Design and Optimization of a Dedicated Cone-Beam CT System for Musculoskeletal Extremities Imaging

W. Zbijewski\textsuperscript{a}, P. De Jean\textsuperscript{a}, P. Prakash\textsuperscript{a}, Y. Ding\textsuperscript{a}, J. W. Stayman\textsuperscript{a}, N. Packard\textsuperscript{b}, R. Senn\textsuperscript{b}, D. Yang\textsuperscript{b}, J. Yorkston\textsuperscript{b}, A. Machado\textsuperscript{c}, J. A. Carrino\textsuperscript{c}, J. H. Siewerdsen\textsuperscript{a}\textsuperscript{*}

\textsuperscript{a}Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD USA 21205
\textsuperscript{b}Carestream Health, Rochester, NY USA 14615
\textsuperscript{c}Department of Radiology, Johns Hopkins University, Baltimore, MD USA 21287

ABSTRACT

The design, initial imaging performance, and model-based optimization of a dedicated cone-beam CT (CBCT) scanner for musculoskeletal extremities is presented. The system offers a compact scanner that complements conventional CT and MR by providing sub-mm isotropic spatial resolution, the ability to image weight-bearing extremities, and the capability for integrated real-time fluoroscopy and digital radiography. The scanner employs a flat-panel detector and a fixed anode x-ray source and has a field of view of \(\sim (20\times20\times20)\) cm\(^3\). The gantry allows a “standing” configuration for imaging of weight-bearing lower extremities and a “sitting” configuration for imaging of upper extremities and unloaded lower extremities. Cascaded systems analysis guided the selection of x-ray technique (e.g., kVp, filtration, and dose) and system design (e.g., magnification factor), yielding input-quantum-limited performance at detector signal of 100 times the electronic noise, while maintaining patient dose below 5 mGy (a factor of \(\sim 2-3\) less than conventional CT). A magnification of 1.3 optimized tradeoffs between source and detector blur for a 0.5 mm focal spot. A custom antiscatter grid demonstrated significant reduction of artifacts without loss of contrast-to-noise ratio or increase in dose. Image quality in cadaveric specimens was assessed on a CBCT bench, demonstrating exquisite bone detail, visualization of intra-articular morphology, and soft-tissue visibility approaching that of diagnostic CT. The capability to image loaded extremities and conduct multi-modality CBCT/fluoroscopy with improved workflow compared to whole-body CT could be of value in a broad spectrum of applications, including orthopaedics, rheumatology, surgical planning, and treatment assessment. A clinical prototype has been constructed for deployment in pilot study trials.

Keywords: musculoskeletal radiology, orthopaedics, rheumatology, extremities imaging, cascaded system analysis, system optimization, cone-beam CT, flat-panel detector

1. INTRODUCTION

X-ray CT and MRI provide powerful, complementary diagnostic modalities for imaging of musculoskeletal (MSK) extremities – the former offering high spatial resolution and the latter exquisite soft-tissue visibility. While each of these tools has numerous strengths, a variety of limitations can be identified, including: (i) difficulty in examining weight-bearing extremities (e.g., tissue impingements that manifest only in loaded or tensioned extremities); (ii) diagnostic challenges in the presence of prosthetic implants (e.g., total knee replacement), where metal artifacts can hamper the visualization of implant loosening and bone disintegration; (iii) potentially large cumulative radiation dose in longitudinal studies (e.g., analysis of fracture healing or response to therapy); and (iv) cost, space, and workflow associated with whole-body scanners used for extremities imaging. One example of this last challenge is the inability to use the same system for tomographic imaging, traditional, planar x-ray radiography or fluoroscopy (which remain standard, valuable exams for many MSK diagnostic tasks). Addressing these challenges could benefit diagnosis, planning, and therapy evaluation in applications such as orthopaedic trauma, cancer, and inflammatory disease (e.g., arthritis).

