantom (b) Virtually Augmented Phantoms

Figure 5.1: Augmentation of Tom phantom



(a) Original Phantom (b) Virtually Augmented Phantoms

Figure 5.2: Augmentation of $Jerry$ phantom

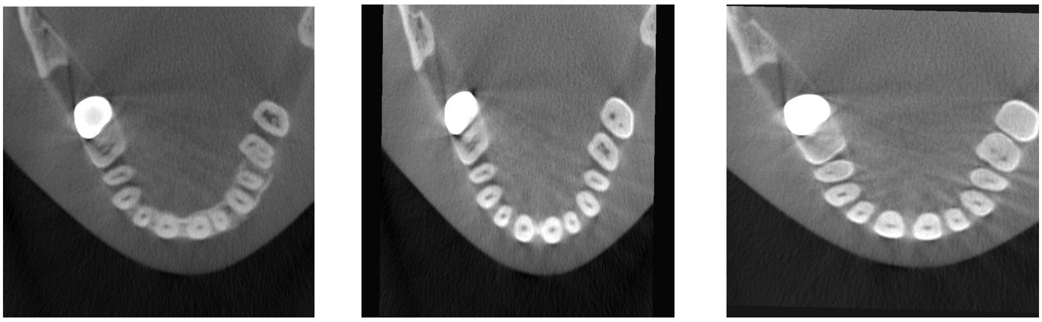

(a) Original Phantom (b) Virtually Augmented Phantoms

Figure 5.3: Augmentation of $Silvestro$ phantom

5.4.1 Motion Simulation

Realistic motions were simulated using the strategy described in the section 3.5. We simulated four different motion types: nodding (**Nod**), tilting (**Tilt**), lateral rotation (**LR**), and tremor (**Trem**). Additionally, three different motion patterns were considered; 1. the patient moves and stays in the final position for the rest of the scan (**sf**), 2. the patient moves and returns back to the initial position after some time (**ri**), 3. the patient moves and returns back to a particular state other than the initial position (**rd**). Spin-Neto et al. (2016, 2018) demonstrated that larger motions affecting image quality are confined to a smaller region. Therefore, the motion is estimated to be localized roughly in a range of 1-5 seconds for a full-scan of 20 seconds. In this study, the motion is simulated in different parts of the scanned projections with varying amplitude between $3\ to\ 8\ degrees$ for rotation and $3\ to\ 8\ mm$ for translation.

5.5 Methodology

A deep learning-based approach is used in this study to detect motion artifacts in CBCT short-scan volumes. The motion detection framework consists of three steps

- Slice generation, Network classification, and Volume averaging as shown in Figure 5.4.

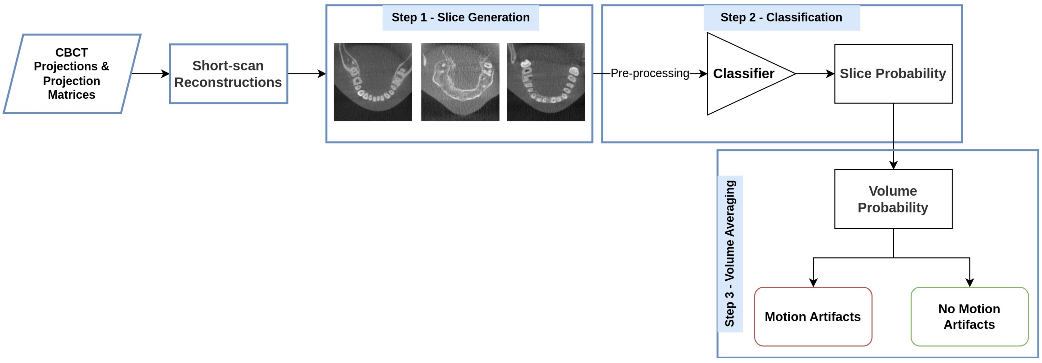


Figure 5.4: Framework of motion detection in CBCT volumes

5.5.1 Slice Generation

The variability of features in the human skull bones is limited and statistically described well in the literature (Braun et al., 1998; Kairalla et al., 2014). Human skull bones and teeth are locally very similar, although they result in faces that look different to human observers. For instance, the sizes and shapes of teeth vary by very small amounts (Leung et al., 2018). Therefore teeth and jaw bones can be good features to train a generalized motion detection framework. We train the network to see a limited space of anatomical features, and the motion-affected images have distinctive features, different from motion-free anatomies. We apply this principle in a simplified setting since the teeth and jaw bones can be well seen in 2D axial slices extracted from the volume. Artifacts are often visible in most axial slices, making them a reliable choice for classification. In this study, each short-scan volume is represented by 300 axial-view slices for classification. Using 3D data in the classification process can make the network architecture more complex and computationally expensive. A simpler network architecture, however, is often sufficient for classifying 2D data.

5.5.2 Network Classification

Different CNN models have been used and their performance on our data has been analyzed. The models are constructed based on transfer learning using pre-trained ImageNet weights. The network architectures include VGG with Batch Normalization (Simonyan and Zisserman, 2014), DenseNet (Huang et al., 2017), ResNet (He et al., 2016), EfficientNet (Tan and Le, 2019), and EfficientNet V2 (Tan and Le, 2021) . In the training phase, Binary Cross Entropy Loss is used. All the networks are trained with the Stochastic Gradient Descent (SGD) optimizer with a learning rate of 1e-3 and momentum of 0.9.

5.5.2.1 Pre-processing

As the networks have been pre-trained on natural images with pixel values ranging from 0 to 255, the pixel values within the slices are normalized to conform to the input specifications of the networks. The input images are also resized to 224×224 to match the networks' specifications.

5.5.3 Volume Averaging

In some slices where the direction of motion is parallel to the direction of strong edges, the artifacts due to motion may be less noticeable or compromised. This is because the strong edges can provide some structural information that can help to compensate for the motion-induced artifacts. To overcome this issue, volume averaging is performed. In the evaluation phase, the final prediction ($\mathbf{y}_{final}$) is based on the mean outcome ($\mathbf{y}_{pred}$) of the network per volume, i.e., the probability scores ($\mathbf{y}$) for all the slices generated from a volume are averaged (Eq. 5.1), and the average score is considered for the final prediction (Eq. 5.2). The advantage of volume averaging is that, e.g., even if only 65% of the slices are correctly predicted with an average probability of 0.8, the volume will still be classified correctly.

$$\mathbf{y}_{pred} = \frac{\sum_i \mathbf{y}_i}{N} \tag{5.1}$$

$$\mathbf{y}_{final} = \begin{cases} No\ Motion\ (-), & \text{if } \mathbf{y}_{pred} < 0.5 \\ Motion\ (+), & \text{if } \mathbf{y}_{pred} >= 0.5 \end{cases} \tag{5.2}$$

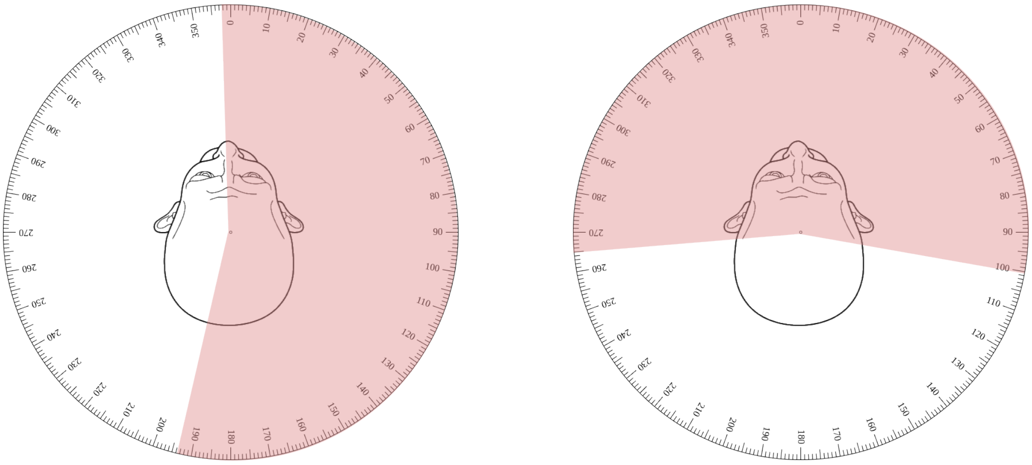
5.6 Implementation

The experiments were conducted with Python programming language (version 3.10) on a Windows-10 PC with Intel(R) Core(TM) i7-2.90 GHz processor, 16GB RAM, and NVIDIA GeForce RTX 3070 (8GB VRAM). The deep learning framework has been implemented using Pytorch (version 1.13) with CUDA (version 11.7). The pre-trained models used in this implementation are from the Torchvision library (version 0.14). Evaluation metrics are from the Scikit-learn library (version 1.1.3).

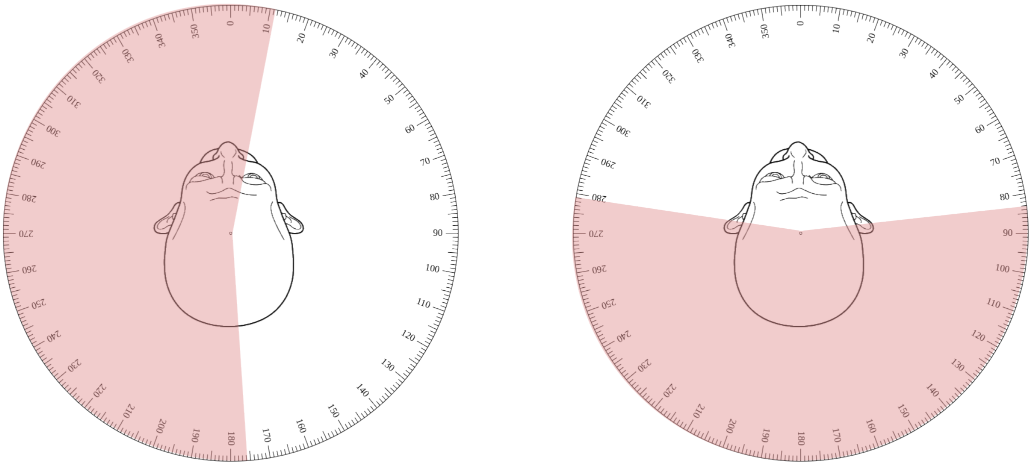
5.6.1 Dataset

The dataset consists of short-scan volumes reconstructed with projection images corresponding to a $194°$ angle. These short-scan volumes are derived from the full $360°$ scan of the head phantoms (also including the augmented ones). We considered 4 default short-scan views to cover the full $360°$ scan as illustrated in Figure 5.5. Since we applied a short-duration motion of 1-5 seconds for a scan of 24 seconds, there should always be at least one short-scan that is free from motion. The training dataset has 80 short-scan volumes, 40 volumes each from Tom and $Jerry$ phantoms. The volumes with and without motion artifacts are distributed equally in the training dataset. $Silvestro$ phantom is used for testing in order to study the robustness of our framework to detect motion in the presence of metal artifacts.

The testing dataset consists of 200 short-scan volumes reconstructed from 50 sets of projection data corresponding to 10 motion types (Table 5.1).



(a) View-1 (b) View-2



(c) View-3 (d) View-4

Figure 5.5: Illustration of regions (indicated in red) considered to perform short-scan reconstructions.

5.6.2 Evaluation Metrics

The proposed method has been evaluated using the Precision-Recall (PR) curve. The Precision-Recall (PR) curve plots the precision values (see Eq. 5.4) on the y-axis against the recall values (see Eq. 5.3) on the x-axis. The ideal model performance has high precision as well as high recall values which means the AUC-PR value is close to 1.0.

$$Recall = \frac{TP}{TP + FN} \tag{5.3}$$

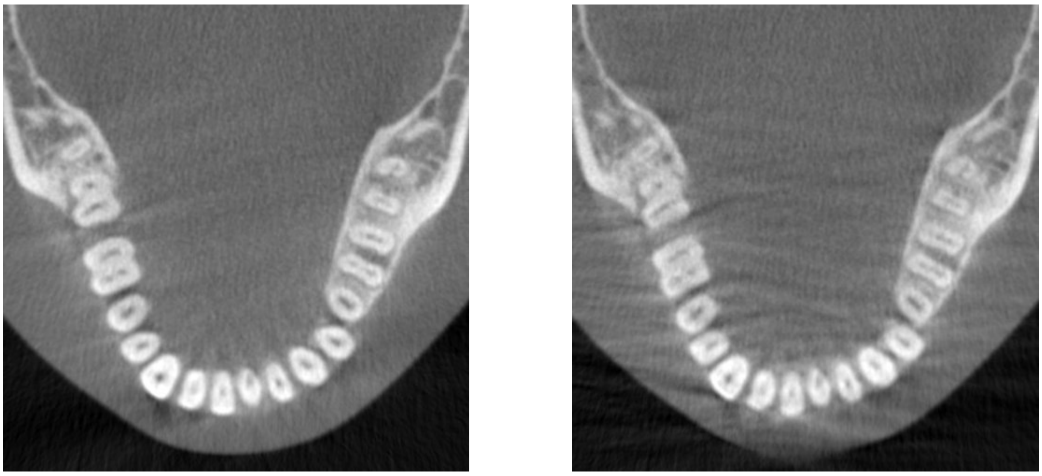
$$Precision = \frac{TP}{TP + FP} \tag{5.4}$$

5.7 Experimental Results

The performance has been evaluated using the area under the Precision-Recall curve (AUC-PR) metric for both slices and volumes in the testing dataset for different network architectures. Five sets of projection data are used per motion type and the short-scan reconstructions are labeled manually. Table 5.1 shows the average AUC-PR per motion type and overall average. ResNet performed well with respect to the number of slices being correctly predicted. However, the goal is to classify volumes, and in this case, EfficientNet V2 outperforms all other architectures. The results show that the framework is capable of specifically detecting motion artifacts in the presence of metal artifacts. Figures 5.6, 5.7, and 5.8 showcase few results. In these figures, a red caption indicates the network's detection of motion artifacts, while a green caption signifies the network's identification of the absence of motion artifacts.

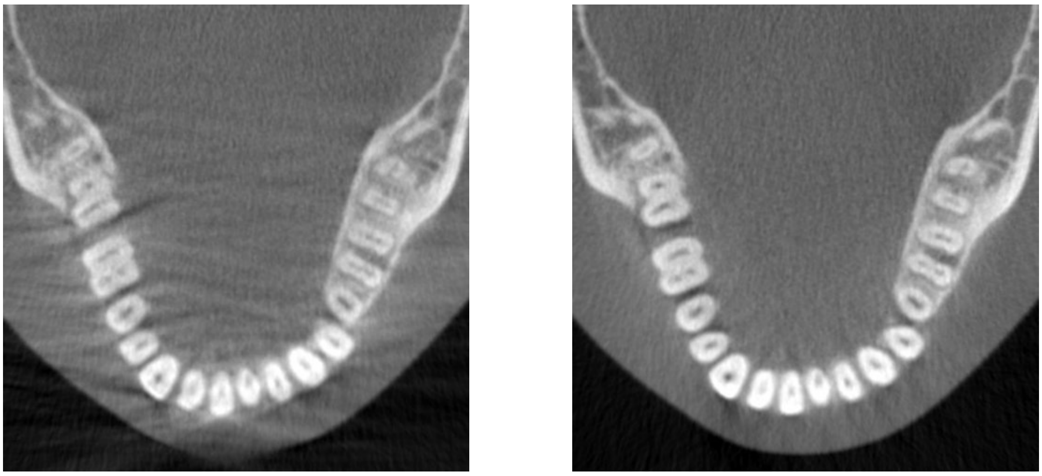
Table 5.1: AUC-PR for the testing ($Silvestro$) dataset on different network architectures. The best figures for each motion type are in bold.

