GPU-accelerated angular spectrum approach for real-time full field acoustic beam prediction

Mario Hlawitschka, Robert J. McGough, Katherine W. Ferrara, Dustin E. Kruse

Introduction

In our preliminary work, we have demonstrated that ultrasonic drug release using mild hyperthermia is feasible; temperature sensitive drug delivery particles are injected and allowed to accumulate within a region of interest, typically over 6-24 hours. Ultrasound is then scanned through the volume of interest to locally increase the temperature by 2 – 4°C for a short period, releasing a drug and increasing its efficacy.

In order to translate these methods, ultrasound AND drug dose must each be predicted and validated. We have made great strides on drug dose—we now tackle ultrasound dose for this special problem of mild hyperthermia. Predicting the 3-D heating profile in tissue as a function of time requires knowledge of the acoustic beam intensity as a function of time as it is steered in 3-D space. Ultrasound beams generated from an aperture large enough to deliver the acoustic power necessary for tissue heating must be focused to some degree in order to work at reasonable depths from the aperture face and at depth-of-fields that are not much larger that the desired heating region. Focusing necessitates that the beam be scanned in order to treat larger volumes, such as the volume encompassing a tumor. The beam scanning is accomplished dynamically, in real time, and knowledge of the beam is necessary in order to make predictions on *where* energy is being deposited and *how much* is being deposited. This information is important for predicting thermal dose according to Arrhenius-based methods such as Cumulative Equivalent Minutes at 43°C (CEM43).

Background

Beam simulations are notoriously time consuming, which has precluded their use in real-time applications. As shown below, ultrasound beam calculations in the monochromatic and linear regime rely heavily upon the evaluation of integrals of complex exponentials:

First Rayleigh-Sommerfeld Method[1]

$$p(\mathbf{r},t) = j\rho c k e^{j\omega t}\int_S u(\mathbf{r})\frac{e^{-jk|\mathbf{r}-\mathbf{r}|}}{2\pi|\mathbf{r}-\mathbf{r}|}d\mathbf{S}$$

Fast Near-field Method (FNM)[2]

$$p(x,y,z,t) = j\rho c u_0 e^{j\omega t}\frac{1}{2\pi}\left(s_1\int_{l_1}^{l_2}\frac{e^{-jk\sqrt{z^2+\sigma^2+s_1^2}}-e^{-jkz}}{\sigma^2+s_1^2}d\sigma + l_1\int_{s_1}^{s_2}\frac{e^{-jk\sqrt{z^2+\sigma^2+l_1^2}}-e^{-jkz}}{\sigma^2+l_1^2}d\sigma - s_2\int_{l_1}^{l_2}\frac{e^{-jk\sqrt{z^2+\sigma^2+s_2^2}}-e^{-jkz}}{\sigma^2+s_2^2}d\sigma - l_2\int_{s_1}^{s_2}\frac{e^{-jk\sqrt{z^2+\sigma^2+l_2^2}}-e^{-jkz}}{\sigma^2+l_2^2}d\sigma\right)$$

The angular spectrum approach (ASA) was first used in optics to propagate fields from a source plane to a destination plane [1]. The ASA works by decomposing the source plane into plane waves using the Fourier transform. The plane wave components are then propagated in the Fourier domain and reconstructed using the inverse Fourier transform. Performing the computations on a uniform and regular grid enables the use of Fast Fourier Transforms (FFTs) to accelerate the computation. Hence, computing a 3-D acoustic field requires computing only one source plane according to either of the approximate methods given above.

Two ASA implementations using the spectral propagator, h_u, and spectral propagator, H_u, are given by:

$$p(x,y,z,t) = j\rho c k e^{j\omega t}u(x,y)\otimes h_u(x,y,z) \qquad h_u(x,y,z) = \frac{e^{-jkr}}{2\pi r}$$

$$P(k_x,k_y,z,t) = j\rho c k e^{j\omega t}U(k_x,k_y)H_u(k_x,k_y,z) \qquad H_u(k_x,k_y,z) = \begin{cases}\dfrac{e^{-jz\sqrt{k^2-k_x^2-k_y^2}}}{j\sqrt{k^2-k_x^2-k_y^2}}, & k_x^2+k_y^2\le k^2\\[2ex] \dfrac{e^{-z\sqrt{k_x^2+k_y^2-k^2}}}{\sqrt{k_x^2+k_y^2-k^2}}, & k_x^2+k_y^2> k^2\end{cases}$$

Here, u(x,y) is the distribution of the normal velocity on the radiator. The spectral propagator has been shown to be faster than the spectral propagator due to the analytical H_u function (obviating the need for an additional FFT at each step) [3], so our analysis focuses on accelerating the spectral propagator. Errors due to aliasing and grid truncation are reduced through the use of a low pass filter and zero padding as detailed in [3].

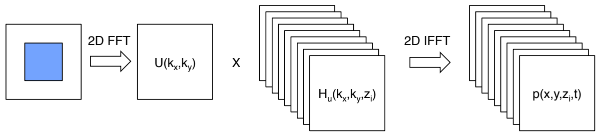
[1] Goodman, JW Introduction to Fourier optics. 3rd ed. Roberts & Co., Englewood, Colo., 2005.
[2] McGough, RJ, "Rapid calculations of time-harmonic nearfield pressures produced by rectangular pistons", *J. Acoust. Soc. Am.*, 115:1934-41, 2004.
[3] Zeng X, McGough RJ, "Evaluation of the Angular spectrum approach for simulations of near-field pressures", *J. Acoust. Soc. Am.*, 123(1):68-76, 2008.
[4] McGough, RJ http://www.egr.msu.edu/~fultras-web/
[5] Kruse DE, Lai C-Y, Stephens DN, Sutcliffe P, Paoli EE, Barnes SH, Ferrara KW, "Spatial and temporal controlled tissue heating on a modified clinical ultrasound scanner for generating mild hyperthermia in tumors", *IEEE Trans Biomed Eng.* 57(1):155-166, 2010.