A dedicated imaging system is under development to address these limitations and provide complementary or alternative functionality to CT and MRI. The system is based upon x-ray cone-beam CT (CBCT) using a high-performance flat-panel detector, a paradigm that has demonstrated success in applications ranging from breast [1] and maxillofacial [2,3] imaging to image-guided interventions [4,5]. The compact form of flat-panel detectors was leveraged to design a geometry that is optimized for imaging both unloaded and weight-bearing extremities. Moreover, using a flat-panel detector yielded a system that combines digital radiography and real-time fluoroscopy with 3D
imaging and could potentially be deployed in-office with site requirements, shielding, and footprint conservative in comparison to conventional CT. The system was designed using an iterative process that considered theoretical modeling using cascaded systems analysis [6,7] (CSA) to optimize x-ray technique selection and critical aspects of scanner geometry. This paper summarizes scanner design, details the model-based optimization process, and reports quantitative and qualitative studies of imaging performance. We also discuss some of the task-specific reconstruction and processing algorithms and acquisition protocols currently under development for this platform.

2. METHODS

2.1 Scanner design.
The design of the extremities scanner is illustrated in Figure 1 with images of a clinical prototype and a CAD rendering of the internal components of the system. The gantry allows scanning in both the standing configuration (Fig. 1A) for imaging of weight-bearing lower extremities, and in the sitting configuration (Fig. 1B) for imaging of upper extremities (with the capability to apply tension) and unloaded lower extremities. The gantry is self-shielded, facilitating compliance with x-ray safety site requirements. The patient enters the bore through a sliding door on the side of the gantry. The inner bore is ~20 cm and incorporates accessories for patient immobilization (e.g., inflatable air bladder). Inside the gantry, a flat-panel detector and fixed-anode source are mounted on a sickle-arm (Fig. 1C) to provide a side opening. The source-detector distance is ~55 cm, the source-isocenter distance is ~42 cm. Full gantry rotation is 240°, which exceeds minimum requirements for fully 3D reconstruction from a short-scan range for this geometry. The volume of reconstruction is ~(20x20x20) cm³. A Varian 3030+ flat panel detector is used (1536x1536 pixels at 0.194 pitch) in full- or half-resolution mode. Dual and dynamic gain readout modes are available. The detector is fitted with a custom 10:1 antiscatter grid. The fixed anode, 0.5 mm focal spot x-ray source can deliver x-ray techniques up to 125 kVp, 7 mA (~0.8 kW).

2.2 Cascaded system model of the imaging chain.
Cascaded system analysis was employed to guide the selection of system components by providing a mechanism for estimating modulation transfer function (MTF), detective quantum efficiency (DQE), and detectability index (d’) of bone and soft-tissue imaging tasks in CBCT reconstructions for various imaging conditions. As described in [7], a complete CSA model of a flat-panel cone-beam CT consists of a 2D cascade, modeling the interactions of x-rays in the converter and sampling and readout of the detector with additive noise, and a 3D reconstruction cascade, involving processing steps from log-transform of the projections to discrete sampling of the 3D volume.

The 2D detector cascade was used to study the influence of focal spot size and magnification factor on system resolution and to investigate DQE as a function of x-ray technique and imaging dose. The system MTF ($MTF_{\text{system}}$) was obtained as a product of presampling detector MTF and the MTF due to focal spot blur. The latter was modeled as a simple Gaussian distribution with characteristic width (focal spot diameter) $a_{\text{spot}}$ [8,9]. $MTF_{\text{system}}$ was computed at the imaging plane (located at the isocenter) over a range of geometric magnifications $M$ and focal spot sizes for a fixed detector pixel size of 0.388 mm (half-resolution read-out) and a 90 kVp beam (0.3 mm Cu and 4 mm Al added filtration).