Motion	**VGG**		**DenseNet**		**ResNet**		**EfficientNet**		**EfficientNet V2**	
Tpye	**Slices**	**Volume**	**Slices**	**Volume**	**Slices**	**Volume**	**Slices**	**Volume**	**Slices**	**Volume**
Nod-sf	**0.907**	**0.983**	0.885	0.962	0.904	**0.983**	0.888	0.929	0.873	0.906
Nod-ri	0.829	**1.000**	0.801	0.857	**0.844**	0.950	0.791	0.917	0.786	**1.000**
Nod-rd	0.842	0.857	0.830	0.833	**0.851**	0.894	0.826	0.885	0.839	**0.917**
Tilt-sf	0.864	0.962	0.846	0.900	**0.902**	**1.000**	0.862	0.900	0.859	**1.000**
Tilt-ri	0.851	0.896	0.805	0.837	**0.872**	**0.934**	0.826	0.896	0.820	0.896
Tilt-rd	0.903	**1.000**	0.865	0.900	**0.908**	0.983	0.855	0.929	0.874	**1.000**
LR-sf	0.837	0.962	0.836	0.900	**0.841**	0.883	0.839	0.900	0.815	**0.983**
LR-ri	0.845	**0.960**	0.816	0.866	**0.854**	0.919	0.826	0.866	0.806	0.919
LR-rd	0.800	0.884	0.768	0.821	**0.831**	**0.925**	0.778	0.884	0.768	**0.925**
Trem	0.884	**1.000**	0.833	0.857	**0.935**	**1.000**	0.854	0.857	0.855	**1.000**
Average	0.856	0.950	0.828	0.873	**0.874**	0.947	0.834	0.896	0.830	**0.955**



(a) View-1

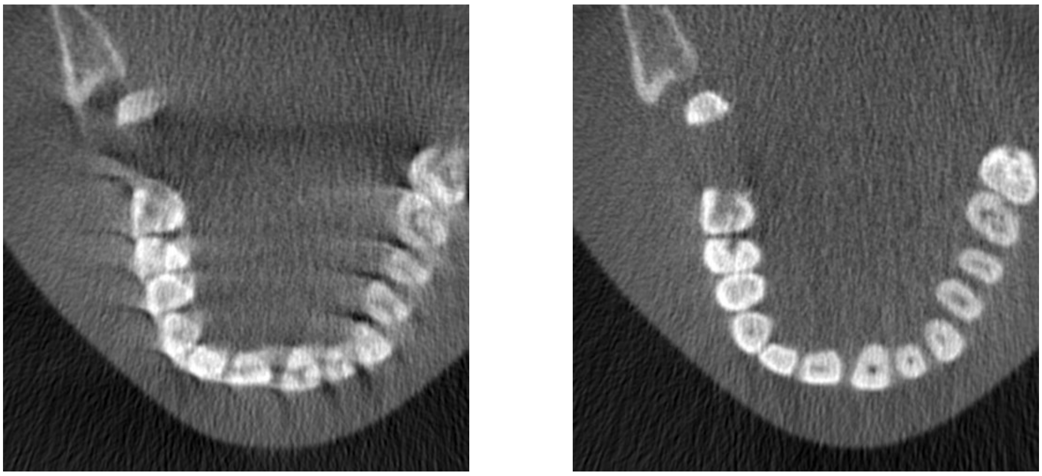
(b) View-2



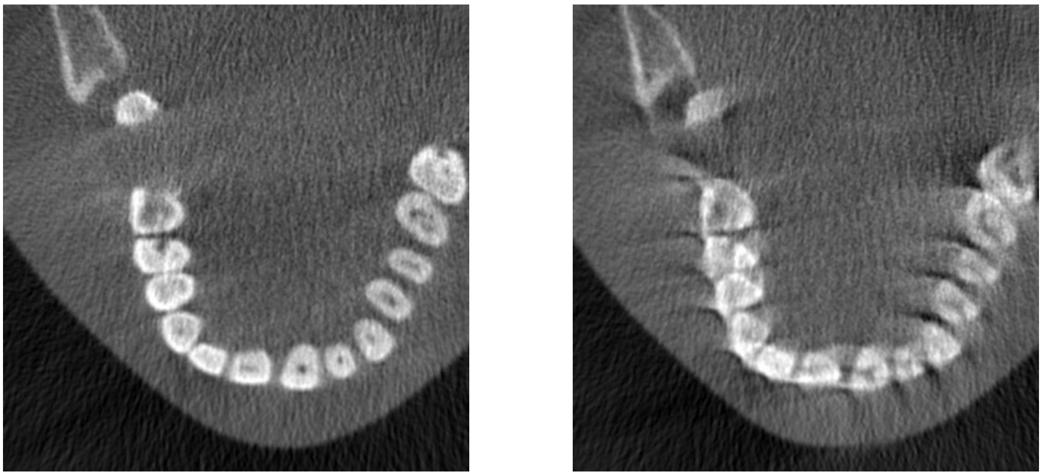
(c) View-3

(d) View-4

Figure 5.6: Short-scan regions for Tom data with Tremble.

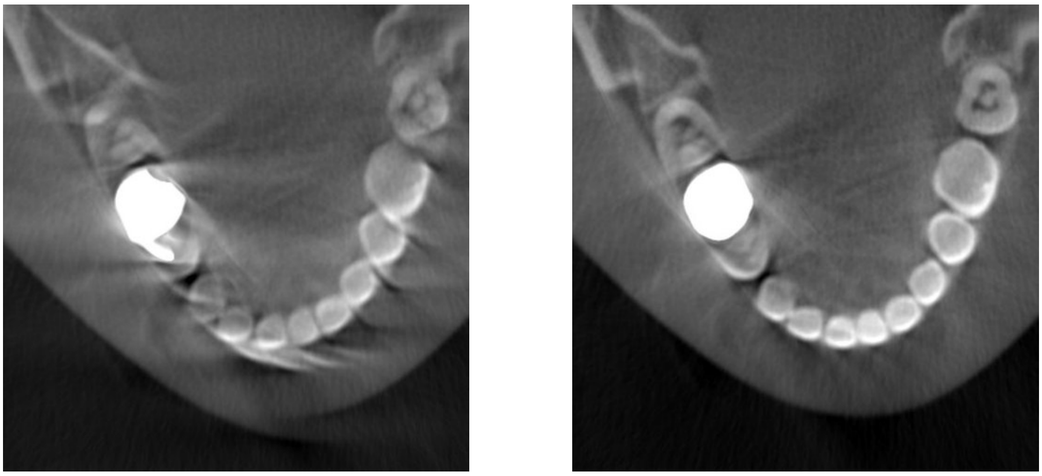


(a) View-1 (b) View-2

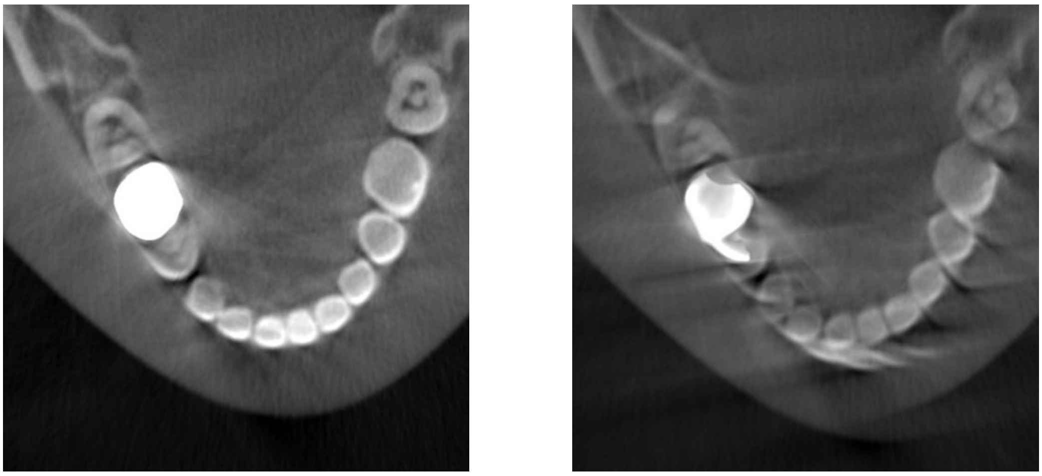


(c) View-3 (d) View-4

Figure 5.7: Short-scan regions for $Jerry$ data with Tilting.



(a) View-1 (b) View-2



(c) View-3 (d) View-4

Figure 5.8: Short-scan regions for $Silvestro$ data with Nodding. Metal artifacts are visible in all the reconstruction views.

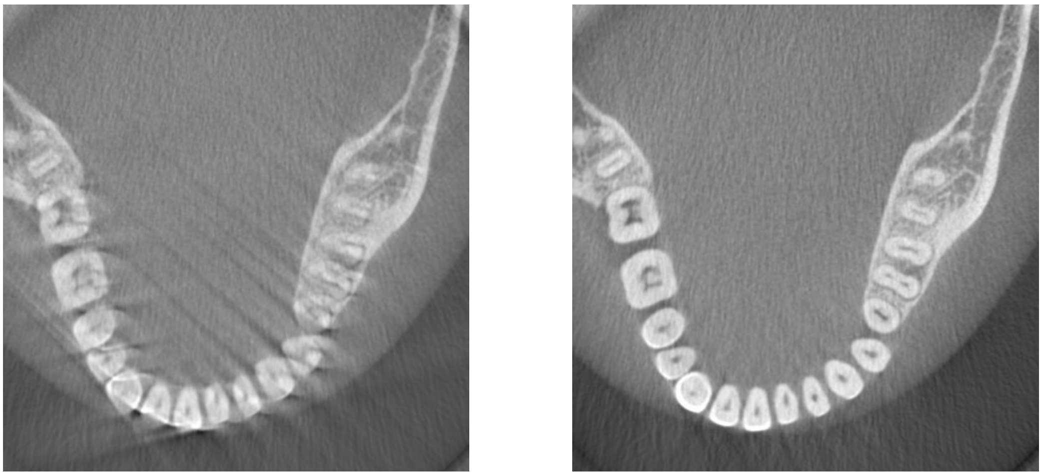
5.7.1 Validation on Real Data

The framework has also been tested in real motion-affected data to evaluate its robustness in real-world scenarios. The data was acquired by moving the phantom during scanning to induce motion. A setup was used to move the phantom during the scan, thereby inducing a motion that is approximately equivalent to lateral rotation. Three scans were performed with motion at different time intervals and short-scan volumes were reconstructed for the 4 selected views. These volumes were then classified using different architectures and the results are presented in Table 5.2. EfficientNet, EfficientNet V2, and VGG were able to cope-up with differences between real and simulated motions better than ResNet and DenseNet (see Figures 5.9, 5.10, 5.11).

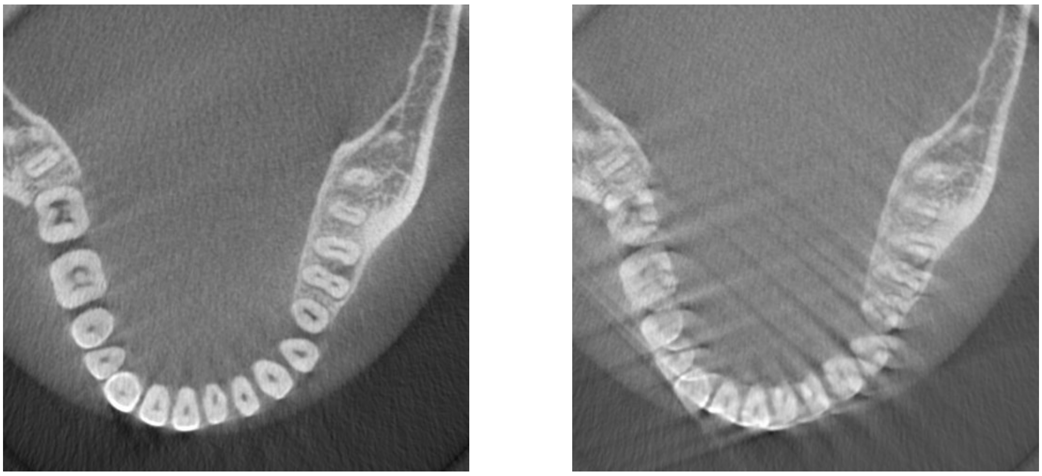
Table 5.2: AUC-PR for real motion-affected data on different network architectures. The best figures are in bold.

Scan No.	**VGG**		**DenseNet**		**ResNet**		**EfficientNet**		**EfficientNet V2**	
Tpye	**Slices**	**Volume**	**Slices**	**Volume**	**Slices**	**Volume**	**Slices**	**Volume**	**Slices**	**Volume**
Scan-1	0.934	**1.000**	0.920	0.875	0.918	0.875	0.937	**1.000**	**0.970**	**1.000**
Scan-2	0.931	**1.000**	0.894	0.875	0.907	0.875	0.933	**1.000**	**0.941**	**1.000**
Scan-3	0.855	**1.000**	0.801	0.750	0.802	0.750	0.855	**1.000**	**0.896**	**1.000**
Average	0.906	**1.000**	0.872	0.833	0.876	0.833	0.908	**1.000**	**0.936**	**1.000**

Simulated motion data is different from real-world scans in two ways. Firstly, the reconstruction is performed from simulated x-ray projections in which the noise due to scattering, photon starvation, and beam hardening are significantly less compared to acquired projections. Secondly, motion is modeled and not acquired from a real-world setup. Overall, the results show that the network is capable of inferring motion artifacts from real-world scans successfully, even though only simulated motion-affected data is used in training.

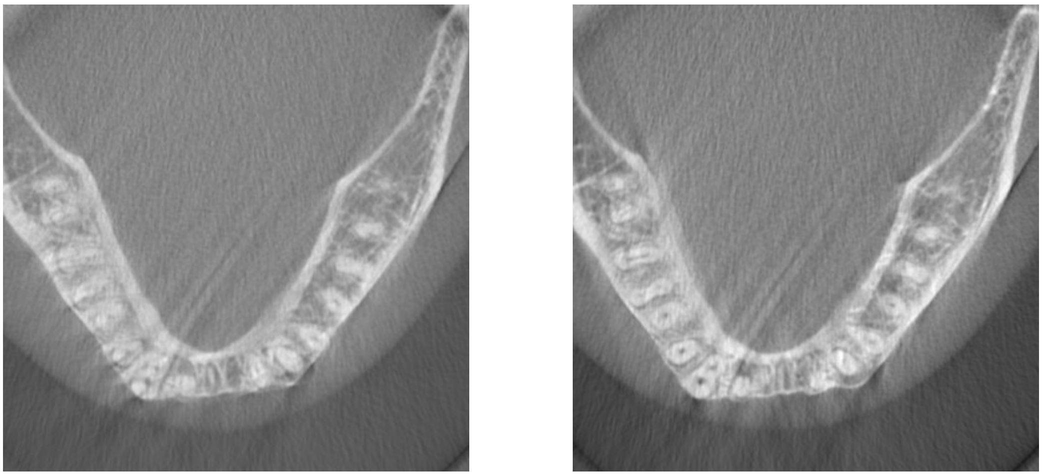


(a) View-1 (b) View-2

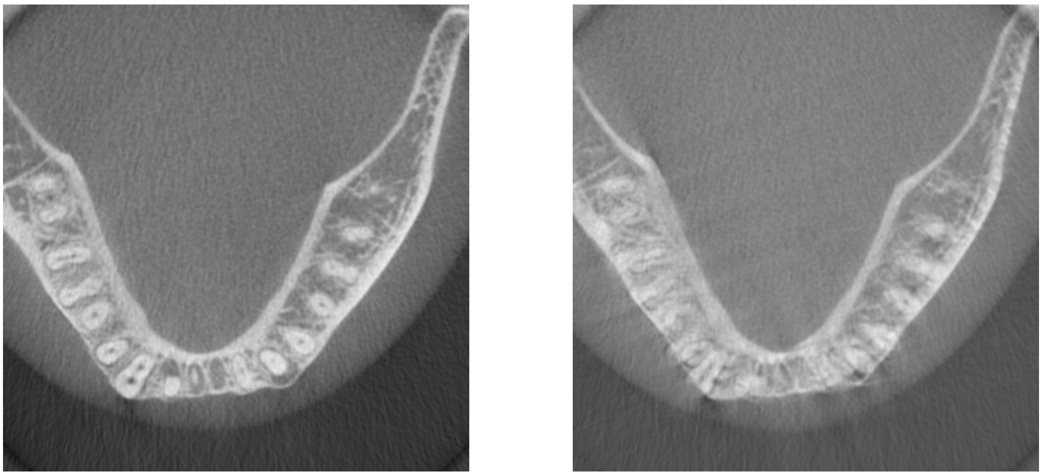


(c) View-3 (d) View-4

Figure 5.9: Short-scan regions for Scan 1.

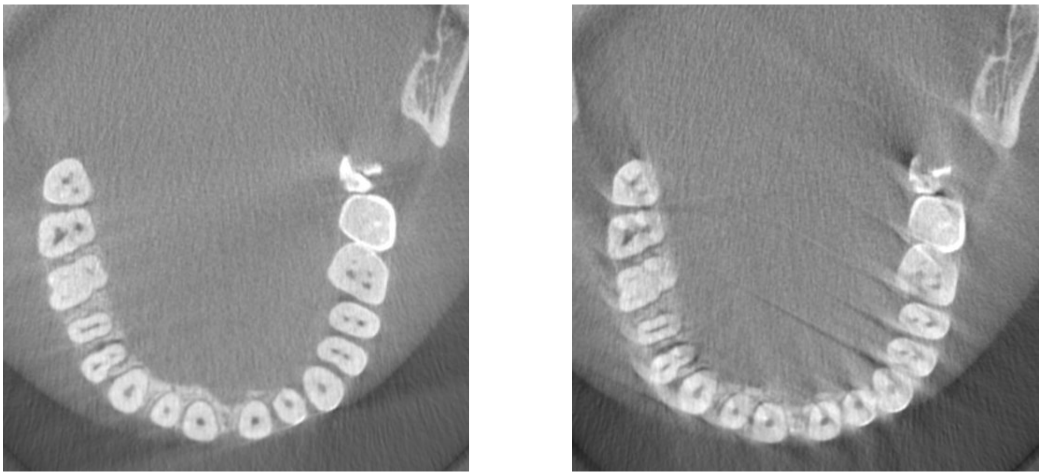


(a) View-1 (b) View-2

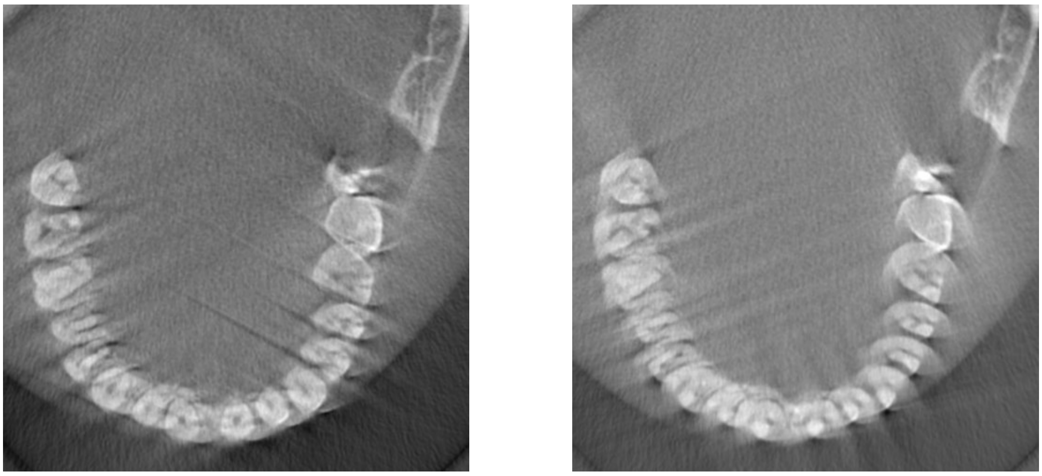


(c) View-3 (d) View-4

Figure 5.10: Short-scan regions for Scan 2.



