DYNAMIC RECEIVE APERTURE DOWNSAMPLING FOR ULTRASOUND IMAGING

M. Albulayli    D. Rakhmatov

ECE Department, University of Victoria, Canada
Email: daler@ece.uvic.ca

ABSTRACT
We propose a simple method for reducing the number of sampled array elements during ultrasound signal reception. Our approach involves dynamic data-driven switching between various downsampling factors applied to the receive aperture in real time. Our simulation results show that one can achieve approximately 70-85% savings in the total number of sampled array elements per image (thus reducing the total number of expensive analog-to-digital conversions during image acquisition), without significant image quality losses.

Index Terms— Ultrasound imaging, array signal processing, dynamic aperture.

1. INTRODUCTION
Our goal is to reduce the number of ultrasound transducer array elements activated during signal reception, but without significantly affecting the quality of resulting images. Let $M$ be the total number of the regularly spaced elements of a given transducer array. We propose that only $\lfloor M/a(t) \rfloor$ array elements be activated at the sampling instance $t$, where $a(t) \in \{1, 2, ..., M\}$ is a time-varying aperture downsampling factor, and $a(t) - 1$ is the number of unused elements between adjacent activated elements (starting from the first element of the array). For example, if $M = 8$ and $a(t) = 2$, the element activation pattern at $t$ is 10101010, indicating that only the first, third, fifth, and seventh elements of the transducer array are activated; on the other hand, if $a(t + 1) = 4$, the element activation pattern at $t + 1$ is 10001000. There are two obvious advantages to this approach: (1) reducing the input snapshot size from $M$ to $\lfloor M/a(t) \rfloor$ reduces the complexity of adaptive beamforming, and (2) more importantly, it allows us to turn off unused transducer elements, thus avoiding expensive analog-to-digital conversions whenever $a(t) > 1$.

The key disadvantage of using $a(t) > 1$ is the emergence of grating lobes in the angular response of the transducer array, and a common solution to this problem is to impose fixed non-uniform spacing between adjacent array elements [1]. Rather than fixing some optimized non-uniform layout of the array elements during image acquisition, we propose to maintain uniform inter-element spacing when downsampling the receive aperture by $a(t)$, but instead allow $a(t)$ to change dynamically based on the sampled input characteristics. Switching between different downsampling factors takes place whenever the value of the temporally smoothed coherence factor (derived from the sampled input and used as its characteristic feature) crosses certain user-defined thresholds, as described in Section 2.

Figure 1 illustrates the benefits of our approach applied to the simulated 4-MHz ultrasound images of a point-scatterer-cyst (PSC) phantom acquired by a 192-element linear array. The top image in Figure 1 corresponds to the transmit/receive aperture using 64 active elements with no downsampling, which is a default configuration. The center image in Figure 1 corresponds to the transmit/receive aperture using 32 active elements with no downsampling, which substantially compromises the image quality for only 50% savings in analog-to-digital conversions. The bottom image in Figure 1 corresponds to the transmit/receive aperture with 64 elements, but during reception the array elements are activated in accordance with the time-varying downsampling factor $a(t)$ (see Sections 2 and 3 for details). Despite its simplicity, our method yields approximately 74% savings in analog-to-digital conversions without significant degradation in the image quality, in comparison to the top image in Figure 1.

2. PROPOSED METHOD FOR DYNAMIC APERTURE DOWNSAMPLING
Let $\overrightarrow{M} = \lfloor M/a(t) \rfloor$ represent the time-varying number of ultrasound transducer array elements activated during signal reception, and let $\overrightarrow{\mathbf{x}}(t)$ denote the resulting input snapshot, which is a vector of $\overrightarrow{M}$ data samples taken by the receive aperture at the sampling instance $t$. We assume that appropriate focusing delays are applied at every $t$, so that $\overrightarrow{\mathbf{x}}(t)$ is a real-valued phase-compensated input snapshot, whose $\overrightarrow{M}$-element steering vector is denoted by $\overrightarrow{\mathbf{a}} = [1 \ 1 \ldots 1]^T$. Each $\overrightarrow{\mathbf{x}}(t)$ is

The authors thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Saudi Arabian Cultural Bureau in Canada for financial support.

