Leveraging multi-layer imager detector design to improve low-dose performance for megavoltage cone-beam computed tomography

Yue-Houng Hu¹, Joerg Rottmann¹, Rony Fueglistaller², Marios Myronakis¹, Adam Wang³, Pascal Huber², Daniel Shedlock⁴, Daniel Morf², Paul Baturin³, Josh Star-Lack³, and Ross Berbeco¹

¹Department of Radiation Oncology, Division of Medical Physics and Biophysics, Brigham and Women’s Hospital, Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts 02115 ²Varian Medical Systems, CH-5405, Baden-Dättwil, Switzerland ³Varian Medical Systems, Palo Alto, California 94304-1030 ⁴Varex Imaging Corporation, Knoxville, Tennessee 37922-9435

Abstract

While megavoltage (MV) cone-beam computed tomography (CBCT) using an electronic portal imaging device (EPID) provides many advantages over kilovoltage (kV) CBCT, clinical adoption is limited by its high doses. Multi-layer imager (MLI) EPIDs increase DQE(0) while maintaining high resolution. However, even well-designed, high-performance MLIs suffer from increased electronic noise from each readout, degrading low-dose image quality. To improve low-dose performance, shift-and-bin addition (ShiBA) imaging is proposed, leveraging the unique architecture of the MLI. ShiBA combines hardware readout-binning and super-resolution concepts, reducing electronic noise while maintaining native image sampling. The imaging performance of full-resolution (FR); standard, aligned binned (BIN); and ShiBA images in terms of noise power spectrum (NPS), electronic NPS, modulation transfer function (MTF), and the ideal observer signal-to-noise ratio (SNR) - the detectability index (d′) – are compared. The FR 4-layer readout of the prototype MLI exhibits an electronic NPS magnitude 6-times higher than a state-of-the-art single layer (SLI) EPID. Although the MLI is built on the same readout platform as the SLI, with each layer exhibiting equivalent electronic noise, the multi-stage readout of the MLI results in electronic noise 50% higher than simple summation. Electronic noise is mitigated in both BIN and ShiBA imaging, reducing its total by ~12 times. ShiBA further reduces the NPS, effectively upsampling the image, resulting in a multiplication by a sinc² function. Normalized NPS (NNPS) show that neither ShiBA or BIN otherwise affects image noise. The LSF shows that ShiBA removes the pixilation artifact of BIN images and mitigates the effect of detector shift, but does not quantifiably improve the MTF. ShiBA provides a pre-sampled representation of the images, mitigating phase dependence. Hardware binning strategies lower the quantum noise floor, with 2×2 implementation reducing the dose at which DQE(0) degrades by 10% from 0.01 MU to 0.004 MU, representing 20% improvement in d′.
I. INTRODUCTION
Advantages of megavoltage (MV) cone-beam computed tomography (CBCT) include financial cost savings, translation between Hounsfield units (HU) and electron density, high-Z material suppression, and co-registration of treatment and imaging beams (Dipasquale et al., 2014; Groh et al., 2002; Morin et al., 2006). The main challenges to MVCBCT have been the increase in radiation dose and decrease in image quality, relative to kV CBCT with an on-board imaging (OBI) system. By stacking several imager (scintillation-readout) layers, a zero-frequency detective quantum efficiency (DQE) equivalent to or surpassing that of kV-CBCT may be achieved (Hu et al., 2017). While a multi-layer imager (MLI) may improve MV image quality at quantum noise limited doses (Myronakis et al., 2017; Rottmann et al., 2016), additional readout layers will also contribute electronic noise, inhibiting low-dose DQE.

To improve low-dose performance of the MLI, we propose an imaging technique called shift-and-bin addition (ShiBA), which leverages the architecture of the MLI by employing concepts of readout binning and shift-and-add super-resolution (Chen et al., 2017). Hardware readout-binning has been shown as an effective means of reducing electronic noise and readout time (Beutel et al., 2000; Srinivas and Wilson, 2004). Shift-and-add super-resolution, which increases sampling by physically offsetting each layer of an MLI by sub-pixel distances, has previously been shown to better represent high-frequency information, typically unresolvable by any single layer of the MLI or typical single-layer imager (SLI) detector (Chen et al., 2017). By offsetting individual MLI layers, shift-and-add super-resolution allows sampling rates greater than the pixel pitch. Although our previous studies have shown that current sampling rates (i.e. 0.336 mm) and resolution are unnecessary for typical clinical activities and tasks (Hu et al., 2017) - including those relevant for MV-CBCT - super-sampling, used in conjunction with detector binning as in the proposed ShiBA technique, may be used to improve low-dose performance while preserving native imager sampling rates.