(a) View-1 (b) View-2

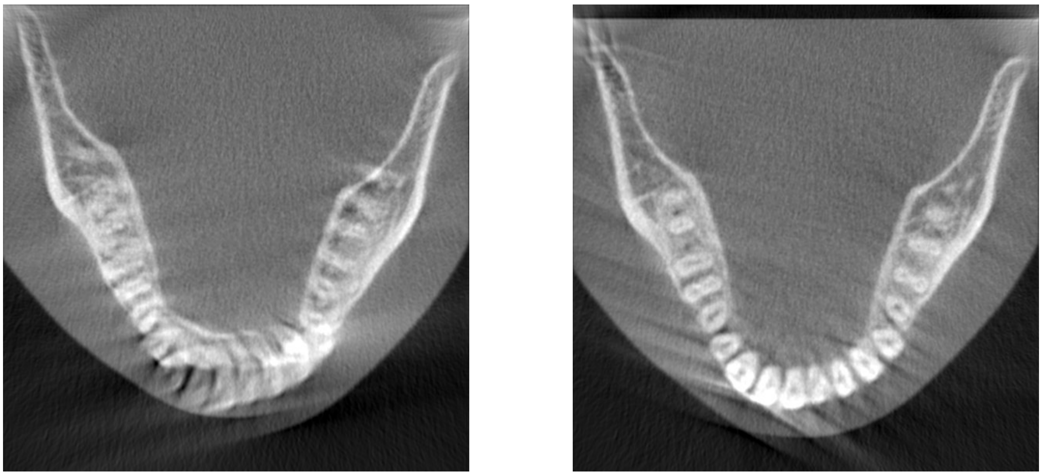


(c) View-3 (d) View-4

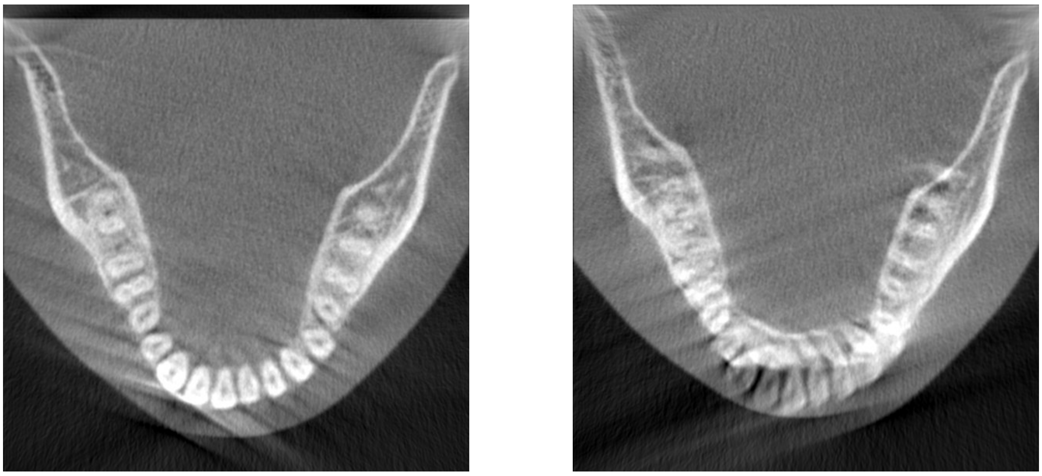
Figure 5.11: Short-scan regions for Scan 3.

5.7.2 Discussion on Complex Cases

As explained in Section 4.5.3, it is relatively common to encounter multiple instances of motion within a single full-scan, particularly in pediatric cases. In such situations, we opted to utilize the volume probability score of the network rather than the final output. Through this, we can identify a subset with lesser motion artifacts, which is used for reconstructing a reference volume.



(a) View-1 (0.90) (b) View-2 (0.88)



(c) View-3 (0.76) (d) View-4 (0.94)

Figure 5.12: Short-scan regions of the multiple motion data discussed in Section 4.5.3

As depicted in Figure 5.12, the short-scan reconstruction with minimal motion artifacts (Figure 5.12c) exhibited the lowest volume probability score (**0.76**). Therefore, it was chosen as the reference volume for the subsequent motion compensation algorithm.

5.8 Conclusion

The findings illustrate the framework's capability to consistently generate a short-scan CBCT reconstruction of satisfactory quality to be the reference for the motion compensation algorithm illustrated in Figure 4.1. Our framework can discriminate between motion artifacts and other types of artifacts. Moreover, our approach establishes the viability of effectively training a network with simulated motion data, showcasing commendable performance when applied to real motion-affected data. Further validation with patient data, would allow the performance to be evaluated in diverse populations, which is essential for developing a reliable and robust clinical tool.

Chapter 6

Motion Compensation without Detection

6.1 Introduction

In Chapter 4, our goal was to develop a non-iterative patient motion compensation algorithm. This algorithm was designed with a clean short-scan reconstruction serving as a reference for the compensation process. In the current chapter, we will delve into a regularization technique that enables us to bypass the motion detection step discussed in Chapter 5 and directly proceed to the compensation step. Due to the absence of motion detection, we do not know which projections are motion-free. Consequently, we executed the compensation process for all projection images.

6.2 Outline of the Chapter

Section 6.3 explains the proposed methodology, providing a comprehensive description of the regularization technique detailed in Subsection 6.3.2. The experimental results and subsequent discussion are delivered in Section 6.4. Finally, Section 6.5

wraps up with a summary, highlighting key insights.

6.3 Method

The pipeline of the proposed method is illustrated in Figure 6.1. An initial reference volume is reconstructed using all the acquired projections. To enhance the registration accuracy, the reference reconstruction undergoes regularization before being reprojected (Section 6.3.2). Each projection is then registered to a synthetic projection, which is created by reprojecting the reference reconstruction.

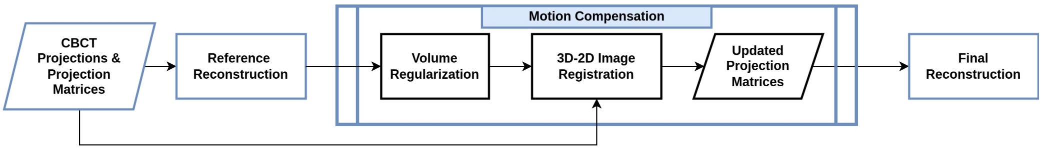
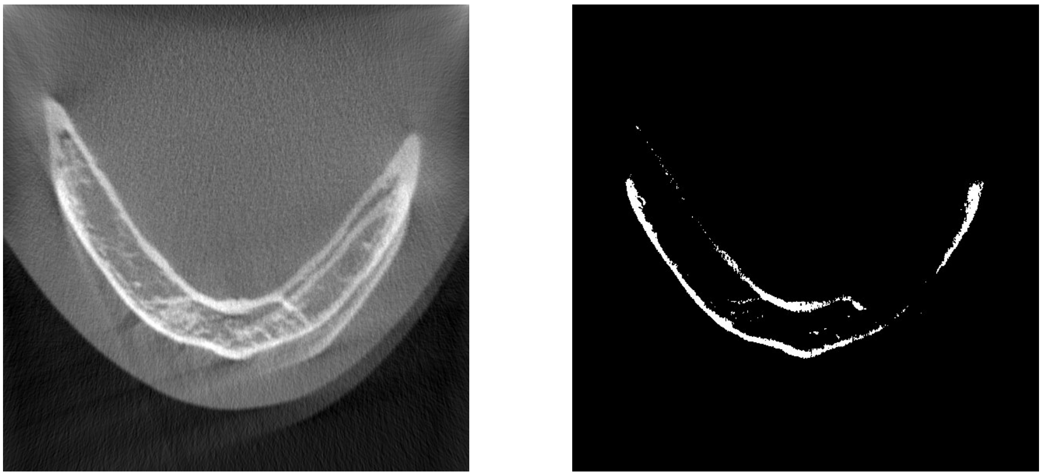


Figure 6.1: Pipeline of the Proposed Method

The process of registering 3D to 2D is executed following the same procedure detailed in Section 4.3.1, with the sole modification being the substitution of the motion-free short-scan reference volume with a regularized full-scan reference volume.

6.3.1 Reference Reconstruction

By utilizing all the acquired original projections, the FDK algorithm generates an initial reference reconstruction (Figure 6.2a). However, this reconstruction is affected by motion artifacts, including blurriness, streaks, and double contours. Reprojecting the reference reconstruction leads to the occurrence of overlapping structures in the forward projection, which makes their registration with the original projections more challenging. A regularization step is implemented to refine the reference reconstruction to overcome this difficulty.



(a) Reference Reconstruction (b) After Regularization

Figure 6.2: Regularization (The threshold is set to 1400 HU)

6.3.2 Regularization

Regularization of the reference reconstruction is necessary for three main reasons. Firstly, as mentioned above, motion artifacts like double contours or strong streaks affect the reference reconstruction. During reprojection, these types of artifacts may generate ambiguous overlapping structures on the forward projections which are not present in the real projections. Second, small structures in 2D projections are very sensitive to motion. Since the optimizer makes a step-by-step modification of the motion parameters, small structures such as those in the trabecular bone of the jaw may appear and disappear at each optimization step. They may overlap with other small or large structures. The third reason for regularization is the presence of Poisson noise in the original projections which weakens low-contrast structures even more.

Consequently, reprojecting the reference reconstruction and comparing the original and reprojected projections directly becomes challenging. To address this issue, we focus on identifying the most prominent structures, such as jawbones and teeth, by employing implicit regularization with a threshold-based segmentation of the reference reconstruction as shown in Figure 6.2b. During this particular step, our focus is on segmenting solely the hard tissues within the reference reconstruction. To accomplish this, we substitute any attenuation values exceeding a specific threshold with a constant value that closely approximates the average attenuation of the hard tissues. Any values below the threshold are discarded.

6.3.2.1 Threshold Selection

The selection of the threshold is based on the notable difference in attenuation coefficient values observed between soft tissues and hard tissues, as illustrated in Figure 6.3. In instances where less than one-fourth of the scan is affected by motion, motion artifacts, such as streaks and double contours, typically exhibit attenuation values lower than those of the jawbones and teeth. Therefore, we determined the threshold heuristically by identifying the attenuation value at which artifacts are effectively removed from the regularized volume. Choosing a threshold within this range (in our case, set at 1400 HU) efficiently eliminates the majority of artifacts found in the reference volume, as their attenuation coefficient values fall below the established threshold. The forward projection of the regularized reference volume results in a synthetic projection with high contrast, as shown in Figure 6.7b. This synthetic projection significantly contributes to achieving robust registration.

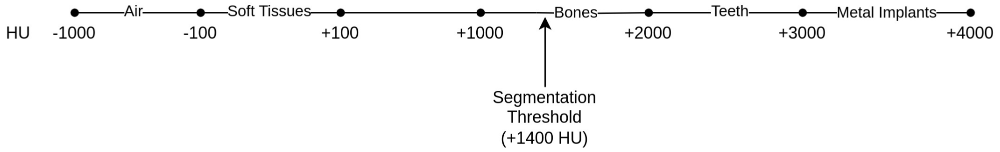
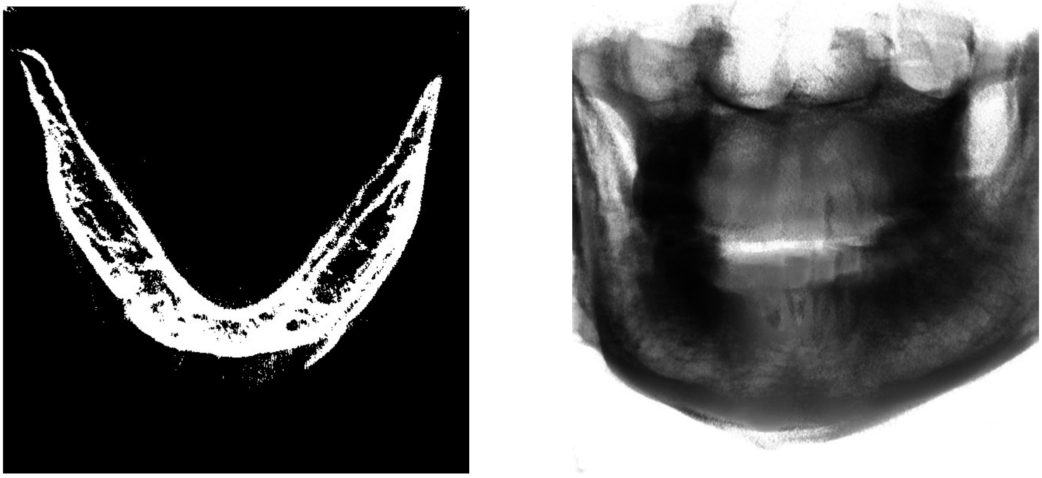
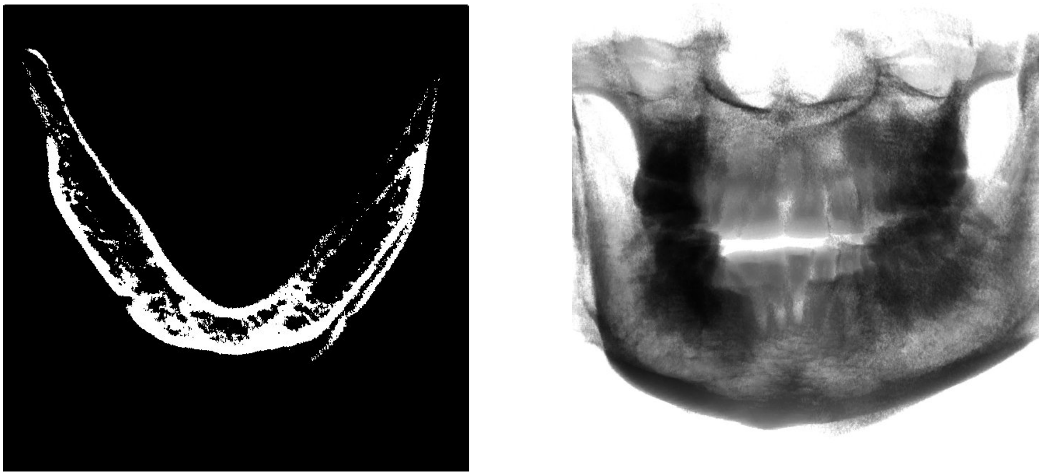


Figure 6.3: Threshold selection based on attenuation coefficient values (HU)



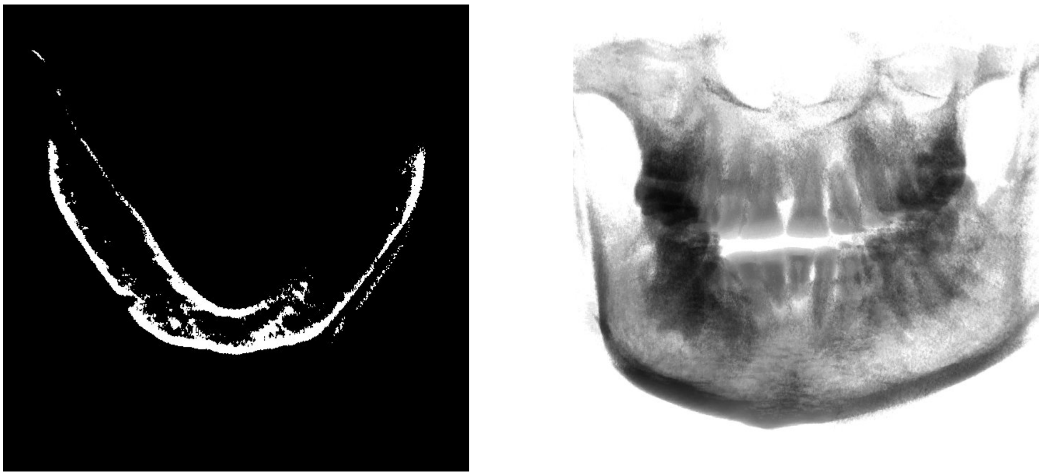
(a) Regularized Volume (b) Forward Projection