The PSC phantom is placed at 60 mm (transmit focus) and consists of a point target, a highly scattering region with the radius of 1.5 mm, and a water-filled cyst region with the radius of 2 mm. The lateral distances are $-14$ mm, $-5$ mm, and 10 mm, respectively. During reception, dynamic receive focusing is performed at 10-mm intervals.
signals, the relative impact of correlation between desired and interfering signals is non-stationary). To mitigate the negative impact, adaptive beamforming can substantially enhance the resolution and contrast of the displayed ultrasound images [1]. The optimal MVDR beamformer weights are given by

$$w_{opt}(t) = \frac{R^{-1}(t)d}{d^H R^{-1}(t)d}, \quad (1)$$

As $R(t)$ is typically unknown, it is estimated using the sample correlation matrix $\tilde{R}(t) = \frac{1}{N} \sum_{t=-N+1}^{t} x(t)x(t)^H$, where $N$ is the number of input snapshots ($N$ is usually very small, as ultrasound signals are non-stationary). To mitigate the negative impact of correlation between desired and interfering signals, $\tilde{R}(t)$ is replaced by a spatially smoothed sample correlation matrix $\tilde{R}(t)$ computed as follows [10]:

$$\tilde{R}(t) = \frac{1}{(M-L+1)N} \sum_{t=-N+1}^{t} \sum_{k=1}^{M-L+1} \tilde{x}_k(t)\tilde{x}_k^H(t), \quad (2)$$

where $\tilde{x}_k(t) = [x_k(t) x_{k+1}(t) \ldots x_{k+L-1}(t)]^T$ is the $k$th subarray of size $\mathcal{L} = \lfloor M/2 \rfloor$ within $x(t)$. Substituting $\tilde{R}(t)$ for $R(t)$ in Equation (1) yields

$$w(t) = \frac{\tilde{R}^{-1}(t)d}{d^H \tilde{R}^{-1}(t)d} \quad (3)$$

which is a weight vector of size $\mathcal{L}$, leading to the following beamformed output [3]:

$$y(t) = \frac{w^H(t)}{M-L+1} \sum_{k=1}^{M-L+1} \tilde{x}_k(t). \quad (4)$$

The MVDR beamformer can be implemented using the GSC structure shown in Figure 2, whose optimal weights and output are computed as follows [1]:

$$\tilde{w}_a(t) = [\tilde{B}^H(\tilde{R}(t)\tilde{B})]^{-1}\tilde{B}^H(\tilde{R}(t))\tilde{w}_q, \quad (5)$$

$$y(t) = \frac{\tilde{w}_q - \tilde{B}\tilde{w}_a(t)}{M-L+1} \sum_{k=1}^{M-L+1} \tilde{x}_k(t). \quad (6)$$

Note that $B$, $w_q$, and $w_a(t)$ from Figure 2 have been replaced by their respective counterparts $\tilde{B}$, $\tilde{w}_q$, and $\tilde{w}_a(t)$ with reduced dimensions to comply with the $L \times L$ size of $\tilde{R}(t)$.

**Coherence Factor.** The coherence factor is defined as

$$\text{CF}(t) = \frac{|d^H x(t)|^2}{M \sum_{i=1}^{M} |x_i(t)|^2} = \frac{|\sum_{i=1}^{M} x_i(t)|^2}{M \sum_{i=1}^{M} |x_i(t)|^2} \in [0, 1]. \quad (7)$$