II. METHODS
II.A. ShiBA imaging technique
The ShiBA technique combines detector binning and shift-and-add super-resolution. Since current implementation of the MLI hardware and software is not capable of these operations (illustrated in Fig. 1), ShiBA images were simulated by constructing composite projection images from the information from each individual MLI readout. The imaging potential and physical underpinnings of the ShiBA technique, shift-and-add super-resolution and detector readout-binning were examined. Results were compared to a conventional EPID (Varian AS-1200) equipped on a Varian TrueBeam linac. Further, to determine the effect of
hardware pixel binning readout in the ShiBA technique on image noise, noise power spectrum (NPS) was measured for quantum noise limited doses (i.e. > 0.5 MU at a source-to-imager distance of 150 cm) as well as for dark-frame images to analyze the electronic noise. Finally, composite projections of a SeeDOS QC-3 phantom was acquired to gain a qualitative understanding of the effect of ShiBA imaging. All images were acquired at a tube energy of 6 MV with source-to-axis distance (SAD) and source-to-imager distance (SID) were 100 and 150 cm, respectively.

Shift-and-add super-resolution was the subject of a prior simulation study and found that bar patterns as fine as 3.3 lp/mm were better resolved after sub-pixel layer shifts (Chen et al., 2017). To study the physical basis of this effect, super-resolution was incorporated post-readout. Each projection was subjected to a N-1 shift prior to either 2×2 or 4×4 software image binning, where N describes the imager layer number (i.e. N=1 is the layer closest to the tube head and N=4 is the bottom layer). The image binning results in sparser image sampling. In order to combine all shifted and binned layers, the data from each readout were upsampled to native resolution via nearest neighbor resampling prior to summation. To separate the effect of upsampling from standard binning (i.e. without upsampling), ShiBA images were compared with both standard 2×2 and 4×4 binned (BIN) images and 2×2 and 4×4 binned images with upsampling (upsampled BIN).

II.B. Analysis of image quality

II.B.1. Analysis of modulation transfer function—Modulation transfer function (MTF) was measured for projection images based on a slanted slit experimental setup. A 100 μm slit phantom composed of a pair of tungsten alloy blocks (~12 cm × 7.7 cm × 18.7 cm) was placed flush with the imager towards the center of its active area at an angle of 1.4 degrees and was imaged 300 times in order to reduce measurement noise. Each frame was acquired at the maximum dose at which the detector dynamic range was not saturated. The slit images were dark-field and flood field corrected and used to produce an oversampled line-spread function (LSF), the Fourier transform of which produces a phase-averaged MTF.

II.B.2. Analysis of noise power spectrum—NPS was measured for projection images by acquiring flat field images (i.e. without any obstruction in the beam path) of composite images. Images were acquired at a 40×40 cm field size and were dark field and flood field corrected. Each projection image was divided into roughly 325 128×128 pixel sub-regions-of-interest (sub-ROI). The projection-space NPS, \( S_p(f_x,f_y) \) was measured as the ensemble average of the NPS of each sub-ROI according to (Beutel et al., 2000; Zhao and Zhao, 2003; Zhao et al., 2009):

\[
S_p(f_x,f_y) = \frac{d_x d_y}{N_x N_y} \left\langle \left| \text{FT} [ I(x, y) - \bar{I}(x, y) ] \right|^2 \right\rangle ,
\]

where \( N \) denotes the number of pixels in the \( x \)- and \( y \)- directions (lateral and superior-inferior directions, respectively) for each sub-ROI, \( d \) is the pixel dimension (0.336 mm pixel size for a full resolution readout) and \( I \) is the image intensity. NPS and NNPS (the NPS divided by the square of the mean intensity of the flat field image) were measured for dark