Figure 6.4: Regularization with threshold 800 HU



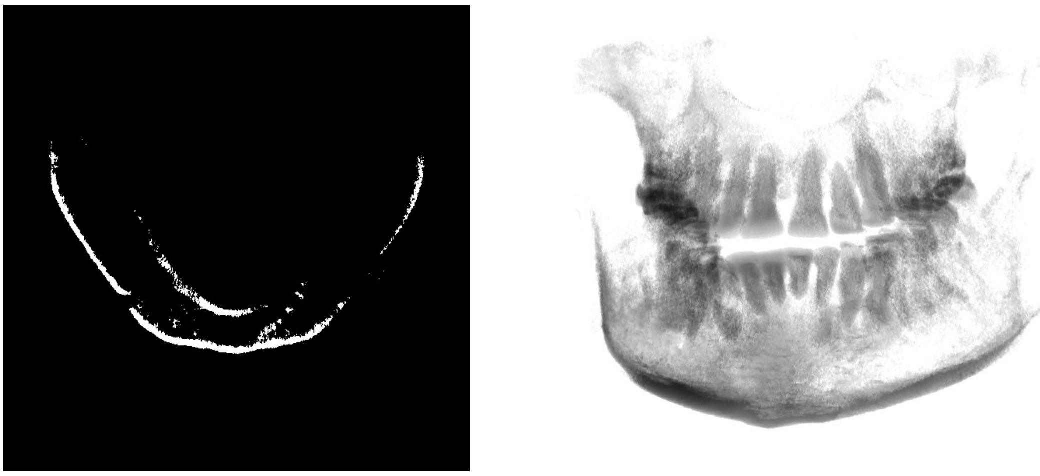
(a) Regularized Volume (b) Forward Projection

Figure 6.5: Regularization with threshold 1000 HU



(a) Regularized Volume (b) Forward Projection

Figure 6.6: Regularization with threshold 1200 HU



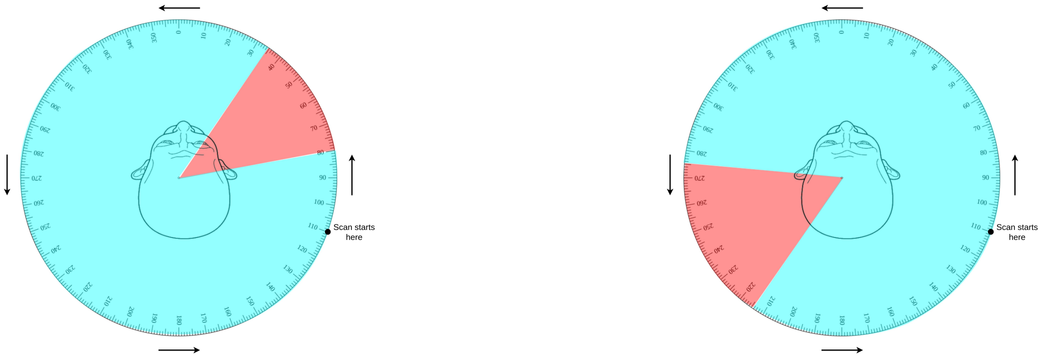
(a) Regularized Volume (b) Forward Projection

Figure 6.7: Regularization with threshold 1400 HU

6.4 Experimental Results

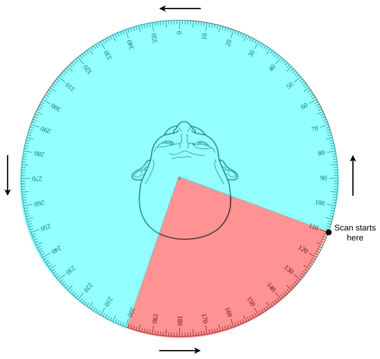
6.4.1 Data

Two returning motion data, **nodding** and **tilting** were simulated for a duration of 3, and 2 seconds respectively, for a full 360° scan of 24 seconds. Additionally, a **non-returning abrupt motion** was also simulated, where the patient remains in one pose for most of the scan duration (approximately 18 seconds), and then suddenly moves to another pose and remains still for the rest of the scan (for 6 seconds). An illustration showing the motion-affected (marked in red) and motion-free (marked in blue) region of scanned projections for the data used in this experiment is presented in Figure 6.8.



(a) Nodding

(b) Tilting



(c) Abrupt motion

Figure 6.8: Illustration showing the motion-affected subset (red) and motion-free subset (blue) of projections

6.4.2 Quantitative Results

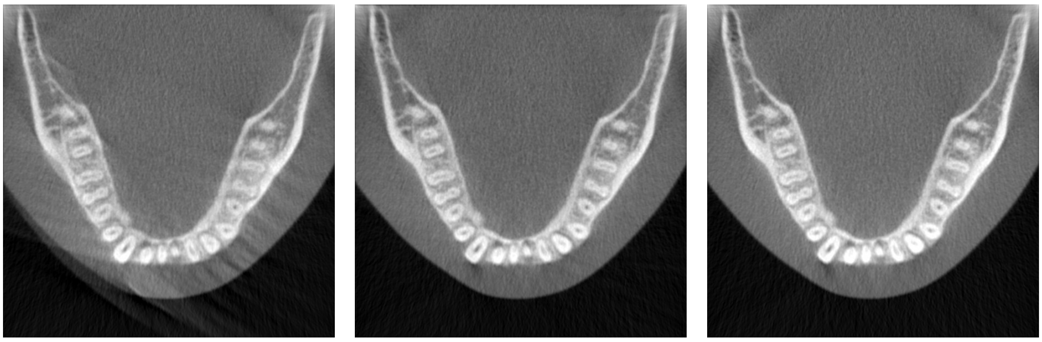
Table 6.1 provides a quantitative comparison of the motion-affected and motion-compensated reconstructions with the ground truth, as evaluated by the SSIM and RMSE metrics. For nodding and abrupt motion, there was a significant improvement in SSIM. For tilting motion the improvement was small. Abrupt motion exhibited a large improvement in RMSE but for nodding and tilting, there was only a moderate improvement in RMSE for the motion-compensated reconstruction.

Table 6.1: SSIM and RMSE values for different motions [MA $=$ motion-affected, MC $=$ motion-compensated]

	SSIM (↑)			RMSE ($\times 10^{-2}$)(↓)		
Motion type	MA	MC	Diff %	MA	MC	Diff %
Nodding	0.893	0.933	+4.02	2.04	1.31	-0.73
Tilting	0.895	0.897	+0.22	2.23	1.54	-0.69
Abrupt motion	0.799	0.917	+11.80	4.44	1.50	-2.94

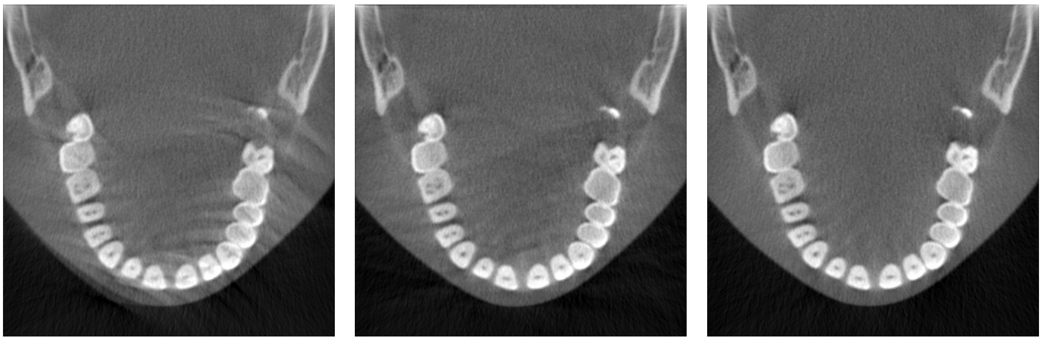
6.4.3 Qualitative Results

The qualitative comparison between the ground truth (motion-free), motion-affected, and motion-compensated reconstructions is shown in Figs. 6.9, 6.10, and 6.11. The proposed method effectively compensated motion-induced artifacts in nodding and non-returning abrupt motion, as almost all artifacts were successfully compensated. In the case of tilting motion, some minor artifacts were still present, but the crucial diagnostic structures such as the teeth were unaffected. The results show that the proposed method can successfully compensate for most motion-induced artifacts, leaving only minor artifacts in some cases.



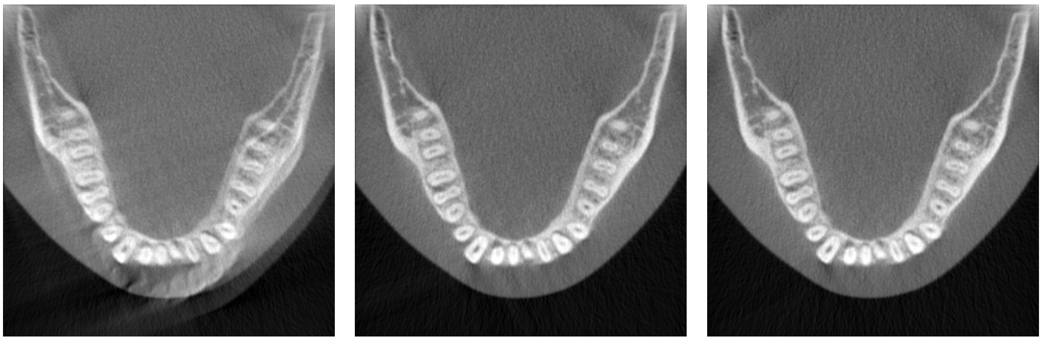
(a) Motion-affected (b) Compensated (c) Ground truth

Figure 6.9: Results for nodding motion



(a) Motion-affected (b) Compensated (c) Ground truth

Figure 6.10: Results for tilting motion



(a) Motion-affected (b) Compensated (c) Ground truth

Figure 6.11: Results for non-returning abrupt motion

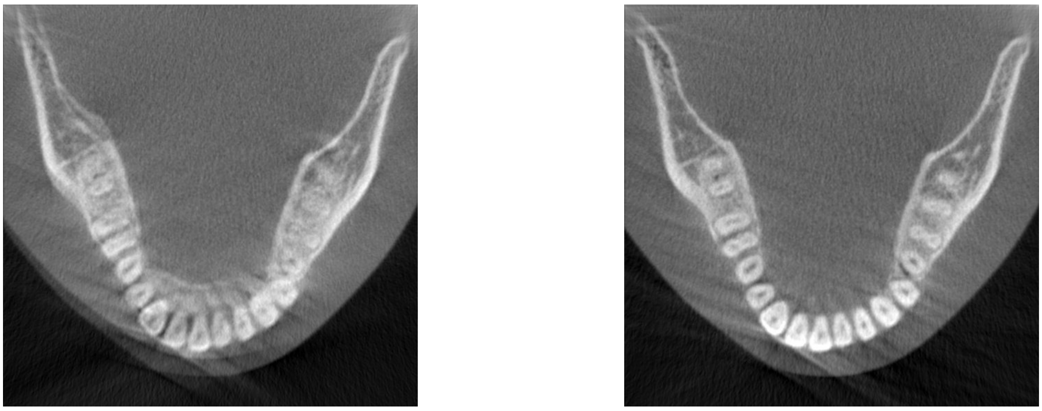
6.4.4 Discussion on Complex Cases

The data presented in Section 4.5.3, involving multiple motions that hinder obtaining a clean short-scan reference volume, was evaluated by incorporating the regularization process outlined in Section 6.3.2. Initially, we applied the regularization step to the entire scan volume and we continued with the 3D-2D registration of original and forward projected projections. As shown in Figure 6.12c, the resulting reconstruction still exhibited a higher presence of artifacts, making the diagnostic process challenging.

In the second test, we opted to regularize the short-scan reference volume, followed by 3D-2D registration for all projections. This adjustment was made because the reference volume still exhibited some motion artifacts. Figure 6.12d demonstrates a significant enhancement in reconstruction quality. Visible artifacts are successfully eliminated. Additionally, notable improvements were observed in both the SSIM and RMSE values for the case where both motion detection and regularization have been implemented.

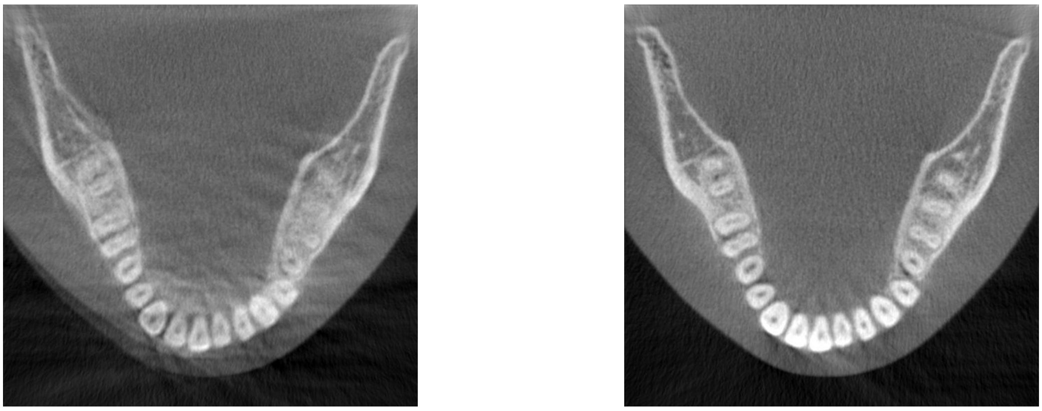
Table 6.2: SSIM and RMSE values for complex motion 4.5.3 [MoD = Motion Detection, REG = Regularization, MA = motion-affected, MC = motion-compensated]

MoD	REG	SSIM (↑)			RMSE ($\times 10^{-2}$)(↓)		
		MA	MC	Diff %	MA	MC	Diff %
Yes	No	0.791	0.898	+10.67	5.33	1.87	-3.46
No	Yes	0.791	0.781	-0.97	5.33	5.16	-0.17
Yes	Yes	0.791	0.906	+11.52	5.33	1.46	-3.87



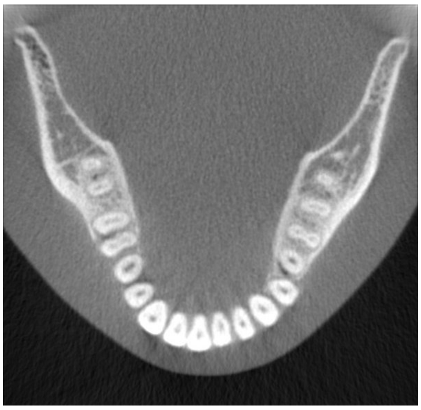
(a) Reconstruction affected by multiple motions

(b) MC reconstruction after motion detection without regularization



(c) MC reconstruction without motion detection with regularization

(d) MC reconstruction after Motion detection with regularization



(e) Ground truth

Figure 6.12: Results for multiple motions with and without regularization/motion detection (MC = motion-compensated).

6.5 Conclusion

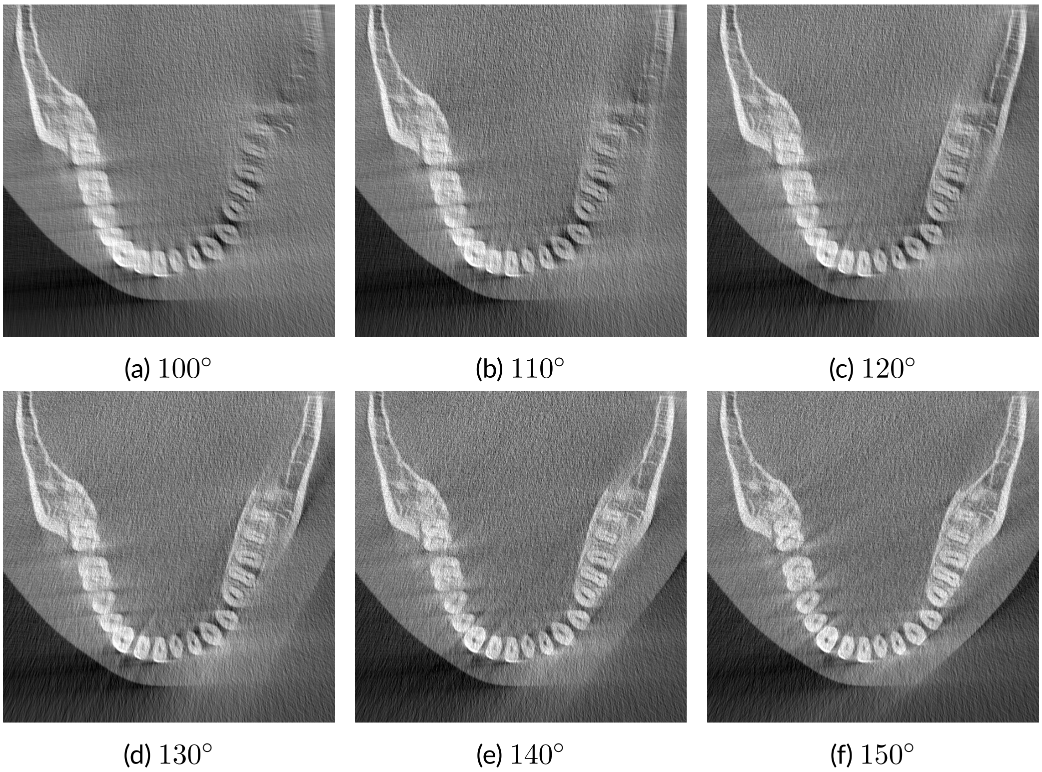
The results show that in cases where motion affects a limited part of the scan, our approach, which integrates regularization into the reference reconstruction without a previous detection step, enables the generation of high-quality reconstructions without requiring multiple full-scan reconstructions during the compensation phase. Nevertheless, when dealing with cases affected by multiple motions, an optimal strategy involves merging both motion detection and regularization to achieve a reconstruction devoid of artifacts. To validate the reliability of our proposed method across diverse scenarios, further testing is necessary, particularly involving complex motions and clinical data.