The estimation of projection domain DQE also assumed 0.388 mm detector pixels (2x2 binning of the detector elements). A range of x-ray techniques (defined by beam kVp and added filtration) was tested. The input spectrum was
attenuated by a model object (simulated knee) prior to entering the detector cascade. The simulated knee consisted of a 16 cm water cylinder containing an 8 cm diameter inner cylinder emulating the bone (femoral condyles). The bone model included a 2 mm thick outer shell made of cortical bone and an inner shell made of trabecular bone. For each x-ray technique, the DQE, patient dose and required x-ray power were computed for a range of exposures, each of them selected to provide a given detector signal behind the simulated knee. The patient dose was calculated at the center of 16 cm diameter water cylinder, assuming a scan consisting of 480 projections. The presence of an antiscatter grid was taken into account by simple scaling of the estimated dose and power by the inverse of the transmission factor (measured at ~0.7). The x-ray power estimation assumed 30 ms pulses.

The task-based detectability index ($d'$) was computed using the complete 3D CSA model to provide initial investigation of the dependence of imaging performance on dose with respect to idealized imaging tasks pertinent to MSK imaging. The calculations included the filtered backprojection cascade [7], anatomical background clutter [10], and simple models of the imaging task and observer, as prescribed in ICRU 54 [11]. The simplest such model is that of the prewhitening observer, for which the “generalized” detectability, which includes anatomical background clutter [10,12], is given by:

$$d'^2 = \frac{\left[MTF_{recon}(f_x,f_y,f_z)W_{task}(f_x,f_y,f_z)\right]^2}{MTF_{recon}(f_x,f_y,f_z)^2S_B(f_x,f_y,f_z) + S_Q(f_x,f_y,f_z) + S_E(f_x,f_y,f_z)},$$

(1)

where $f_x,f_y,f_z$ are the 3D frequency coordinates. The imaging task is represented by a task function, $W_{task}(f_x,f_y,f_z)$, given by the difference of the Fourier transform of two hypotheses (e.g., in a detection task, signal-present and signal-absent hypotheses). $MTF_{recon}(f_x,f_y,f_z)$ is the reconstruction MTF, $S_B(f_x,f_y,f_z)$ and $S_Q(f_x,f_y,f_z)$ are the reconstruction-domain power spectra of the quantum and electronic noise, respectively. All these quantities are obtained using the 3D CSA model. $S_B(f_x,f_y,f_z)$ is the power spectrum associated with anatomical background noise. In other applications [10,13], $S_B$ has been widely modeled simply as a power-law characteristic:

$$S_B(f) = \frac{\kappa}{(af)^\beta},$$

(2)

where $\beta$ is the degree of correlation (taken equal to 3.0 for 3D data, consistent with previous work [10,13]). The magnitude of the background power spectrum, $\kappa$, was scaled from measurements of Gang et al. [10] according to the contrast (squared) of muscle and fat, approximating 3D clutter in soft-tissue visualization tasks of the knee or hand. The scale factor $a$ is 1 mm$^{-1}$ and yields a dimensionless denominator.

The detectability index was computed over a dose range up to 10 mGy for two imaging tasks: (i) a bone feature detection task consisting of a 0.3 mm cube in a uniform water background; and (ii) a soft-tissue detection task involving a Gaussian object (2.0 mm full-width-at-half-maximum) of muscle tissue in a water background. The calculations assumed a 90 kVp beam (0.3 mm Cu and 4 mm Al added filtration), 480 total projections, a Hann reconstruction filter, and bi-linear interpolation of the projection data for voxel-driven reconstruction. The detector pixel size was varied from 0.194 mm (full resolution read-out of the Varian 3030+ panel) to 0.388 mm (half-resolution read-out), and the reconstruction voxel size was matched to the detector pixel size (divided by the magnification) for all cases.

### 2.3 Test-bench.

Further optimization and investigation of image quality was performed using an experimental CBCT bench. The bench consisted of a rotary object table, Varian 4030CB flat-panel detector (having the same pixel architecture as the panel used in the extremities scanner) and a rotating anode source. The source, detector, and rotary table were mounted on translation stages that allow for a variety of imaging geometries. For the experiments described here, the bench was set to emulate the geometry of the extremities scanner. Unless mentioned otherwise, all studies used a 90 kVp x-ray beam (0.3 mm Cu and 4 mm Al added filtration), in accordance with the results of the cascaded systems analysis; half-resolution (0.388 mm) detector readout was employed. Unless mentioned otherwise, the acquisitions involved a 240° scan (the same scanning arc as in the proposed system) with 480 projections at 0.1 mAs per projection. The custom 10:1 anti-scatter grid was mounted on the panel. The images were reconstructed using the Feldkamp algorithm with extended Parker weights [14], Hann apodization filter, and a voxel size of 0.5 mm.