\textit{Phys Med Biol. Author manuscript; available in PMC 2019 January 30.}
frame images, as well as doses of 0.5, 1.0, 1.5, 2.0, 2.5 MU. Dark frame images were acquired in pairs and subtracted prior to application of Eq. (1) to remove the effect of fixed-pattern noise. Since electronic noise is uncorrelated, the combination of the two dark frame images results in NPS double that of any single frame (Richard and Siewerdsen, 2008a). $S_p$ for the paired dark frame images were divided by two to determine the electronic noise contribution of any single exposure. For doses above 0.5 MU, the NPS were found to be linear with respect to tube output, confirming quantum noise limited performance.

The total NPS is defined as the sum of the electronic noise and the quantum noise of a particular image. Since quantum noise is linear as a function of exposure, total noise may be estimated after measurement of NPS at quantum noise limited exposures (Beutel et al., 2000). DQE(0) is calculated as the MTF(0) divided by NPS(0), multiplied by the number of input quanta at a given exposure.

II.B.3. Simulation of image reconstruction—Image reconstruction is modeled based on analytical reconstruction platforms such as filtered backprojection (FBP) (Richard and Siewerdsen, 2008a; Richard et al., 2005; Tward and Siewerdsen, 2008; Tward et al., 2008). Essentially, 3D resolution and noise are derived from planar measurements of projection space MTF and NPS as measured in section II.B.1 and II.B.2.

The effect of ramp ($H_{RA}$) and apodization ($H_{AP}$) filters is handled in Fourier space, where their frequency-response is multiplied with the projection space ($P$) MTF and NPS according to (Hu et al., 2008; Zhao and Zhao, 2008):

$$T_P(f_x, f_y) = T_P(f_x, f_y) H_{RA} H_{AP}$$
$$S_P(f_x, f_y) = S_P(f_x, f_y) H_{RA}^2 H_{AP}^2 , \quad (2)$$

where $F$ denotes the filtering stage of the model for the MTF ($T$) and NPS ($S$). $H_{RA}$ and $H_{AP}$ are defined in standard Fourier forms according to:

$$H_{RA} = \frac{|f_x|}{f_{Ny}}$$
$$H_{AP} = \frac{1}{2} (1 + \cos(2\pi f_x d_x)) , \quad (3)$$

where $f_{Ny}$ is the detector Nyquist frequency and $d_x$ is the readout pixel pitch (i.e. 0.336 mm for full resolution readout and 0.672 mm and 1.344 cm for 2×2 BIN and 4×4 BIN, respectively). Two-dimensional (2D) to three-dimensional (3D) conversion is based on the central-slice theorem (Hu et al., 2016; Hu and Zhao, 2011; Tward and Siewerdsen, 2008; Tward et al., 2008; Zhao and Zhao, 2008). Mathematically, for an acquisition of full angular sampling and $N$ views, the 3D NPS and MTF may be calculated using:
where $x'$, $y'$ and $z'$ represent coordinates in 3D reconstructed space for lateral, superior-inferior, and anterior-posterior directions, respectively and $\delta$ describes sampling along angle $\theta_i$. Conceptually, the central-slice theorem describes 3D transfer functions (NPS, MTF) as the mapping of each of the 2D projection transfer functions along its angle of acquisition with respect to gravity. As total scan angle and total projection numbers increases, frequency space is better and more completely sampled. In the present study, an acquisition of 400 views over 180 degrees of equiangular projections was modeled.

\[ T_B(f_x', f_y', f_z') = \frac{\sum_{i=1}^{N} T_B(f_x, f_y) \delta(f_x \sin \theta_i - f_x \cos \theta_i)}{T_B(0,0,0)} \]
\[ S_B(f_x', f_y', f_z') = \frac{\sum_{i=1}^{N} S_B(f_x, f_y) \delta(f_x \sin \theta_i - f_x \cos \theta_i)}{S_B(0,0,0)} \]

\( (4) \)