Chapter 7

Addressing Short-scan Reconstructions

7.1 Introduction

In Chapter 6, we introduced a non-iterative strategy for motion compensation in a complete 360° scan, employing a regularization technique for reference reconstruction. Building upon the findings from the preceding chapter, the current chapter proposes a two-stage approach designed to handle patient movements in short-scan reconstructions (comprising 180° plus the fan angle). Short scans are typically favored in clinical scenarios, as they involve collecting the minimum necessary number of projections to attain diagnostic quality. This typically results in a shorter scan duration and reduced radiation exposure for the patient. In contrast to full-scan reconstruction, it is difficult to extract a motion-free subset in short-scan reconstruction. This is because the scans shorter than 160° have more additional artifacts in the reconstruction arising due to the limited projection angle (Figure 7.1), making it challenging for the motion detection algorithm to distinguish motion artifacts from those arising due to the limited projection angle. Consequently, we propose a solution here that does not rely on motion detection.


(a) $100°$ (b) $110°$ (c) $120°$

(d) $130°$ (e) $140°$ (f) $150°$

Figure 7.1: Reconstructions with different projection angles.

7.2 Outline of the Chapter

Section 7.3 details the methodology. In Section 7.4, we discuss the outcomes of the experiments and explore different cases. Section 7.5 outlines adjustments made to the proposed method for cases with extended patient motion. Lastly, Section 7.6 concludes with a brief summary.

7.3 Method

The proposed method is illustrated in Figure 7.2. The process begins with an initial motion-affected reference reconstruction generated from the acquired projections.

This is followed by the same motion compensation technique explained in Section 6.3. After the compensation process is completed, an intermediate reconstruction is performed using the estimated motion parameters. This intermediate reconstruction is used as the reference for the next stage of the compensation algorithm. This process can be repeated until the artifacts are eliminated. A final high-quality reconstruction is performed using the motion parameters estimated in the n^{th} stage.

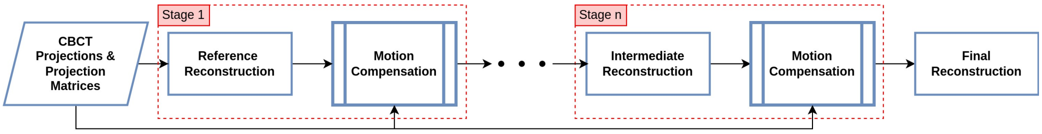


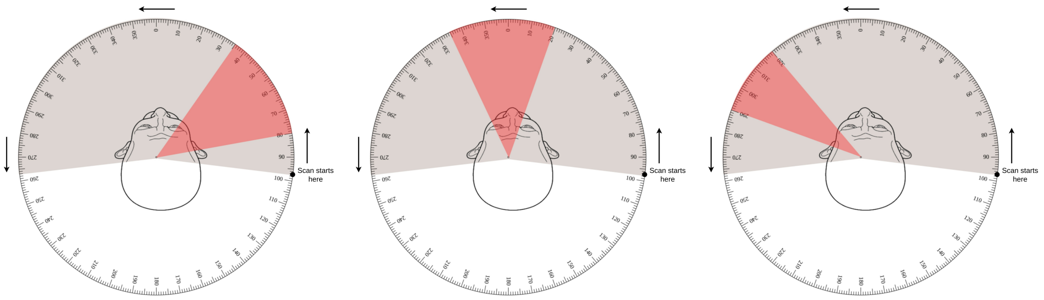
Figure 7.2: Pipeline of the proposed method

From our experiments, it turned out that after two stages the algorithm reached convergence. Further intermediate stages were not necessary since they did not provide any significant improvement in the final result.

7.4 Experimental Results

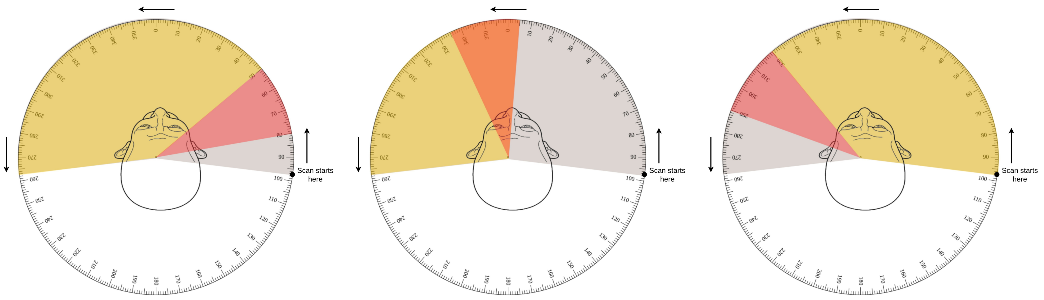
7.4.1 Data

Our previous findings demonstrate the effectiveness of the compensation algorithm in addressing various types of motion (nodding, tilting, lateral rotation), including scenarios involving multiple motions. In this experiment, we sought to assess the algorithm's performance across different positions defined by the relative orientation of the detector to the patient's head. We simulated nodding motions in three distinct positions: first with the detector on the right side, followed by one with the detector in front, and finally, one with the detector on the left side of the patient's head. This experimental setup was designed to investigate the algorithm's limitations concerning the detector's position. We simulated three returning nodding motions, where the patient returns to the initial position, as illustrated in Figure 7.3. Additionally, we simulated three non-returning nodding motions, where the patient does not return to the initial pose after movement, as illustrated in Figure 7.4. These scenarios are challenging because non-returning motions create double contours in the reconstruction (Figure 7.10), affecting the majority of projections and complicating diagnosis. To evaluate the quality of the final reconstruction, we performed a motion-free short-scan reconstruction of 194° to serve as a ground truth for comparison.



(a) View - 1 (b) View - 2 (c) View - 3

Figure 7.3: Illustration showing the motion-affected subset (red) and motion-free subset (gray) of projections in the short-scan angle for nodding motion (returning)

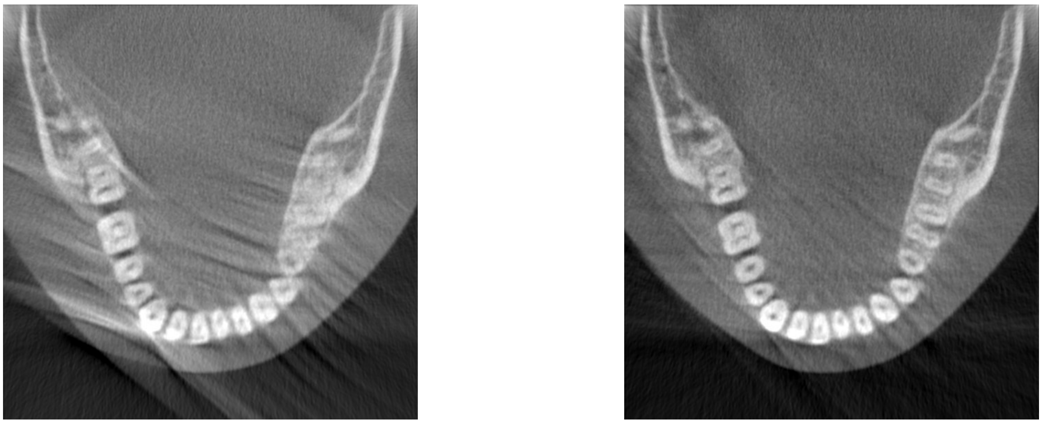


(a) View - 1 (b) View - 2 (c) View - 3

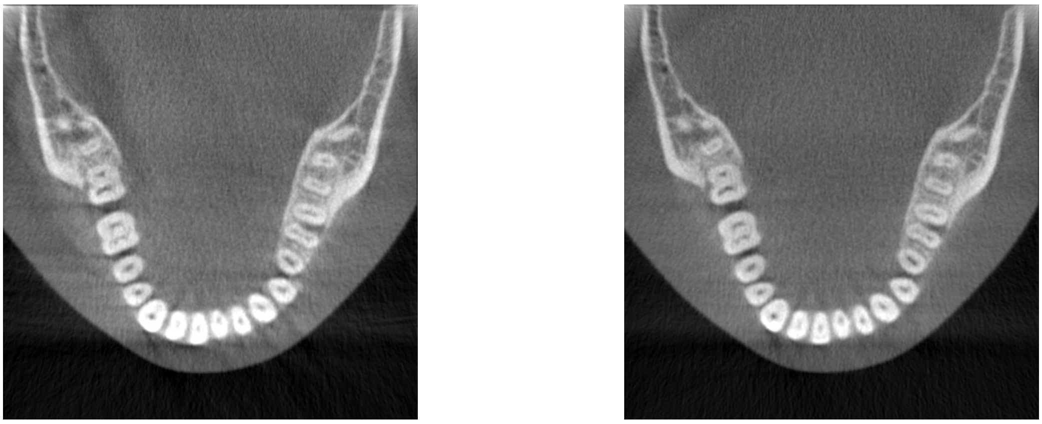
Figure 7.4: Illustration showing the motion-affected and motion-free subsets for nodding motion (non-returning). 'red' indicates patient movement, while 'gray' and 'yellow' signify instances where the patient remains stationary but in distinct poses.

7.4.2 Returning Motion Results

Figures 7.5, 7.6, and 7.7 present a comparison among the reconstructions affected by motion, motion-compensated reconstructions, and the ground truth for the three test cases of returning nodding motions mentioned above. In all three scenarios, the second stage of motion compensation notably reduced motion artifacts. It's important to highlight that although some minor artifacts are noticeable, crucial diagnostic structures remain free from motion artifacts and are clearly visible. The maximum duration of applied motion is 3 seconds, accounting for approximately one-fourth of the short-scan duration.

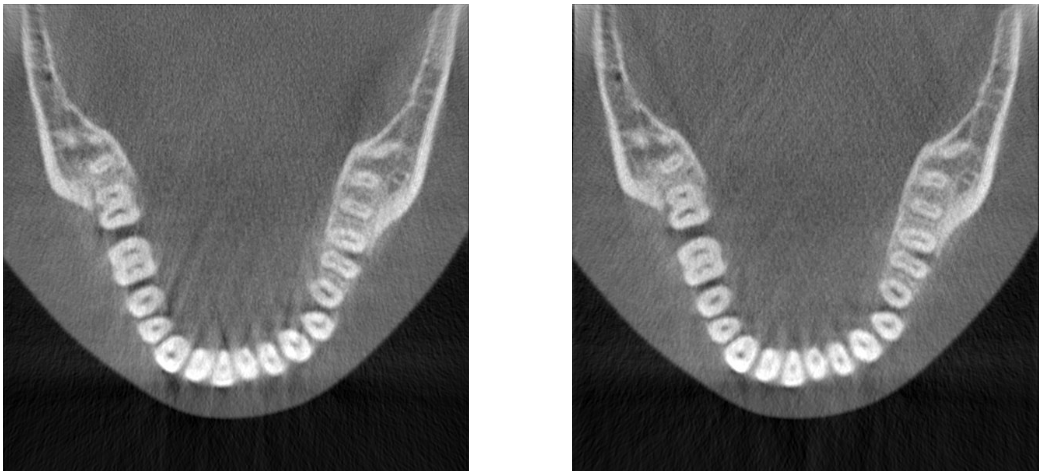


(a) Motion-affected (b) Stage 1

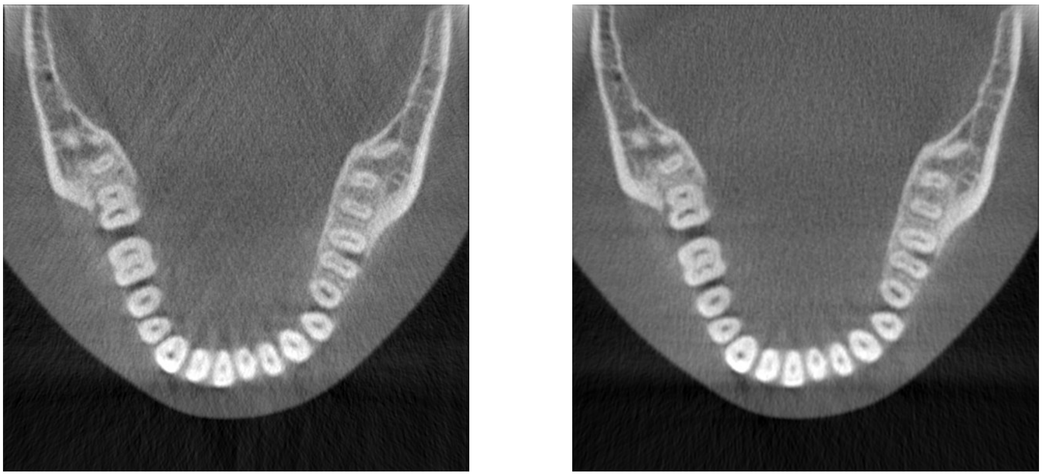


(c) Stage 2 (d) Ground truth

Figure 7.5: Results for returning nodding motion (View-1)

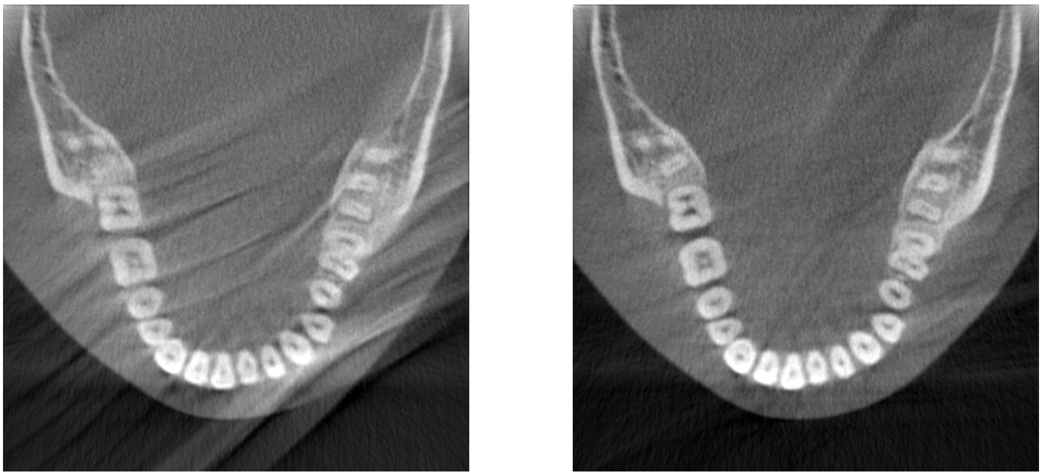


(a) Motion-affected (b) Stage 1

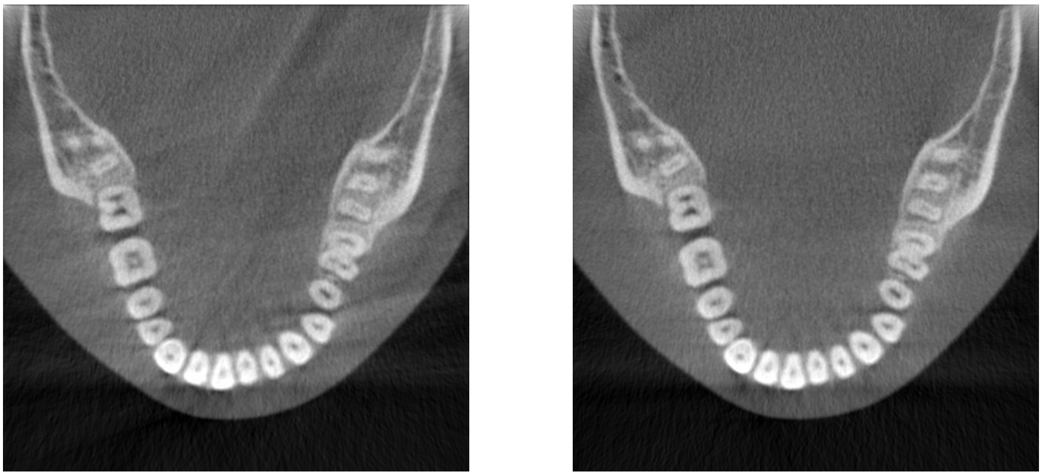


(c) Stage 2 (d) Ground truth

Figure 7.6: Results for returning nodding motion (View-2)



(a) Motion-affected (b) Stage 1



(c) Stage 2 (d) Ground truth

Figure 7.7: Results for returning nodding motion (View-3)

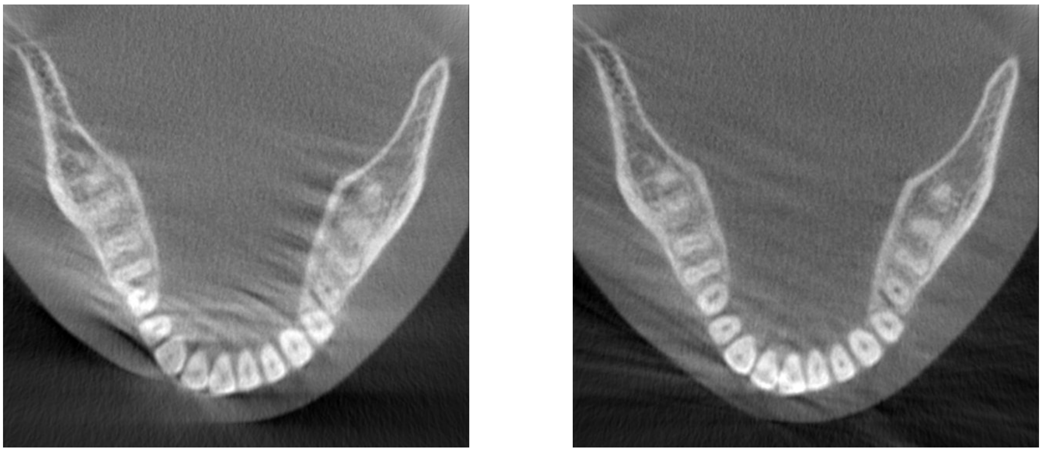
Table 7.1 provides a summary of the quantitative evaluations for the three tests of returning nodding motions. Notably, there is a substantial enhancement in View-1 and a slight improvement in View-3. However, in View-2, the SSIM score decreases for the motion-compensated reconstruction. This discrepancy arises because the SSIM score is computed for the entire 3D volume, encompassing a significant portion of free space where diagnostic structures are absent. Nevertheless, despite this decrease in SSIM score, the clinical value of the motion-compensated image surpasses that of the motion-affected reconstruction.