### 2.4 Dose measurements.

The imaging dose was measured on the imaging bench using a RadCal Accu-Pro 9096 multi-purpose radiation meter (Radcal, Monrovia CA) with a high dose-rate, 6 cc active volume ionization chamber. The bench was in the geometry emulating the proposed extremities scanner. The dosimeter was placed at the isocenter of the system, in the center of a
16 cm CTDI phantom (Radcal model 20CT6). This setup, in which a long phantom is used in concert with a Farmer chamber, is consistent with the approach proposed for volume CT dosimetry in AAPM Task Group 111 report. Measurements were obtained for a range of kVp values (80, 90, and 110 kVp) and tube output levels (0.1 – 0.5 mAs per frame). The additional beam filtration was fixed at 4 mm Al + 0.3 mm Cu. For each (kVp, mAs) pair, 30-40 exposures were collected and averaged to yield the dose per projection. The total dose per scan was then computed by simple scaling by the number of exposures in a typical scan, equal to 480. The effective dose can be computed from the absolute dose by scaling it by the scan length of 20 cm and by the organ-specific weighting factor (equal to 0.0005 mSv/mGy·cm for distal extremities [16]).

2.5 Antiscatter grid.

The performance of a custom, focused 10:1 antiscatter grid (200 lines/inch) was evaluated on the test-bench. A 16 cm diameter, 30 cm long plastic cylinder phantom with plugs of various tissue-mimicking materials was imaged with and without the grid at various imaging doses (5 mGy – 15 mGy per scan). Signal-difference to noise ratio (SDNR) was computed for one of the plugs to assess the influence of the trade-off between scatter rejection and reduction in the primary radiation caused by the grid.

2.6 Benchtop cadaver studies.

Qualitative assessment of imaging performance was performed in fresh cadaver specimens (hand and knee) on the test-bench. The nominal imaging protocol described in Section 2.4 was used in this study.

3. RESULTS

Figure 2 illustrates the trade-offs between focal spot size, magnification factor, and system resolution. A summary metric is provided by $f_{50}$ (the spatial frequency at which MTF drops by 50%), displayed as a function of magnification factor and focal spot size in Figure 2A. Once the magnification increases above a certain threshold (whose value depends on $a_{spot}$), $MTF_{system}$ becomes dominated by focal spot blur. As a result, the system resolution begins to decrease as a function of magnification. For a focal spot size of 0.5 mm (typically the smallest size available for fixed anode sources with sufficient power for the extremities scanner), the transition occurs at a magnification of ~1.3 and initially leads to a broad plateau of optimal magnification. This is further illustrated in Figure 2B. The $MTF_{system}$ for $a_{spot}=0.5$ mm is only slightly broadened by the increase in $M$ from 1.3 to 1.6, whereas a significant broadening over the case of $M=1$ is apparent. Based on this analysis and on the mechanical constraints of maintaining a natural stance during scanning in a standing configuration, a magnification of 1.3 was selected for the system under construction.

![Figure 2. (A) Frequency at which the MTF$_{system}$ drops by 50% ($f_{50}$) as a function of focal spot size and magnification factor. (B) Complete system MTF as a function of the magnification factor M for a fixed focal spot size of 0.5 mm. For this focal spot size (same as in the source used in the prototype system), the trade-off between the focal spot blur and detector blur leads to a broad region of optimal MTF$_{system}$ in the region of magnification factors spanning from ~1.3 to ~2.](image-url)