II.B.4. Ideal observer SNR: detectability index—To quantify the low-dose performance of binning strategies in comparison to full resolution readout of the MLI and a standard single-layer imager (SLI), the ideal-observer signal-to-noise ratio (SNR) was invoked. Its integration over all frequencies yields a single figure-of-merit (FOM), which is referred to as the detectability index ($d'$). In the present manuscript, $d'$ is defined and measured for in-plane (IP) structures of three-dimensional (3D) volumes, which is typical of how CBCT images are viewed. The SNR and $d'$ are defined as in previously published work (Gang et al., 2011; Gang et al., 2010; Hu et al., 2016; Hu and Zhao, 2011; Ikejimba et al., 2012; Richard and Siewerdsen, 2008a, b; Richard et al., 2008; Vennart):

\[ SNR_{IP}^2(f_x, f_z) = \frac{\int |task(f_x, f_y, f_z)| T(f_x, f_y, f_z) df_y}{\int S(f_x, f_y, f_z) df_y} \]

\( (5) \)

and

\[ d':_{IP} = \sqrt{\int \int SNR_{IP}^2(f_x, f_z) df_x df_z} \]

\( (6) \)

where $T$ is the reconstructed 3D MTF and $S$ the 3D NPS, the $z$-direction describes the anterior-posterior direction, and $task$ is a detection task function with contrast $K_C$, which compares the difference of signal present vs. signal absent according to:

\[ task(f_x, f_y, f_z) = O(f_x, f_y, f_z) - 0 \]

\( (7) \)

where $O$ is the object function: the Fourier definition of a shape such as a sphere (Desponds et al., 1991):
or a cylinder (Desponds et al., 1991):

\[
O_{\text{cyl}}(f_x, f_y) = \frac{2rL}{2f_x} J_1(2\pi r f_x) \frac{\sin(L f_y)}{L f_y},
\]  

where \( r \) is the radius of the sphere or cylinder, \( L \) is the length of the long-axis of the cylinder, and \( J_1 \) is a first-order Bessel function of the first kind.

A range of clinically relevant shapes were examined. To simulate the effect of fiducials and tumor masses, spherical objects with radii of 2 mm and 30 mm, respectively were calculated according to Eq. (8). The spine was simulated simply as a cylinder by applying Eq. (9) with a radius and length of 22.5 mm and 1000 mm.

### III. RESULTS

#### III.A. Projection space Fourier metrics

**III.A.1. Measurement of MTF**—A comparison of MTF for full resolution readout with 2x2 (a) and 4x4 (b) binning schemes (including BIN, upsampled BIN, and ShiBA) is plotted in Fig. 3. Due to the larger readout width of the binning schemes, the MTF is lower for all BIN cases, with 4x4 binning resulting in greater MTF degradation than 2x2 counterparts. In both 2x2 and 4x4 cases, standard BIN treatments exhibit lower MTF than upsampled BIN and ShiBA treatments.

The degradation of BIN MTF without upsampling is also expressed in the measured LSF as seen in Fig. 4. In both 2x2 and 4x4 cases, upsampled BIN and ShiBA treatments exhibit a narrower LSF.

Although the MTF as plotted in Fig. 3 exhibits little difference between upsampled treatments, the LSF in Fig. 4 reveals smoother response for the ShiBA treatments over the standard BIN scheme.

Fig. 5 presents software processed images of the 0.43 lp/mm feature of a QC-3 phantom, employing full resolution (left); 2x2 and 4x4 treatments (top and bottom rows, respectively) of BIN (center); and 4x4 ShiBA (right). The images were acquired at quantum noise limited doses to illustrate the effect of ShiBA and BIN processing on imager resolution.

**III.A.2. Measurement of NPS**—Fig. 6 plots the electronic NPS and the NNPS for 2.5 MU flat-field images. Hardware binning reduces electronic noise by a factor of approximately 12. Upsampling (for both BIN upsampled and ShiBA cases) further decreases the NPS. This is accounted for by multiplication with a sinc^2 function, which accounts for the larger aperture of the binned pixels, which is observable only when upsampling is performed.
Although the measured electronic NPS for binned readout is lower by a factor of 12 in comparison to full resolution readout, this quantity is affected by lower digital gain – instituted to mitigate effects of reduced dynamic range. Accounting for this reduced gain, the effective decrease in electronic noise for binned readouts is ~3. A comparison of the NNPS between full resolution and 2×2 BIN for a quantum-noise limited projection view as in Fig. 6(b) reveals no appreciable difference after pixel binning aside from the aforementioned sinc² effect for upsampled treatments.