The CF value, ranging from 0 and 1, can be interpreted as the ratio of the on-axis power to the total received power [9]. Multiplying the beamformer output $y(t)$ by $\text{CF}(t)$ has been shown to improve the image quality (e.g., see [9, 11, 12, 13]), and we follow this practice as well. Additionally, we calculate the temporally smoothed coherence factor, denoted by $\tilde{\text{CF}}(t)$, which we define as a simple moving average taken over the last $T$ computed CF values:

$$\tilde{\text{CF}}(t) = \frac{\sum_{n=1}^{T} \text{CF}(n)}{T} \in [0, 1]. \quad (8)$$

We use $\tilde{\text{CF}}(t)$ to decide on the receive aperture downsampling factor $a(t+1)$ as described next.
Aperture Downsampling. We restrict \( a(t) \) to be from a user-defined set of values \( \{\alpha_1, \alpha_2, ..., \alpha_K\} \), such that \( K > 1 \) and \( M \geq \alpha_1 > \alpha_2 > ... > \alpha_K \geq 1 \). We also partition the interval \([0, 1]\) into \( K \) regions denoted by \( \Delta_1 = [0, \beta_1) \), \( \Delta_2 = [\beta_1, \beta_2), ..., \Delta_K = [\beta_{K-1}, 1) \), where \( 0 < \beta_1 < \beta_2 < \ldots < \beta_{K-1} < 1 \) represent user-defined region boundaries. If \( CF(t) \in \Delta_j \), then \( a(t+1) = \alpha_j \), where \( j = 1, 2, ..., K \). In other words, whenever the temporally smoothed coherence factor is detected to be in the region \( \Delta_j \), the aperture downsampling factor for the next sampling instance is set to \( \alpha_j \).

Our proposed method for using dynamically downsampled receive aperture during signal reception is summarized below, where initially \( a(0) = 1 \):

1. Starting from the first element of the transducer array, activate \( M = \lceil M/a(t) \rceil \) elements, so that adjacent activated elements are separated by \( a(t) - 1 \) unused element(s).
2. Acquire \( M \)-element input snapshot \( \mathbf{x}(t) \), compute \( CF(t) \), and update \( CF(t) \).
3. Determine \( j \in \{1, 2, ..., K\} \) such that \( \widetilde{CF}(t) \in \Delta_j \), and let \( a(t+1) = \alpha_j \).
4. Compute spatially smoothed GSC-beamformed output \( y(t) \), and let \( y(t) \leftarrow CF(t) \cdot y(t) \).

The amount of savings in analog-to-digital conversions per image is given by

\[
\text{Savings/Image} = \left(1 - \frac{\sum_t \lceil M/a(t) \rceil}{\sum_t M}\right) \times 100\%, \quad (9)
\]

where \( \sum_t \) represents the summation over the total number of inputs snapshots taken to form an entire image. This amount varies depending on several factors: (1) the choice of \( \{\alpha_1, \alpha_2, ..., \alpha_K\} \) and \( \{\beta_1, \beta_2, ..., \beta_{K-1}\} \), (2) the choice of a smoothing equation for the coherence factor, and (3) the input data affecting the \( CF(t) \) values. Additionally, there are computational savings in beamforming-related calculations. Computing \( y(t) \) with the GSC involves a matrix inversion requiring \( O(L^3) \) multiplications, where \( L \) depends on \( a(t) \).

### 3. EVALUATION RESULTS

In this section, we evaluate several downsampling scenarios based on the simulated 4-MHz ultrasound images of a 12-point phantom acquired by a phased array with \( M = 96 \) and \( N = 1 \), and a PSC phantom acquired by a linear array with \( M = 64 \) and \( N = 2 \) (as described in Section 1).\(^2\) The simulations were performed using the FIELD-II tool [14].

Table 1 lists our simulated downsampling scenarios S1-S12. Scenario S1 involves no downsampling, i.e., \( a(t) = 1 \), whereas in S2, S3, and S4, \( a(t) \) is always fixed at 2, 4, and 8, respectively (no switching). For the remaining scenarios S5-S8 (with \( T = 2 \)) and S9-S12 (with \( T = 4 \)), we use \( K = 4 \) and let \( a(t) \in \{8, 4, 2, 1\} \) vary in time. Table 1 also shows quantitative quality indicators for images obtained using our downsampling scenarios, as well as the resulting savings in analog-to-digital conversions (% ADC). We use scenario S1 (no downsampling) as our reference for comparison purposes.