Figure 3 illustrates the results of the DQE study. Figure 3A displays a surface plot of zero-frequency DQE as a function of x-ray energy and added filtration for a fixed signal level of 100x above the electronic noise floor. The dependence of imaging dose and required source power on the added filtration and beam kVp are illustrated for the same signal conditions in Figures 3B and Figure 3C, respectively. The DQE increases as the x-ray energy decreases – at a cost, however, of increased patient dose and x-ray power required to achieve the desired signal level. Maintaining the radiation dose below 5 mGy (~1/3 of a conventional CT knee exam [17]) and x-ray power below 0.5 kW (typical for the...
fixed anode sources) suggests an x-ray technique of 90 kVp + 0.3 mm Cu. Figure 3D plots the DQE(0) as a function of the desired signal level for added filtration of 0.3 mm Cu and various kVp values. The DQE(0) becomes largely independent of the signal level (and thus the imaging dose) for signal levels ~50-80x above the noise floor, indicating quantum-limited operation. For a 90 kVp beam, this transition to a quantum-limited regime occurs at approx. 1.5-2 mGy.

Figure 3. (A) Zero-frequency DQE as a function of kVp and added filtration for detector signal level of 100x electronic noise floor. (B) Dose as a function of added Cu filtration for three kVp values and a signal level of 100x noise floor. (C) Source power required to achieve a signal level of 100x noise floor as a function of added Cu filtration and kVp. X-ray technique of approx 90 kVp+0.3 mm Cu allows to maintain the patient dose below 5 mGy and source power output at ~0.5 kW. (D) DQE(0) as a function of detector signal level for a range of kVp values and added filtration fixed at 0.3 mm Cu. For a 90 kVp beam, the system operation becomes quantum limited at signal levels of ~50x noise floor.

Figure 4 summarizes the detectability index as a function of dose. The high-contrast bony detail task maintains high detectability for relatively low dose ($d' > 1$ for dose $> 0.3$ mGy), and performance can be improved by judicious selection of detector pixel size and reconstruction voxel size. At low dose (below 2 mGy per scan), it is advantageous to use larger pixels (corresponding to 2x2 detector binning for the Varian 3030+ panel) in order to reduce image noise. For dose above ~2-3 mGy, quantum and electronic noise become less of a concern for imaging high-contrast details, and the detectability can be improved by acquiring at full resolution (no binning). The situation is different for the soft-tissue task, where the detectability is only marginally affected by detector pixel size across the range considered. Imaging dose plays a much more significant role in the discrimination of soft-tissue details. Detectability slowly increases with dose, exceeding a level of $d' \sim 1$ at approximately ~2-3 mGy. The results of this study indicate that moving to a full-resolution panel read-out will be advantageous for diagnostic tasks that involve precise assessment of bone morphology, such as in the diagnosis and staging of arthritis. Further experimental optimization of such high resolution imaging protocols is ongoing.
Figure 4. Task-based detectability index computed as a function of dose. The two imaging tasks modeled are a high-contrast bone detail detection task (a 0.3 mm Calcium cube; top two curves) and a low-contrast soft tissue detection task (a 2 mm Gaussian of muscle tissue in water background; bottom two curves). Calculations were performed as a function of detector pixel size and dose for a scan of 480 projections. Acquisition at full resolution (0.194 mm pixel size) improves the detectability for the bone task but has little effect on soft-tissue task. Soft-tissue detectability increases gradually with dose, suggesting $d' > 1$ for dose per scan above ~2-3 mGy.

The results of the study of imaging dose as a function of tube output for various x-ray techniques (always including added filtration of 0.3 mm Cu and 4 mm Al) are presented in Table 1. For tube voltage of 90 kVp, measurements in the projections of a cadaveric knee indicate that signal level of ~100x noise floor is attained for tube output of 0.1 mAs (the imaging protocol chosen for the cadaver studies described below). According to the data presented in Table 1, this corresponds to a dose of 6.4 mGy. The theoretical dose estimate obtained with the CSA model for this signal level and tube voltage was 4-5 mGy, with the discrepancies likely due to inaccuracies in the modeling of the x-ray source and x-ray scatter.

Table 1. Imaging dose for a scan of 480 projections for various values of kVp and tube output (mAs) per projection.