Fig. 7 plots the modeled total NNPS (the NNPS integrated over all measured frequencies) as a function of projection dose (a) as well as the modeled DQE(0) normalized to the value at 2 MU/projection (b). Plotted for comparison is a measurement for the conventional EPID, employing a typical SLI architecture (solid, thin-point line). As plotted in Fig. 7(a), the MLI exhibits lower noise due to the additional layers of scintillator material. Employing hardware binned readout further lowers the NNPS, which is particularly appreciable at projection doses below 0.10 MU. The curve for the normalized DQE(0) plotted in Fig. 7(b) should be flat and equal to approximately 1 for quantum-noise limited operation. A standard SLI EPID exhibits losses of approximately 10% at a projection dose of approximately 0.006 MU. Due to the increased electronic noise from the additional 3 readout layers for the MLI, the DQE(0) exhibits losses of 10% at a higher dose of approximately 0.01 MU. Implementation of readout binning strategies lowers this dose to 0.004 MU.

III.B. Effect on CBCT performance

A band-limited total NNPS is plotted in Fig. 8. The NNPS was integrated over all frequencies between $-f_{\text{lim}}$ and $+f_{\text{lim}}$, where $f_{\text{lim}}$ describes the maximum frequency for the task function modeled (i.e. a 2 mm spherical fiducial marker, a 30 mm spherical tumor, and a cylindrical spine model with a diameter of 45 mm). As in Fig. 8(a), the 2×2 hardware binned readout exhibits lower NNPS than the full resolution case at low-doses. The dose where the 2×2 hardware binned readout deviates from the full resolution readout for the MLI detector is approximately 0.00075 MU, 0.00015 MU, and 0.0001 MU for 2 mm spherical fiducial markers, 30 mm spherical tumors, and the cylindrical spine model, respectively.

Fig. 9 is a dual-axis plot where a comparison of reconstructed $d^\prime_{\text{IP}}$ for each task function (2 mm sphere, 30 mm sphere, and spine, respectively) is presented on the primary axis and the ratio between $d^\prime_{\text{IP}}$ for binned and full resolution images is plotted on the secondary axis, both as a function of dose per projection for a 400-projection scan. For each of the modeled objects, binned readout represents an improvement in $d^\prime_{\text{IP}}$ of approximately 10-12% for all doses above 0.03 MU per projection. Below 0.03 MU, improvement in $d^\prime_{\text{IP}}$ for binned acquisitions increase more rapidly, up to 37-38% at 0.001 MU per projection view.

IV. DISCUSSION

The MTF curves plotted in Fig. 3 show a slight improvement for upsampled treatments of binned readouts compared to standard 2×2 BIN. This improvement is likely due to misalignment of MLI layers. Presently, no engineering is devoted to precise alignment of individual imager layers in the prototype detector. As such the full resolution LSF curves lack smoothness typical of oversampled measurements. Misalignment and addition of large,
binned pixels result in more severe blurring of objects. Pre-addition resampling (as in both upsampled BIN and ShiBA) results in improved phase response, which allows the imager to output a narrower LSF despite detector misalignments and shifts as seen in Fig. 4. Further, by executing a pixel offset prior to image binning and subsequently resampling, as in ShiBA treatments, a phase-averaged image results, which will improve consistency of image quality (removing dependence on object location) and resolution, as well as improving pixilation artifact inherent in widely binned images. This is observable in the LSF curves in Fig. 4. While both BIN and upsampled BIN cases result in characteristic pixilation, ShiBA retains a smooth sampling of the slit image. Examples of this effect may be seen in images of a QC-3 phantom are presented in Fig. 5.

In Fig. 5, the effect of improved phase sampling of ShiBA is apparent. In the presented 0.43 lp/mm feature, effects of phase, detector shift, and pixilation artifact are apparent in the 4×4 BIN image. While ShiBA does not present the feature in equivalent contrast as the full resolution readout, the phase-averaged response allows differentiation of the individual lines with little obvious pixilation. As may be seen in Fig 6., binning strategies such as BIN, upsampled BIN, and ShiBA represents a large improvement in electronic noise, without sacrificing quantum-noise limited performance.