For the 12-point phantom images, we rely on the FWHM (full width at half maximum) as an indication of the resolution quality and the sidelobe energy \( E_{SL} \) (calculated for attenuation greater than 25 dB) as an indication of the contrast quality. Both quantities are calculated at the 60-mm transmit focus point. Scenario S9 yields FWHM = 0.3003 mm and \( E_{SL} = -36.67 \) dB that practically match those in S1 (FWHM = 0.3008 mm and \( E_{SL} = -37.02 \) dB), while offering 85.3% in ADC savings.

For the PSC phantom images, we rely on the contrast values (calculated with respect to the speckled background) for the scatterer and the cyst. Scenario S6 yields the scatterer contrast value of 3.235, which is approximately 5.55% worse than that in S1 (equal to 3.425); however, S6 offers 73.9% in ADC savings, which is a sensible tradeoff. The cyst contrast values due to S6 and S1 are practically the same (equal to 0.970 and 0.967, respectively).

Figures 3 and 4 show several 12-point and PSC phantom images acquired using scenarios S1 (no downsampling), S4 (worst simulated static downsampling), S6 (best simulated dynamic downsampling in terms of scatterer contrast), S9 (best simulated dynamic downsampling in terms of \( E_{SL} \)), and S12 (worst simulated dynamic downsampling). If we are to recommend a single downsampling scenario for both phantoms, it would be S6 (also see Figure 1 and Table 1).

\(^2\)The 12-point phantom consists of 12 single point targets placed at the 10-mm intervals starting at 30 mm from the transducer surface. During transmission, the focus is fixed at 60 mm, and during reception, dynamic receive focusing is performed at 10-mm intervals.
Table 1. Quantitative image quality indicators for simulated downsampling scenarios.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>{α_1, α_2, α_3, α_4}</th>
<th>{β_1, β_2, β_3}</th>
<th>T</th>
<th>FWHM (mm)</th>
<th>(E_{SL}) (dB)</th>
<th>Savings (% ADC)</th>
<th>Contrast (Scatterer)</th>
<th>Contrast (Cyst)</th>
<th>Savings (% ADC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>(a(t) = 1)</td>
<td>—</td>
<td>—</td>
<td>0.3008</td>
<td>-37.02</td>
<td>0</td>
<td>3.425</td>
<td>0.967</td>
<td>0</td>
</tr>
<tr>
<td>S2</td>
<td>(a(t) = 2)</td>
<td>—</td>
<td>—</td>
<td>0.3024</td>
<td>-35.83</td>
<td>50</td>
<td>3.063</td>
<td>0.954</td>
<td>50</td>
</tr>
<tr>
<td>S3</td>
<td>(a(t) = 4)</td>
<td>—</td>
<td>—</td>
<td>0.3015</td>
<td>-34.98</td>
<td>75</td>
<td>2.613</td>
<td>0.965</td>
<td>75</td>
</tr>
<tr>
<td>S4</td>
<td>(a(t) = 8)</td>
<td>—</td>
<td>—</td>
<td>0.3793</td>
<td>-26.56</td>
<td>87.5</td>
<td>1.891</td>
<td>0.965</td>
<td>87.5</td>
</tr>
<tr>
<td>S5</td>
<td>{8, 4, 2, 1}</td>
<td>{0.125, 0.25, 0.5}</td>
<td>2</td>
<td>0.3021</td>
<td>-35.74</td>
<td>84.7</td>
<td>3.151</td>
<td>0.978</td>
<td>70.6</td>
</tr>
<tr>
<td>S6</td>
<td>{8, 4, 2, 1}</td>
<td>{0.125, 0.25, 0.75}</td>
<td>2</td>
<td>0.3023</td>
<td>-35.56</td>
<td>84.8</td>
<td>3.235</td>
<td>0.970</td>
<td>73.9</td>
</tr>
<tr>
<td>S7</td>
<td>{8, 4, 2, 1}</td>
<td>{0.125, 0.5, 0.75}</td>
<td>2</td>
<td>0.3022</td>
<td>-35.38</td>
<td>85.1</td>
<td>3.036</td>
<td>0.974</td>
<td>76.5</td>
</tr>
<tr>
<td>S8</td>
<td>{8, 4, 2, 1}</td>
<td>{0.25, 0.5, 0.75}</td>
<td>2</td>
<td>0.3014</td>
<td>-35.34</td>
<td>86.8</td>
<td>2.862</td>
<td>0.963</td>
<td>78.9</td>
</tr>
<tr>
<td>S9</td>
<td>{8, 4, 2, 1}</td>
<td>{0.125, 0.25, 0.5}</td>
<td>4</td>
<td>0.3003</td>
<td>-36.67</td>
<td>85.3</td>
<td>3.059</td>
<td>0.951</td>
<td>70.9</td>
</tr>
<tr>
<td>S10</td>
<td>{8, 4, 2, 1}</td>
<td>{0.125, 0.25, 0.75}</td>
<td>4</td>
<td>0.3004</td>
<td>-36.41</td>
<td>85.4</td>
<td>2.929</td>
<td>0.973</td>
<td>74.4</td>
</tr>
<tr>
<td>S11</td>
<td>{8, 4, 2, 1}</td>
<td>{0.125, 0.5, 0.75}</td>
<td>4</td>
<td>0.3004</td>
<td>-36.22</td>
<td>85.5</td>
<td>2.730</td>
<td>0.964</td>
<td>76.8</td>
</tr>
<tr>
<td>S12</td>
<td>{8, 4, 2, 1}</td>
<td>{0.25, 0.5, 0.75}</td>
<td>4</td>
<td>0.3089</td>
<td>-31.62</td>
<td>87.1</td>
<td>2.560</td>
<td>0.963</td>
<td>79.8</td>
</tr>
</tbody>
</table>