<table>
<thead>
<tr>
<th>kVp</th>
<th>0.1 mAs</th>
<th>0.16 mAs</th>
<th>0.2 mAs</th>
<th>0.25 mAs</th>
<th>0.4 mAs</th>
<th>0.5 mAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 kVp</td>
<td>mGy</td>
<td>4.02</td>
<td>6.12</td>
<td>7.52</td>
<td>9.27</td>
<td>14.27</td>
</tr>
<tr>
<td>90 kVp</td>
<td>mGy</td>
<td>6.38</td>
<td>9.78</td>
<td>12.26</td>
<td>15.20</td>
<td>22.84</td>
</tr>
<tr>
<td>110 kVp</td>
<td>mGy</td>
<td>13.59</td>
<td>20.14</td>
<td>24.68</td>
<td>30.05</td>
<td>45.51</td>
</tr>
</tbody>
</table>

The tight geometry of the extremities scanner leads to relatively high scatter to primary ratios for larger objects, such as the knee. The use of antiscatter grids to combat the resulting artifacts has been evaluated with a custom 10:1 grid. The magnitude of the cupping artifact was reduced by 50% when employing the grid. Moreover, the SDNR was found to improve with the antiscatter grid over a gridless acquisition across the range of imaging dose considered, indicating that the decrease in signal-to-noise ratio caused by the loss of primary radiation is offset by the rejection of scattered x-rays. Example images of cadaveric extremities are shown in Fig. 5. High isotropic spatial resolution is apparent in the visualization of fine trabecular details. Assessment by expert MSK radiologists suggests that the delineation of fat, muscle, and tendons is comparable to conventional CT. Total effective dose for these scans was 0.064 mSv (see Table 1), which is approximately 1/2 - 1/3 of that reported for diagnostic CT of the knee. The scanning protocol used here (0.1 mAs/projection) delivers signal levels similar to those assumed in the theoretical model of Section 2.2. Satisfactory soft-tissue contrast achieved in the cadaver study further corroborates the basic usefulness of CSA models in the optimization of imaging technique. Further improvement in resolution and contrast-to-noise ratio is expected through improved reconstruction and artifact correction techniques tailored to extremities imaging.
4. DISCUSSION

A novel cone-beam CT scanner for extremities imaging has been developed by combining task-specific mechanical design and optimization of x-ray technique and system geometry by means of cascaded system analysis. Results obtained in cadaver studies using x-ray techniques suggested by the cascaded systems model demonstrate that soft-tissue visibility can be achieved within the power limits of the scanner (~0.5 kW) and at low radiation dose (~1/3 the dose of a conventional CT knee exam [17]). Since the geometry of the scanner involves a very small air-gap, a custom focused antiscatter grid was manufactured for the prototype system. Initial evaluation of the grid demonstrated improved image quality at equivalent or reduced dose to the patient. Overall, the proposed system will deliver low dose, high resolution imaging of the extremities with soft tissue contrast approaching that of conventional CT. It complements existing imaging modalities and offers additional capabilities in imaging of weight-bearing extremities. It also offers potential advantages in cost and workflow relative to whole-body scanners, including the ability to perform digital radiography and dynamic real-time fluoroscopy on the same machine. In addition, the task-specific design facilitates the development of dedicated image acquisition and reconstruction techniques to optimize image quality. Ongoing developments involve advanced image acquisition techniques, such as dual-energy CBCT and ultra-high resolution imaging of bone morphology and joint spaces, as well as novel reconstruction techniques (viz., statistical, model-based reconstruction for correction of metal artifacts originating from prosthetic implants using prior knowledge of prosthetic shape to constrain the reconstruction). Other approaches currently under investigation include improved scatter correction, dual-energy cone-beam CT, and functional 3D/4D motion studies. The prototype scanner has been deployed at our institution for use in pilot studies targeting orthopedic imaging of the knee (including weight-bearing exams and total knee replacement) and the hand (including monitoring of fracture healing and arthritis therapy response).

REFERENCES