In Fig. 7, this is represented as improved noise response and DQE(0) at low doses. This results in relative improvement in detectability for all objects in Fig. 8, with the $d'_{\text{bin}}/d'_{\text{FR}}$ dramatically improving below 0.01 MU/projection.

Although resolution in phantom images is somewhat degraded as in Fig. 5, at the very low projection doses for CBCT applications - typically 1-3 cGy and 8-10 cGy for kV and MV modalities respectively (Alaei and Spezi, 2015) - hardware binning implementation of ShiBA can provide improved image quality and detectability due to suppression of electronic noise.

Hardware binning has previously been employed as a method for reduction of electronic noise and increasing readout speed. One potential reason for electronic noise reduction through hardware binning may be the reduced amplifier noise. Since each data-line of the detector requires a preamplifier, operating fewer lines of readout will reduce electronic noise (Beutel et al., 2000). Further, by operating fewer data-lines, the distributed resistance and capacitance along detector lines may be affected, reducing injected charge (Beutel et al., 2000). Reduction of electronic noise without concurrent losses in quantum noise performance will allow a reduction of patient dose without additional losses in DQE.

**V. CONCLUSIONS**

We report on a new method for leveraging the unique construction of a multi-layer flat-panel imager (MLI) to reduce the effect of electronic noise by implementing detector binning while also retaining native image sampling. Using this methodology, electronic noise may be drastically reduced without the penalty in degraded image quality normally associated with detector binning. For low dose imaging, such as MV-CBCT, ShiBA may allow a dose
reduction of up to 2.5-fold for 2×2 binned treatments, without noticeable pixilation artifacts. This strategy furthers the potential for clinical MV-CBCT imaging in radiation therapy.

Acknowledgments

The project described was supported, in part, by Award Number R01CA188446 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

References


Fig. 1.
An illustration of the shift-and-bin steps of the ShiBA technique for each (n = 1-4) MLI layer (L). The thicker lines depict pixel bins, which are offset according by \((n-1)\times(n-1)\) pixels and the thin lines represent the standard detector readout.
Fig. 2.
An illustration of the effect of the ShiBA technique on sampling of an arbitrary signal (right) in comparison to standard, aligned MLI orientation (left).
Fig. 3.
Comparison of MTF for 2×2 (a) and 4×4 (b) binned modes of detector operation to a full resolution readout.
Fig. 4.
Comparison of the LSF for 2x2 (a) and 4x4 (b) implementations of full resolution and binned modes (BIN and ShiBA) of readout.
Fig. 5.
Images of the 0.43 lp/mm feature of a QC-3 phantom, employing full resolution images (left), and 2×2 (top) and 4×4 (bottom) BIN (center), and ShiBA (right) techniques.
Fig. 6.
Comparison of 2×2 hardware binned techniques with full resolution imaging for electronic NPS (a) and NNPS for a 2.5 MU flat-field image (b).
Fig. 7.
Total modeled NNPS (a) and modeled normalized DQE(0) (b) as a function of projection dose for SLI (light-point lines); full resolution readout of an MLI detector (heavy-point lines); and binned readout of an MLI detector (heavy-dashed lines).
Fig. 8.
Modeled Total band-limited NNPS as a function of projection dose where $f_{\text{lim}} = 0.5$ cycles/mm (a); $f_{\text{lim}} = 0.3$ cycles/mm (b); $f_{\text{lim}} = 0.05$ cycles/mm (c). $f_{\text{lim}}$ was designated isotropically for each Cartesian direction and roughly corresponds to the maximum frequency of each task function (2 mm sphere, 30 mm sphere, and spine, respectively). In each case, the reconstructed NNPS is represented for a 400 projection acquisition.
Fig. 9.
Dual-axis plots where \( d'_{IP} \) for a 2 mm spherical fiducial marker (a); a 30 mm spherical tumor (b); and a cylindrical spinal model (c) is presented on the primary axis and the ratio between \( d' \) for binned images and full resolution (FR) images is on the secondary axis both as a function of dose per projection. The presented quantities represent \( d' \) for 400 projection scans.

Phys Med Biol. Author manuscript; available in PMC 2019 January 30.