4. RELATION TO PRIOR WORK

Sparse sampling of array elements has been under investigation by many researchers over several decades, e.g., some of the recent advances in this area are reported in [15, 16, 17, 18, 19, 20, 21]. Design of sparse arrays typically aims at eliminating grating lobes in the array angular response and achieving beampatterns with a narrow mainlobe width and low sidelobe levels. Finding good sparse samplers of transmit/receive apertures entails solving hard computational problems, and it is impractical to optimize such samplers dynamically in a non-stationary environment. Rather than committing to a single spatial sampling pattern, our method uses a set of simple downsampling-based activation patterns during signal reception and dynamically switches from one pattern to another (i.e., turning certain array elements on and off as needed). Switching is triggered by the sampled input itself, when its smoothed coherence factor is detected to cross certain thresholds. Coherence-factor-based thresholding was also used in [22] to decide on the number of samples (within a full-aperture sampled input snapshot) that were involved in computing an optimal beamformed output. However, unlike this work, reference [22] was not concerned with reducing the size of sampled input snapshots. Reference [23] mentions the idea of using an increased aperture size at greater imaging depths, which helps maintain uniform resolution. In our case, the aperture size (in terms of the number of wavelengths spanned) remains the same, but it is subject to dynamic downsampling, which is decided based on the actual input data. To the best of our knowledge, the proposed method is novel, and as our preliminary simulation results demonstrate, it is rather effective despite its simplicity. As the next step, this work may be extended in several important directions, such as dynamic nonuniform aperture sampling, adaptive control of switching thresholds, alternative mechanisms for activation pattern switching, and combining spatial sampling sparsity with compressed sensing in time [24, 25].

Fig. 4. PSC phantom (from top to bottom): S1, S4, S6, S9, S12 (see Table 1). Horizontal/vertical axes are lateral/axial distances in mm.
5. REFERENCES