Table 7.1: SSIM and RMSE values for nodding motions (returning) [MA = motion-affected, MC = motion-compensated]

	SSIM (↑)			RMSE ($\times 10^{-2}$)(↓)		
View	MA	MC	Diff %	MA	MC	Diff %
View - 1	0.783	0.826	+4.33	4.22	4.17	-0.05
View - 2	0.837	0.817	-1.98	2.49	2.47	-0.02
View - 3	0.825	0.835	+0.91	3.32	2.85	-0.47

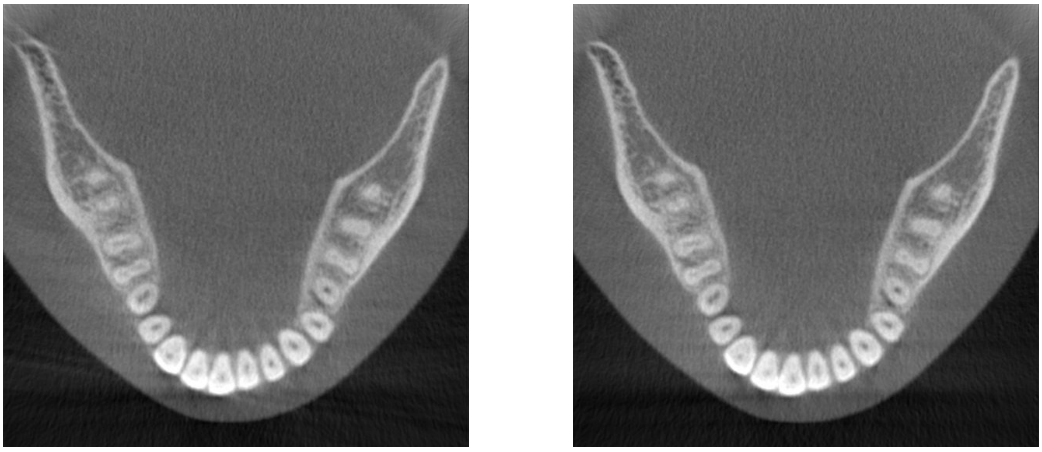
7.4.3 Non-returning Motion

Figures 7.8, 7.10, and 7.11 present a comparison among reconstructions affected by motion, motion-compensated reconstructions, and the ground truth for the three instances of non-returning nodding motions outlined in Section 7.4.1. In the case of View-1, the motion compensation algorithm demonstrated strong performance, notably reducing motion artifacts. This success is due to the fact that nearly three-fourths of the scan duration remained unaffected by motion. Consequently, the regularization technique effectively removed the minor artifacts present in the reference reconstruction. The SSIM and RMSE scores presented in Table 7.2 are consistent with the qualitative analysis, reflecting similar outcomes.



(a) Motion-affected

(b) Stage 1

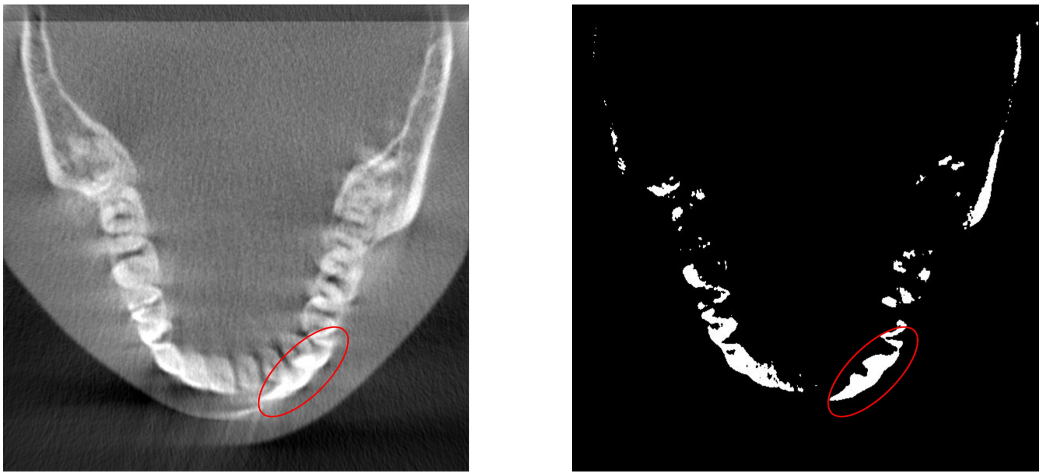


(c) Stage 2

(d) Ground truth

Figure 7.8: Results for non-returning nodding motion (View-1)

However, in the remaining two scenarios (View-2 & View-3), despite the simulated patient movement being brief, the challenge arises from it being a non-returning motion. This poses challenges for the regularization technique, particularly when artifacts have higher attenuation values like those of bones and teeth (Figure 7.9). Consequently, the compensation algorithm struggled to effectively eliminate motion artifacts in these instances. To address such complex cases, a modification of the proposed method is detailed in Section 7.5.



(a) Motion-affected (b) Stage 1

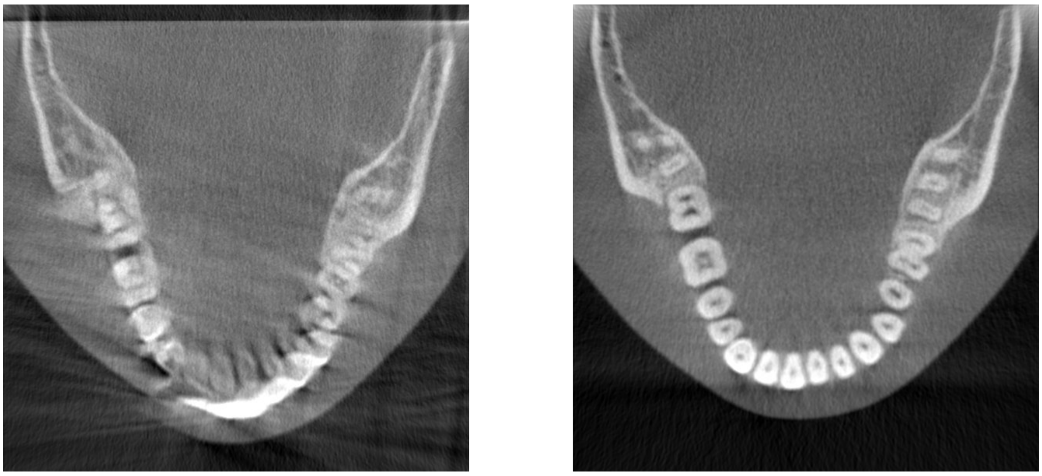
Figure 7.9: Regularization results of non-returning nodding motion (View-2). The highlighted area indicates an artifact with a higher attenuation value that persists even after regularization.

Table 7.2: SSIM and RMSE values for nodding motions (non-returning) [MA = motion-affected, MC = motion-compensated]

	SSIM (↑)			RMSE ($\times 10^{-2}$)(↓)		
View	MA	MC	Diff %	MA	MC	Diff %
View - 1	0.743	0.833	+8.97	5.78	2.47	-3.31
View - 2	0.628	0.558	-7.02	7.53	8.26	+0.73
View - 3	0.726	0.750	+2.38	5.95	3.50	-2.45

(a) Motion-affected

(b) Stage 1

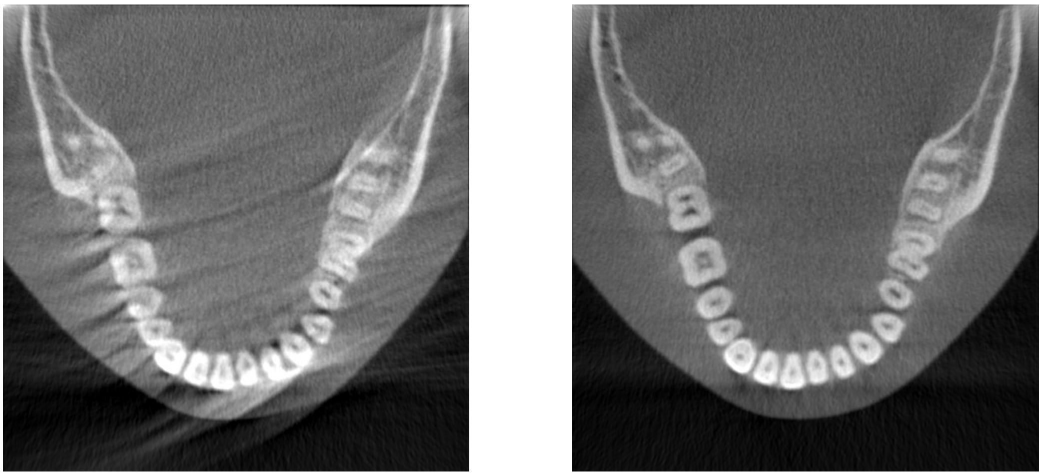


(c) Stage 2

(d) Ground truth

Figure 7.10: Results for non-returning nodding motion (View-2)

(a) Motion-affected (b) Stage 1



(c) Stage 2 (d) Ground truth

Figure 7.11: Results for non-returning nodding motion (View-3)

7.5 Partial Angle Method

In cases where a significant number of projections are impacted by motion in a short-scan reconstruction (as illustrated in Figures 7.10 and 7.11), the compensation algorithm encounters limitations. To tackle this issue, we introduced the partial angle compensation approach (Hahn et al., 2017), replacing the first stage of motion compensation with an incremental compensation strategy. Here, a very small subset of projections, preferably at the beginning or end of the scan, is considered to reconstruct an ultra-short-scan reference volume. This method works under the assumption that the patient remains motionless for a short duration, either at the scan's beginning or end. The workflow of the proposed method is illustrated in Figure 7.12. Starting with the ultra-short-scan reference volume, we gradually incorporate a small fraction of the remaining projections, reconstructing a new reference volume, and compensating these additional projections as they are included. Just like the previously mentioned short-scan technique outlined in Section 7.3, our modified approach demonstrated convergence after two stages. Additional intermediary stages proved unnecessary, as they did not yield any substantial improvement in the final result.

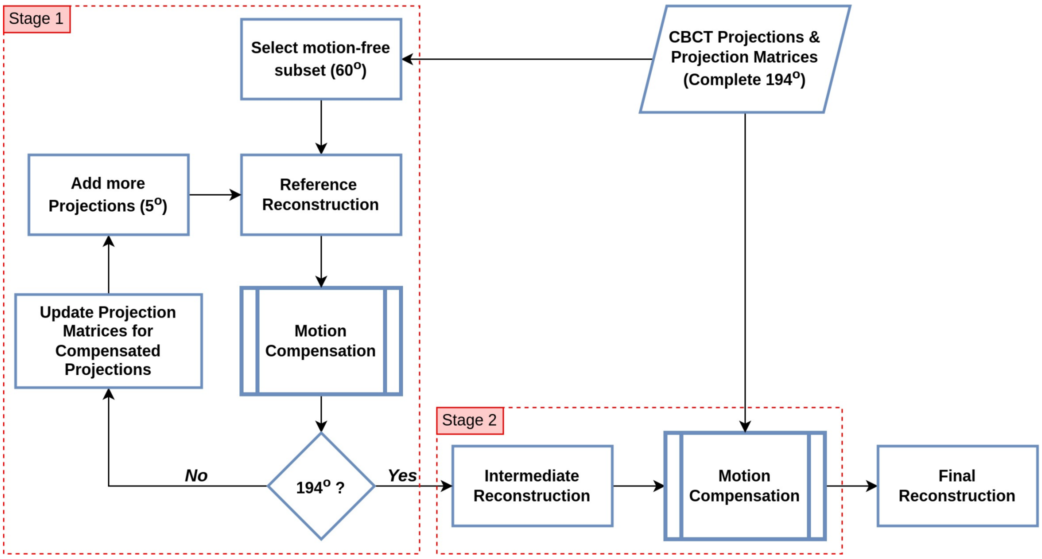
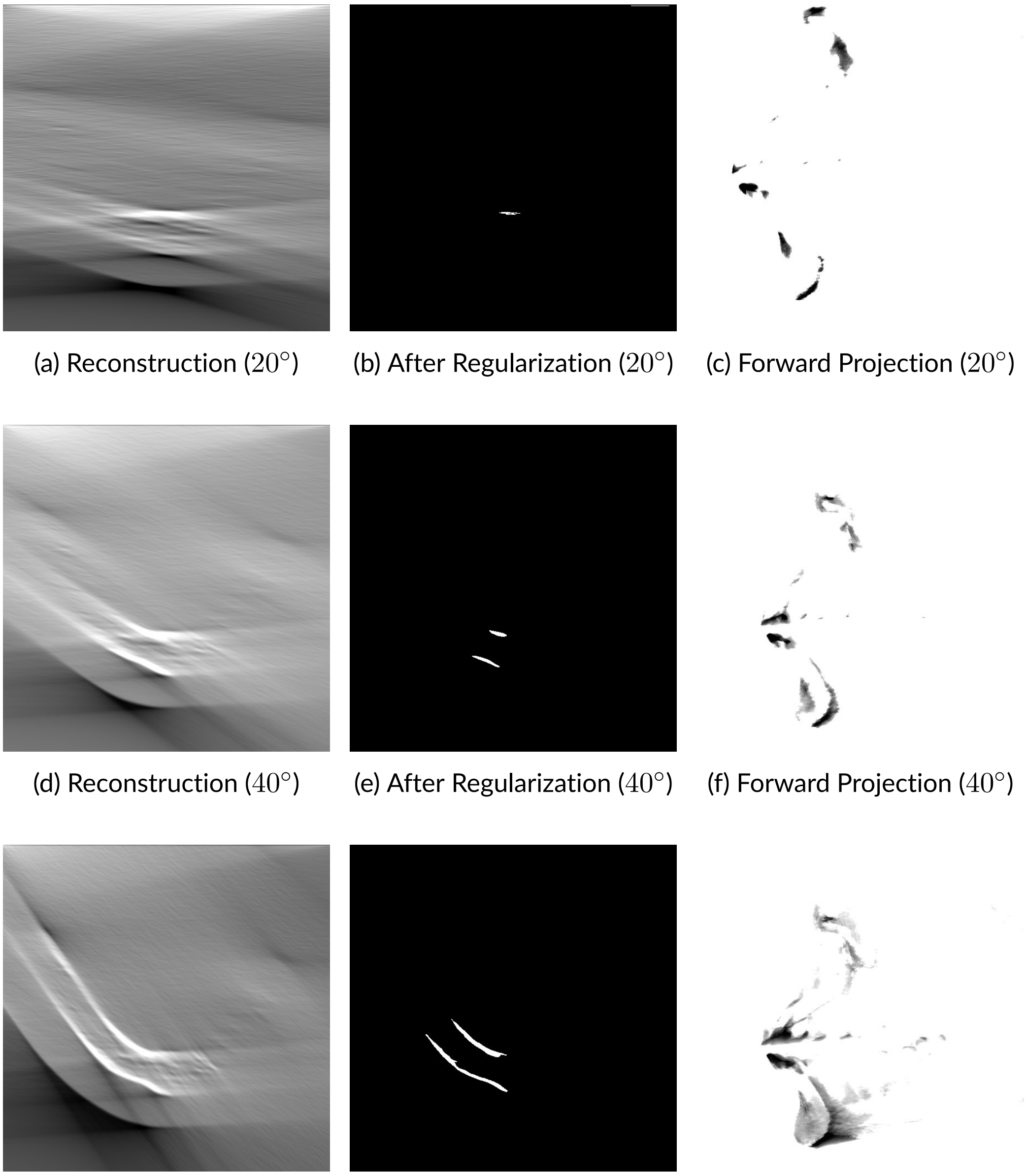


Figure 7.12: Pipeline of the proposed method

Our experiments revealed that the minimum number of projections needed for the ultra-short-scan reference volume is 60° of projections, which is less than one-third of the scan. This necessity arises from the absence of gradients in the synthetic forward projection when the projection angle is below 60°, as depicted in Figure 7.13. Furthermore, we observed that adding 5 degrees of projections at a time proved effective for compensation.