Objective

The objective of this work is to employ the versatility and speed of modern graphics hardware to increase the speed of ultrasound beam predictions using the ASA method.

Implementation on the GPU

The GPU implementation of the spectral propagator is illustrated below. First, we use the CUDA implementation of the FFT to perform the forward Fourier transform of a zero-padded, u(x,y), obtaining $U(k_x,k_y)$, which is stored and reused for each plane. The second step of our method is implemented as a single CUDA kernel that computes $H_u(k_x,k_y,z)$ for this point in z and multiplies it by the source plane in Fourier space, $U(k_x,k_y)$. In the final step, the result is inverse Fourier transformed and multiplied by a complex scale factor. To increase the speed depending on the setup, individual implementations exist for the attenuating and non-attenuating media. All implementations exist using angular restriction and without angular restriction, but due to the previously proven increased accuracy of the restricted propagator [1], this is the default implementation and the implementation we use for speed comparison.



Speed

The algorithm was implemented in C++ and CUDA, a C-like language designed for GPUs. We provide interfaces to C++ and MATLAB for convenience and speed in addition to wrapper functions to FOCUS [3]. Our reference implementations are written in C++ and MATLAB.

Table I. Comparison of GPU computation times of different methods on the CPU and GPU. All times are in units of seconds.

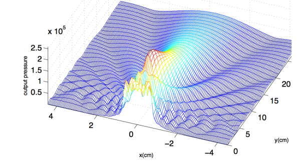
	Calculation Mode	Pressure, Spectral, Attenuated	Pressure, Spectral, Attenuated	Pressure, Spectral, Attenuated	Pressure, Spectral	Pressure, Spectral, Attenuated	Pressure, Spatial	Pressure, Spectral, Attenuated
	Processor	NVIDIA Quadro FX 3800 (GT200GL)	NVIDIA 1060	NVIDIA 1060	Intel Xeon E5620@2.40GHz	Intel Xeon E5620@2.40GHz	Intel Xeon E5620@2.40GHz	Intel Xeon X5550@2.67GHz
	Precision	float	float	double	double	double	double	double
Field Points	In: 151x151 Out:301x301x200 FFT 512	1.8	1.5	1.8	10	10	13	5
	In: 151x151 Out:301x301x200 FFT 1024	6.8	4.8	6.7	35	35	42	14
	151x151 Out:301x301x200 FFT 2048	34.5	14.3	17.3	94	94	124	67
	Software	GPU	GPU	GPU	FOCUS	FOCUS	FOCUS	FOCUS

All timings are made including copying the data to MATLAB.

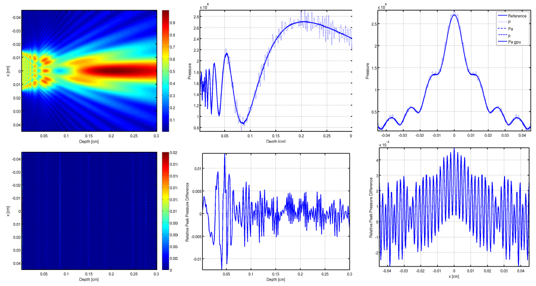
ACKNOWLEDGMENTS: We wish to acknowledge financial support from the National Institute of Biomedical Imaging and Bioengineering through NIHR21EB009434 and ARRA funding through NIHR21EB009434-01S1.

3-D Beam Simulations

We test the accuracy of our calculations based on the same test case presented in [3] for a 3cm x 3cm piston, with a transverse spatial sampling of λ/5 (0.3mm), f=1MHz, c=1500m/s. The initial plane is computed at z=0.15cm using the FNM with 100 Gauss abscissas. Accuracy is measured against the full FNM computation using 100 Gauss abscissas.

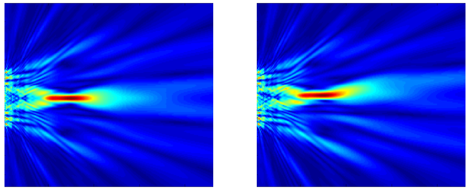


We show that our single precision GPU implementation reaches a precision close to existing implementations of the Angular Spectrum Approach. The dominant error in the GPU implementation is the round off error as a function of position along the z-axis. The following plots show a comparison of different computation methods (top) and the resulting error relative to the peak pressure of the GPU computation compared to the reference implementation (bottom), showing a maximum relative error just over 1%.



Array Beam Simulations

The simulation is independent of array complexity, including geometry, apodization, and phasing as long as the pressure can be given on an initial plane. Here, array simulations are shown for an array of 10 x 10 elements with a beam focused at the array's center and off-center.



Discussion and Conclusion

We have shown that an efficient implementation of the Angular Spectrum Approach on the GPU is possible and performs at reasonable precision for both single-precision and double-precision floating point calculations. Even with the overhead of copying data from and to the GPU, the implementation performs significantly better than previous implementations run on a fast CPU, with a 3–50 times speed improvement for the field sizes tested.

CENTER FOR MOLECULAR AND GENOMIC IMAGING
DEPARTMENT OF BIOMEDICAL ENGINEERING

SCIENTIFIC VISUALIZATION GROUP