(a) Reconstruction ($20°$) (b) After Regularization ($20°$) (c) Forward Projection ($20°$)

(d) Reconstruction ($40°$) (e) After Regularization ($40°$) (f) Forward Projection ($40°$)

(g) Reconstruction ($60°$) (h) After Regularization ($60°$) (i) Forward Projection ($60°$)

Figure 7.13: Illustration of partial angle reconstructions with different projection angles. For $< 60°$, the forward projection doesn't have sufficient edges (gradients).

7.5.1 Results

We assessed the performance of the partial angle method using the two challenging scenarios (non-returning View-2 and View-3) detailed in Section 7.4.1, where the short-scan motion compensation algorithm failed. As presented in Table 7.3, the performance of the partial angle method is significantly larger when compared to the short-scan method, as evaluated through SSIM and RMSE metrics. Additionally, the quality of the motion-compensated reconstructions depicted in Figures 7.14 and 7.15 has substantially improved, successfully eliminating almost all artifacts. This approach shows potential for applications involving large patient movements.

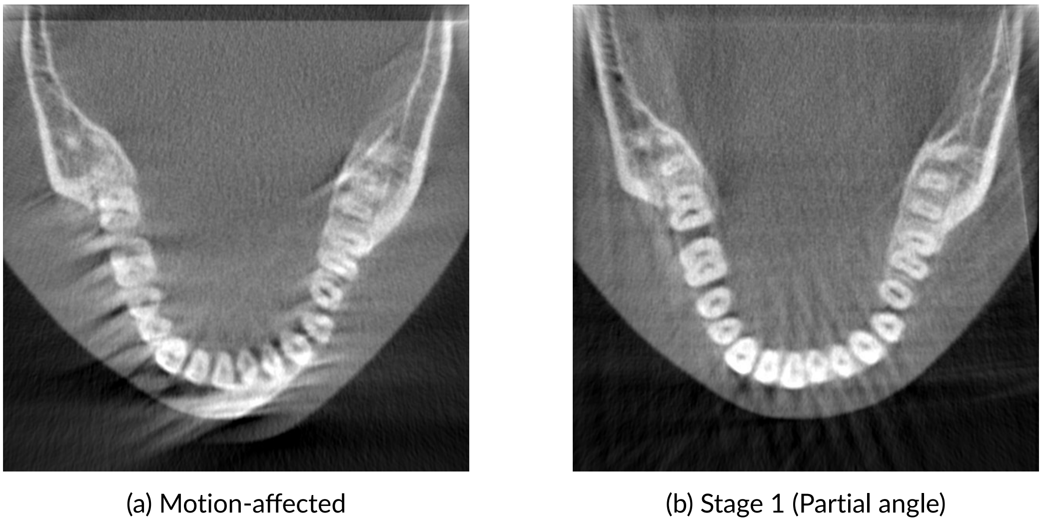
Table 7.3: SSIM and RMSE values for nodding motions (non-returning) using Partial Angle Method [MA = motion-affected, MC = motion-compensated]

	SSIM (↑)			RMSE ($\times 10^{-2}$)(↓)		
View	MA	MC	Diff %	MA	MC	Diff %
View - 2	0.628	0.774	+14.52	7.53	3.22	-4.31
View - 3	0.726	0.814	+8.78	5.95	2.51	-3.44

(a) Motion-affected (b) Stage 1 (Partial angle)

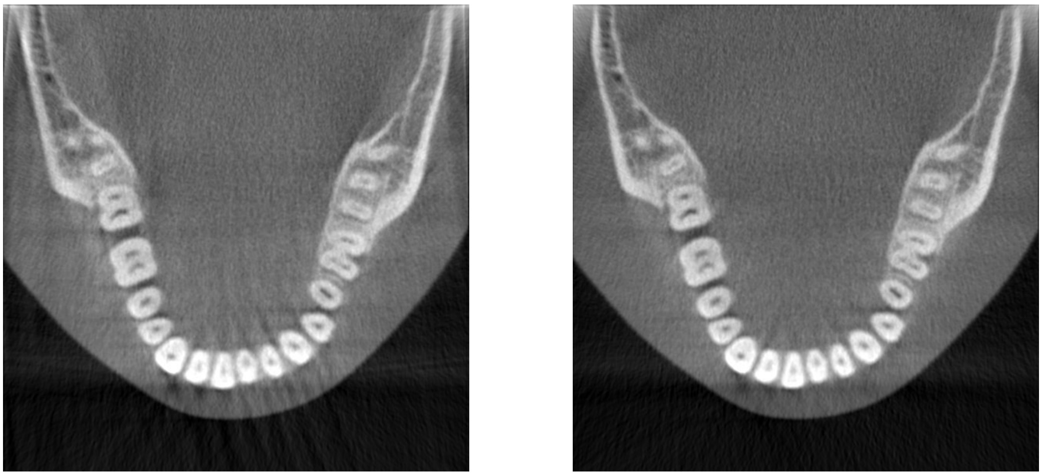
(c) Stage 2 (d) Ground truth

Figure 7.14: Results for non-returning nodding motion (View-2)



(a) Motion-affected

(b) Stage 1 (Partial angle)



(c) Stage 2

(d) Ground truth

Figure 7.15: Results for non-returning nodding motion (View-3)

7.6 Conclusion

The partial angle motion compensation method emerges as a promising solution to address challenges in maxillofacial cone-beam computed tomography (CBCT) imaging, particularly in instances of prolonged patient movements. Utilizing an incremental compensation strategy and reconstructing an ultra-short-scan reference volume, this method effectively improves the quality of motion-compensated reconstructions. The demonstrated performance enhancement, indicated by higher SSIM and lower RMSE metrics, suggests its potential applicability in clinical settings. However, it assumes that the patient is stationary either at the beginning or end of the scan, introducing uncertainty given the unpredictable duration of patient motion. To ensure proper functionality in all scenarios, a strategy to obtain a motion-free subset corresponding to a 60° projection angle is necessary.

Chapter 8

Conclusion

8.1 Summary of the Thesis

This thesis outlines methods for compensating motion-induced artifacts in maxillofacial CBCT imaging. Several factors like the low dose nature of CBCT, the standing posture of the patient while scanning, and the need for high-resolution reconstructions pose challenges to the compensation algorithms. Consequently, we simplified the compensation algorithm by considering only rigid motions that follow the standard protocols of CBCT devices. Another critical concern tackled is the time complexity associated with existing compensation algorithms. While many employ iterative reconstruction methods, our implementation adopts a non-iterative approach, utilizing the FDK reconstruction algorithm to speed up the compensation process and mitigate the time-intensive nature of iterative reconstruction algorithms.

The first method addresses the compensation of motion-induced artifacts in full-scan (360°) reconstructions. We implemented three strategies: one involving motion detection, another incorporating regularization, and a third utilizing both motion detection and regularization. The first strategy is suitable for scenarios where a motion-free subset of projection images, covering approximately 180° projection angle, is available. This subset is employed to reconstruct a reference volume, which, in turn, is utilized in the compensation process. The second strategy is applicable when motion affects less than one-fourth of the scan. In this case, regularization is applied to the initial reconstruction to eliminate motion-induced artifacts, and the regularized reconstruction serves as the reference volume for compensation. However, both of these approaches have limitations due to the unpredictable duration of patient motion. This brings us to the third approach, which combines both motion detection and regularization. Here, a subset of projections least affected by motion is considered for reference reconstruction. The resulting reference volume is then subjected to regularization to eliminate remaining artifacts and is subsequently used in the compensation process. This approach proves effective in most scenarios, provided the patient's movement is not excessive, at least during a part of the scan.

Secondly, we addressed the compensation of motion-induced artifacts in short-scan ($\approx 180^{\circ}$) reconstructions. Short scans are particularly favored by clinicians as they enable the reconstruction of a volume with sufficient diagnostic quality by collecting the minimum necessary projections. This approach not only reduces scan time but also minimizes radiation exposure to the patient. However, detecting motion in short-scan reconstructions presents a significant challenge, as the shorter scans exhibit artifacts very similar to motion artifacts. Consequently, we chose the regularization-based strategy to address motion artifacts in short scans. We implemented a multi-stage approach, repeating the compensation process until achieving an artifact-free reconstruction. With only two stages, this method effectively compensated for motion affecting less than one-third of the scan duration. However, challenges arose in scenarios involving prolonged motion. To address this, we introduced a partial-angle motion compensation approach with an incremental compensation strategy. This approach is effective in compensating for prolonged motions (e.g., non-returning motion), provided the patient remains still for a short duration, either at the beginning or at the end of the scan.

8.2 Perspectives

Upon concluding this thesis, we identified certain aspects of the proposed motion compensation approach that deserve additional investigation and analysis. These aspects have been categorized into two main issues: Scientific and Engineering. This categorization is based on the nature of the investigations, with Scientific issues focusing on theoretical foundations and conceptual understanding, while Engineering issues focus on practical implementation, optimization, and technical refinements.

8.2.1 Scientific Issues

8.2.1.1 Motion Detection

The current implementation of the motion detection algorithm assesses multiple short-scan volumes extracted from a full scan, classifying them as either affected by motion or motion-free. It is required that these short-scan reconstructions encompass a minimum projection angle of 180°. To broaden its applicability in diverse scenarios, it is valuable to explore alternative network architectures capable of detecting motion across varying projection angles below 180°, rather than a fixed one. A notable challenge emerges when handling projection angles less than 180°, especially when the angle is very low. In such cases, the reconstruction lacks sufficient edges (gradients) and also presents more artifacts due to the limited angle of projections used. These factors should be taken into account as they create complexities for the network in identifying a motion-free subset.

8.2.1.2 Partial-angle Method

The partial-angle compensation method outlined in this thesis shows greater potential in compensating for prolonged and persistent motion. This can be very helpful for cases where the patient cannot be idle due to some physiological conditions like in the case of a Parkinson's patient. However, it necessitates a motion-free ultra-short-scan reference volume corresponding to a 60° projection angle, to initiate the compensation process. In such instances, it is beneficial to investigate methods for obtaining an ultra-short-scan reference to initiate compensation. Various approaches can be considered, including the utilization of compensation techniques like the one suggested by Sun et al. (2021), which involves an iterative reconstruction algorithm during the compensation phase. Upon obtaining projection images corresponding to a 60° angle, the partial angle method can be implemented. This approach ensures a quicker compensation process compared to applying iterative compensation for the entire scan.

An alternative approach involves employing motion detection to directly classify motion-free subsets within the reconstructed volume, as illustrated in Figure 8.1. This method would also allow us to extract an ultra-short-scan reference volume for implementing partial-angle compensation.

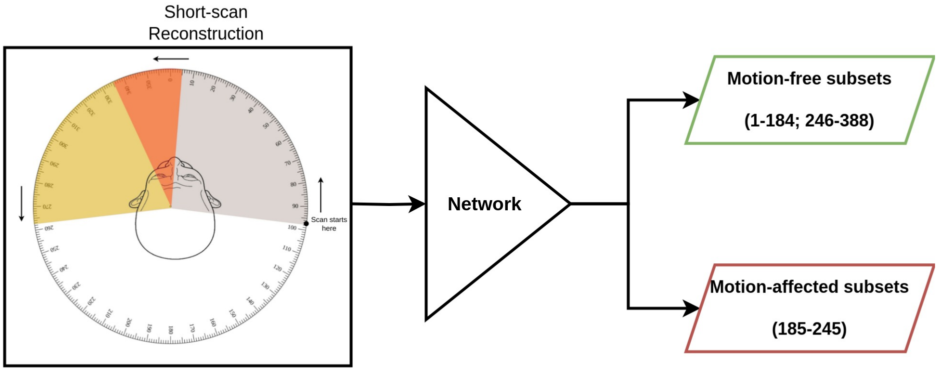


Figure 8.1: Illustration of motion detection framework for Partial-angle compensation method

8.2.1.3 Clinical Evaluation

While the rigid motion assumption enables the simulation of realistic motions, it is crucial to validate the algorithm through its application to actual patient data obtained in a clinical setting. Conducting additional research and validation across various clinical settings can significantly contribute to establishing the robustness and versatility of this method in improving the diagnostic quality of maxillofacial CBCT imaging.

8.2.2 Engineering Issues

The utilization of the FDK reconstruction algorithm in our compensation method already significantly enhanced the speed of the method. To further expedite the procedure, we can introduce an interpolation process. Given that head motion is continuous, certain projections can be skipped during compensation. For example, if p_i represents the current projection for compensation, we can directly compensate for the p_{i+5} projection skipping the projections in between. For the projections (p_{i+1}, p_{i+2}, p_{i+3}, and p_{i+4}), the motion parameters can be interpolated using the updated motion parameters of p_i and p_{i+5}.

Another aspect to consider is integrating an outlier detection mechanism. Given the continuous nature of motion, if the compensation algorithm results in out-of-plane motion parameters for a projection over a sequence of successive projections, this result could be substituted with interpolated motion parameters. This approach aims to enhance the algorithm's robustness by handling occasional instances of compensation irregularities, ensuring a smoother and more consistent motion compensation process.

Bibliography

Abdurahman, S., Frysch, R., Pfeiffer, T., Beuing, O., and Rose, G. (2022). Comparative evaluation of cone-beam consistency conditions for mono-material beam hardening correction in flat detector computed tomography. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 7(4):354–371.

Ahmed, O., Song, Y., Xie, Z., and Ayoub, A. A. (2020). Calculation of beam hardening in industrial x-ray computed tomography and its correction using filtration and linearization methods. In *IOP Conference Series: Earth and Environmental Science*, volume 461(1), page 012082. IOP Publishing.

Ahmed, O. M. H., Song, Y., and Xie, Z. (2019). Theoretical beam hardening correction for industrial x-ray computed tomography. *J. Eng. Technol. Sci.*, 51(6):870–881.

Amirian, M., Montoya-Zegarra, J. A., Herzig, I., Eggenberger Hotz, P., Lichtensteiger, L., Morf, M., Züst, A., Paysan, P., Peterlik, I., Scheib, S., et al. (2023). Mitigation of motion-induced artifacts in cone beam computed tomography using deep convolutional neural networks. *Medical Physics*.

An, K., Wang, J., Zhou, R., Liu, F., and Wu, W. (2020). Ring-artifacts removal for photon-counting ct. *Optics Express*, 28(17):25180–25193.

Andersen, A. H. and Kak, A. C. (1984). Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm. *Ultrasonic imaging*, 6(1):81–94.

Arrowsmith, C., Reiazi, R., Welch, M. L., Kazmierski, M., Patel, T., Rezaie, A., Tadic, T., Bratman, S., and Haibe-Kains, B. (2021). Automated detection of dental artifacts for large-scale radiomic analysis in radiation oncology. *Physics and Imaging in Radiation Oncology*, 18:41–47.

Bhadra, S., Villa, U., and Anastasio, M. A. (2022). Mining the manifolds of deep generative models for multiple data-consistent solutions of ill-posed tomographic imaging problems. *arXiv preprint arXiv:2202.05311*.

Birklein, L., Niebler, S., Schömer, E., Brylka, R., Schwanecke, U., and Schulze, R. (2023). Motion correction for separate mandibular and cranial movements in cone beam ct reconstructions. *Medical physics*, 50(6):3511–3525.

Botticelli, S., Verna, C., Cattaneo, P., and et al. (2011). Two- versus three-dimensional imaging in subjects with unerupted maxillary canines. *European Journal of Orthodontics*, 33(4):344–349.

Braun, S., Hnat, W. P., Fender, D. E., and Legan, H. L. (1998). The form of the human dental arch. *The Angle Orthodontist*, 68(1):29–36.

Buchanan, A. (2023). Cone beam computed tomography: What's left to know? *The Journal of Prosthetic Dentistry*, 129(2):241–242.

Bushberg, J. T. and Boone, J. M. (2011). *The essential physics of medical imaging*. Lippincott Williams & Wilkins.

De Silva, T., Uneri, A., Ketcha, M., Reaungamornrat, S., Kleinszig, G., Vogt, S., Aygun, N., Lo, S., Wolinsky, J., and Siewerdsen, J. (2016). 3d–2d image registration for target localization in spine surgery: investigation of similarity metrics providing robustness to content mismatch. *Physics in Medicine & Biology*, 61(8):3009.

Deng, F., Tie, C., Zeng, Y., Shi, Y., Wu, H., Wu, Y., Liang, D., Liu, X., Zheng, H., Zhang, X., et al. (2021). Correcting motion artifacts in coronary computed tomography angiography images using a dual-zone cycle generative adversarial network. *Journal of X-Ray Science and Technology*, 29(4):577–595.

Donaldson, K., O'Connor, S., and Heath, N. (2013). Dental cone beam ct image quality possibly reduced by patient movement. *Dentomaxillofacial Radiology*, 42(2):91866873.

Ens, S., Ulrici, J., Hell, E., and Buzug, T. M. (2010). Automatic detection of patient motion in cone-beam computed tomography. In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1257–1260. IEEE.

Feldkamp, L. A., Davis, L. C., and Kress, J. W. (1984). Practical cone-beam algorithm. *Josa a*, 1(6):612–619.

Fu, T., Wang, Y., Zhang, K., Zhang, J., Wang, S., Huang, W., Yao, C., Zhou, C., and Yuan, Q. (2023). Deep-learning-based ring artifact correction for tomographic reconstruction. *Journal of Synchrotron Radiation*, 30(3).

Gallastegui, A. (2022). Computed tomography (ct). In *Comparative Veterinary Anatomy*, pages 47–56. Elsevier.

Gupta, K., Colvert, B., Chen, Z., and Contijoch, F. (2023). Difir-ct: Distance field representation to resolve motion artifacts in computed tomography. *Medical Physics*, 50(3):1349–1366.

Gupta, R., Krull, W., Hecla, J., Cramer, A., Kenyon, S., Arzoumanian, Z., and Gendreau, K. (2021). Tomographic imaging system. US Patent 10,895,540.

Haacke, E. M. (1999). Magnetic resonance imaging: physical principles and sequence design. *(No Title)*.

Hahn, J., Bruder, H., Rohkohl, C., Allmendinger, T., Stierstorfer, K., Flohr, T., and Kachelrieß, M. (2017). Motion compensation in the region of the coronary arteries based on partial angle reconstructions from short-scan ct data. *Medical physics*, 44(11):5795–5813.

Hanzelka, T., Dusek, J., Ocasek, F., Kucera, J., Sedy, J., Benes, J., Pavlikova, G., and Foltan, R. (2013). Movement of the patient and the cone beam computed tomography scanner: objectives and possible solutions. *Oral surgery, oral medicine, oral pathology and oral radiology*, 116(6):769–773.

Hartley, R. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press.

Hartshorne, J. (2018). Essential guidelines for using cone beam computed tomography (cbct) in implant dentistry. part 1: Technical considerations. *International Dentistry - African Edition*, 8(3):54–72.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

Hein, D., Liappis, K., Eguizabal, A., and Persson, M. (2022). Deep learning ring artifact correction in photon-counting spectral ct with perceptual loss. In *7th International Conference on Image Formation in X-Ray Computed Tomography*, volume 12304, pages 615–621. SPIE.

Holmes, T. W., Schmidt, B., Flohr, T., Ulzheimer, S., Bluemke, D. A., and Pourmorteza, A. (2022). Beam hardening artifacts in computed tomography: comparison of photon-counting and energy-integrating detectors. In *Medical Imaging 2022: Physics of Medical Imaging*, page PC120310M. SPIE.

Hsieh, J. (2003). *Computed tomography: principles, design, artifacts, and recent advances*. SPIE press.

Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.

Huang, H., Siewerdsen, J., Lu, A., Hu, Y., Zbijewski, W., Unberath, M., Weiss, C., and Sisniega, A. (2023). Multi-stage adaptive spline autofocus (masa) with a learned metric for deformable motion compensation in interventional cone-beam ct. In *Medical Imaging 2023: Physics of Medical Imaging*, volume 12463, pages 222–227. SPIE.

Huang, H., Siewerdsen, J. H., Zbijewski, W., Weiss, C. R., Unberath, M., Ehtiati, T., and Sisniega, A. (2022). Reference-free learning-based similarity metric for motion compensation in cone-beam ct. *Physics in Medicine & Biology*, 67(12):125020.

Ibad, H. A., de Cesar Netto, C., Shakoor, D., Sisniega, A., Liu, S. Z., Siewerdsen, J. H., Carrino, J. A., Zbijewski, W., and Demehri, S. (2023). Computed tomography: state-of-the-art advancements in musculoskeletal imaging. *Investigative radiology*, 58(1):99–110.

Kairalla, S. A., Scuzzo, G., Triviño, T., Velasco, L., Lombardo, L., and Paranhos, L. R. (2014). Determining shapes and dimensions of dental arches for the use of straight-wire arches in lingual technique. *Dental press journal of orthodontics*, 19:116–122.

Kau, C., Richmond, S., and Palomo, J. (2010). Diagnostic and prognostic value of dental cone-beam ct. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 109(2):312–317.

Kim, J., Sun, T., Alcheikh, A., Kuncic, Z., Nuyts, J., and Fulton, R. (2016). Correction for human head motion in helical x-ray ct. *Physics in Medicine & Biology*, 61(4):1416.

Ko, Y., Moon, S., Baek, J., and Shim, H. (2021). Rigid and non-rigid motion artifact reduction in x-ray ct using attention module. *Medical Image Analysis*, 67:101883.

Kumar, V., Ludlow, J. B., Mol, A., and Cevidanes, L. (2007). Comparison of conventional and cone beam ct synthesized cephalograms. *Dentomaxillofacial Radiology*, 36(5):263–269.

Lee, C.-L., Park, J., Nam, S., Choi, J., Choi, Y., Lee, S., Lee, K.-Y., and Cho, M. (2021). Metal artifact reduction and tumor detection using photon-counting multi-energy computed tomography. *Plos one*, 16(3):e0247355.

Lell, M. and Kachelrieß, M. (2023). Computed tomography 2.0: new detector technology, ai, and other developments. *Investigative Radiology*, 58(8):587–601.

Leung, E. M. Y., Yang, Y., Khambay, B., Wong, R. W. K., McGrath, C., and Gu, M. (2018). A comparative analysis of tooth size discrepancy between male and female subjects presenting with a class i malocclusion. *The Scientific World Journal*, 2018.

Loubele, M., Maes, F., Schutyser, F., and et al. (2006). Assessment of bone segmentation quality of cone-beam ct versus multislice spiral ct: a pilot study. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 102(2):225–234.

Ludlow, J. B. and Ivanovic, M. (2008). Comparative dosimetry of dental cbct devices and 64-slice ct for oral and maxillofacial radiology. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 106(1):106–114.

Mansour, H., Kadu, A., Boufounos, P., and Liu, D. (2021). Tomographic imaging system. US Patent 11,204,317.

Maur, S., Stsepankou, D., and Hesser, J. (2016). Ct auto-calibration by consistent contours. In *2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD)*, pages 1–7. IEEE.

Maur, S., Stsepankou, D., and Hesser, J. (2018). Auto-calibration by locally consistent contours for dental cbct. *Physics in Medicine & Biology*, 63(21):215018.

Maur, S., Stsepankou, D., and Hesser, J. (2019). Cbct auto-calibration by contour registration. In *Medical Imaging 2019: Physics of Medical Imaging*, volume 10948, pages 413–421. SPIE.

Motro, M., Ribeiro, A. A., Campos, P. S. F., and et al. (2019). Application of cbct in the diagnosis and treatment of temporomandibular disorders: a systematic review. *Journal of Applied Oral Science*, 27:e20180091.

Murata, K. and Ogawa, K. (2020). Ring-artifact correction with total-variation regularization for material images in photon-counting ct. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 5(4):568–577.

Nardi, C., Molteni, R., Lorini, C., Taliani, G. G., Matteuzzi, B., Mazzoni, E., and Colagrande, S. (2016). Motion artefacts in cone beam ct: an in vitro study about the effects on the images. *The British journal of radiology*, 89(1058):20150687.

Nardi, C., Taliani, G. G., Castellani, A., De Falco, L., Selvi, V., and Calistri, L. (2017). Repetition of examination due to motion artifacts in horizontal cone beam ct: comparison among three different kinds of head support. *Journal of International Society of Preventive & Community Dentistry*, 7(4):208.

Nemtoi, A., Czink, C., Haba, D., and Gahleitner, A. (2013). Cone beam ct: a current overview of devices. *Dentomaxillofacial Radiology*, 42(8):20120443.

Nickenig, H.-J., Eitner, S., Rellecke, C., and et al. (2006). Visualization of mandibular and maxillary fractures by cone beam ct in comparison with conventional multi-slice spiral ct. *Dentomaxillofacial Radiology*, 35(3):152–157.

Niebler, S., Schömer, E., Tjaden, H., Schwanecke, U., and Schulze, R. (2019). Projection-based improvement of 3d reconstructions from motion-impaired dental cone beam ct data. *Medical physics*, 46(10):4470–4480.

Osipov, S. P., Chakhlov, S. V., Zhvyrblia, V. Y., Sednev, D. A., Osipov, O. S., and Usachev, E. Y. (2023). The nature of metal artifacts in x-ray computed tomography and their reduction by optimization of tomography systems parameters. *Applied Sciences*, 13(4):2666.

Ouadah, S., Jacobson, M., Stayman, J. W., Ehtiati, T., Weiss, C., and Siewerdsen, J. H. (2017). Correction of patient motion in cone-beam ct using 3d–2d registration. *Physics in Medicine & Biology*, 62(23):8813.

Patel, S., Dawood, A., Whaites, E., and Pitt Ford, T. (2009). New dimensions in endodontic imaging: part 2. cone beam computed tomography. *International Endodontic Journal*, 42(6):463–475.

Patel, S., Durack, C., Abella, F., Shemesh, H., Roig, M., and Lemberg, K. (2015). Cone beam computed tomography in e ndodontics–a review. *International endodontic journal*, 48(1):3–15.

Saad, M. S. and Moore, S. (2022). Chapter.17 - computed tomography scan. In Hughes, G., editor, *A Medication Guide to Internal Medicine Tests and Procedures*, pages 78–84. Elsevier, Philadelphia.

Šalplachta, J., Zikmund, T., Zemek, M., Břínek, A., Takeda, Y., Omote, K., and Kaiser, J. (2021). Complete ring artifacts reduction procedure for lab-based x-ray nano ct systems. *Sensors*, 21(1):238.

Scarfe, W. C. and Farman, A. G. (2008). What is cone-beam ct and how does it work? *Dental Clinics of North America*, 52(4):707–730, v.

Scarfe, W. C., Farman, A. G., and Sukovic, P. (2006). Clinical applications of cone-beam computed tomography in dental practice. *Journal of the Canadian Dental Association*, 72(1):75–80.

Schofield, R., King, L., Tayal, U., Castellano, I., Stirrup, J., Pontana, F., Earls, J., and Nicol, E. (2020). Image reconstruction: Part 1–understanding filtered back projection, noise and image acquisition. *Journal of cardiovascular computed tomography*, 14(3):219–225.

Sharma, S., Kaushal, A., Patel, S., Kumar, V., Prakash, M., and Mandeep, D. (2021). Methods to address metal artifacts in post-processed ct images–a do-it-yourself guide for orthopedic surgeons. *Journal of Clinical Orthopaedics and Trauma*, 20:101493.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Sinha, A., Mishra, A., Srivastava, S., Sinha, P. M., and Chaurasia, A. (2016). Understanding artifacts in cone beam computed tomography. *International Journal of Maxillofacial Imaging*, 2(2):51–54.

Sisniega, A., Stayman, J., Cao, Q., Yorkston, J., Siewerdsen, J., and Zbijewski, W. (2016). Image-based motion compensation for high-resolution extremities cone-beam ct. In *Medical Imaging 2016: Physics of Medical Imaging*, volume 9783, pages 156–162. SPIE.

Sisniega, A., Stayman, J. W., Yorkston, J., Siewerdsen, J., and Zbijewski, W. (2017). Motion compensation in extremity cone-beam ct using a penalized image sharpness criterion. *Physics in Medicine & Biology*, 62(9):3712.

Sisniega, A., Thawait, G. K., Shakoor, D., Siewerdsen, J. H., Demehri, S., and Zbijewski, W. (2019). Motion compensation in extremity cone-beam computed tomography. *Skeletal radiology*, 48:1999–2007.

Spin-Neto, R., Matzen, L. H., Schropp, L., Gotfredsen, E., and Wenzel, A. (2015). Factors affecting patient movement and re-exposure in cone beam computed tomography examination. *Oral surgery, oral medicine, oral pathology and oral radiology*, 119(5):572–578.

Spin-Neto, R., Matzen, L. H., Schropp, L., Gotfredsen, E., and Wenzel, A. (2016). Movement characteristics in young patients and the impact on cbct image quality. *Dentomaxillofacial Radiology*, 45(4):20150426.

Spin-Neto, R., Matzen, L. H., Schropp, L. W., Sørensen, T. S., and Wenzel, A. (2018). An ex vivo study of automated motion artefact correction and the impact on cone beam ct image quality and interpretability. *Dentomaxillofacial Radiology*, 47(5):20180013.

Spin-Neto, R., Mudrak, J., Matzen, L., Christensen, J., Gotfredsen, E., and Wenzel, A. (2013). Cone beam ct image artefacts related to head motion simulated by a robot skull: visual characteristics and impact on image quality. *Dentomaxillofacial Radiology*, 42(2):32310645.

Su, J., Wang, C., Li, Y., Shang, K., and Liang, D. (2023). Retinexflow for ct metal artifact reduction. *arXiv preprint arXiv:2306.10520*.

Sun, T., Jacobs, R., Pauwels, R., Tijskens, E., Fulton, R., and Nuyts, J. (2021). A motion correction approach for oral and maxillofacial cone-beam ct imaging. *Physics in Medicine & Biology*, 66(12):125008.

Sun, X., Huang, F., Lai, G., Yu, D., Zhang, B., Guo, B., and Ma, Z. (2020). Motion artifacts detection from computed tomography images. In *International Conference on Advanced Data Mining and Applications*, pages 347–359. Springer.

Tan, M. and Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR.

Tan, M. and Le, Q. (2021). Efficientnetv2: Smaller models and faster training. In *International Conference on Machine Learning*, pages 10096–10106. PMLR.

Tyndall, D. A., Price, J. B., Tetradis, S., Ganz, S. D., Hildebolt, C., and Scarfe, W. C. (2012). Position statement of the american academy of oral and maxillofacial radiology on selection criteria for the use of radiology in dental implantology with emphasis on cone beam computed tomography. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, 113(6):817–826.

Vijayan, R. C., Sheth, N., Mekki, L., Lu, A., Uneri, A., Sisniega, A., Maggaragia, J., Kleinszig, G., Vogt, S., Thiboutot, J., et al. (2022). Motion-compensated targeting in pulmonary interventions using cone-beam ct and locally rigid/globally deformable 3d-2d registration. In *Medical Imaging 2022: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 12034, pages 622–627. SPIE.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612.

Welch, M. L., McIntosh, C., Purdie, T. G., Wee, L., Traverso, A., Dekker, A., Haibe-Kains, B., and Jaffray, D. A. (2020). Automatic classification of dental artifact status for efficient image veracity checks: effects of image resolution and convolutional neural network depth. *Physics in Medicine & Biology*, 65(1):015005.

Wicklein, J., Kyriakou, Y., Kalender, W. A., and Kunze, H. (2013). An online motion- and misalignment-correction method for medical flat-detector ct. In *Medical Imaging 2013: Physics of Medical Imaging*, volume 8668, pages 466–472. SPIE.

Würfl, T., Hoffmann, M., Aichert, A., Maier, A. K., Maaß, N., and Dennerlein, F. (2019). Calibration-free beam hardening reduction in x-ray cbct using the epipolar consistency condition and physical constraints. *Medical physics*, 46(12):e810–e822.

Yang, Q., Fullagar, W., Myers, G., Latham, S., Varslot, T., Sheppard, A., and Kingston, A. (2020). X-ray attenuation models to account for beam hardening in computed tomography. *Applied Optics*, 59(29):9126–9136.

Yıldızer Keriş, E. (2017). Effect of patient anxiety on image motion artefacts in cbct. *BMC oral health*, 17(1):1–9.

Zeng, S., Luo, Q., Xiong, H., Guo, W., and Lv, X. (2022). Tomographic imaging method. US Patent 11,346,782.

Zhang, J., He, B., Yang, Z., and Kang, W. (2023). A novel reconstruction of the sparse-view cbct algorithm for correcting artifacts and reducing noise. *Mathematics*, 11(9):2127.

Zhang, Y., Hu, P., He, Y., Yu, H., Tan, H., Liu, G., Gu, J., and Shi, H. (2022). Ultrafast 30-s total-body pet/ct scan: a preliminary study. *European Journal of Nuclear Medicine and Molecular Imaging*, 49(8):2504–2513.