

Focused Ultrasound - Efficient GPU Simulation Methods for Therapy Planning

J. Georgii^{†1} and C. v. Dresky¹ and S. Meier¹ and D. Demedts¹ and C. Schumann¹ and T. Preusser^{1,2}

¹Fraunhofer MEVIS, Germany

²School of Engineering and Science, Jacobs University Bremen, Germany

Abstract

Over the past years, high intensity focused ultrasound therapy has become a promising therapeutic alternative for non-invasive tumor treatment. The basic idea of this interventional approach is to apply focused ultrasound waves to the tumor tissue such that the cells are heated and hence destroyed. Since it is quite difficult to assess the quality of this non-invasive therapy, there is a dire need for computer support in planning, conduction, and monitoring of such treatments.

In this work, we propose efficient simulation techniques for focused ultrasound waves as well as their heat dissemination using current graphics hardware as a numerical co-processor. We achieve speed-ups between 10 and 700 for the single simulation steps compared to an optimized CPU solution, overall resulting in a significant performance gain over previous approaches for simulation of focused ultrasound.

Categories and Subject Descriptors (according to ACM CCS): I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Physically based modeling I.6.7 [Simulation and Modeling]: Simulation Support Systems— J.3 [Life and Medical Sciences]: Health—

1. Introduction

During the last decade, High Intensity Focused Ultrasound (HIFU) therapy has become a promising treatment alternative to the surgical resection of tumors [Jol09]. This approach is based on the application of focused ultrasound waves to the tumor tissue, which leads to its heating, the coagulation of proteins and thus the destruction of cells. The energy density of the acoustic wave is more than 1,000 times higher than for the well-known diagnostic ultrasound. The short time application of high intensity focused ultrasound causes sharply delimited small lenses (sonications) of thermally destroyed tissue. Through a refocusing and a replacement of the ultrasound transducer the complete tumor tissue and a safety margin can be destroyed.

Due to its few side effects and many further advantages, this non-invasive therapy has great potential for tumor treatment. However, many physical processes related to this treatment are not yet well-understood. For this reason the

application to malign tumors is still limited to prostate cancer, while this therapy is already approved for the treatment of some benign tumors and metastases such as uterine fibroids, adenomyosis and bone metastases. In order to overcome the occurring difficulties and to expand this promising treatment to further indications, there is an urgent need for further research. In current studies, the application of HIFU to the liver, kidney, breast, and brain are examined.

Some of the current difficulties are related to the non-invasiveness of HIFU therapies, e.g. it is quite difficult to assess the quality of a HIFU treatment. In contrast to surgical resection no pathological workup of the malignant tissue is possible and the success of the therapy must be assessed using imaging techniques only. Moreover, the control of the thermal ablation is quite challenging as it results from the complex interplay of a variety of bio-physical and bio-chemical processes. In fact the extra-corporeally created focused ultrasound wave is reflected and refracted multiple times on the path towards the target site. Furthermore, the conversion of the acoustic pressure gradients into heat depends on the absorption rate of the tissue and on the complex vascular structures in the vicinity of the lesion. This dis-

[†] joachim.georgii@mevis.fraunhofer.de

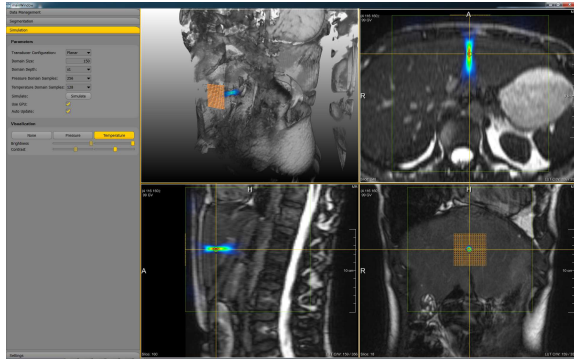


Figure 1: Screenshot of the software prototype. A single sonication in the liver is simulated in the MRI data set. The temperature field is visualized using a color map ranging from blue to red. The ultrasound transducer is shown in brown color.

cussion motivates the development of computer support for planning, conduction and monitoring of HIFU treatments.

In this paper we deal with the numerical simulation of the heat development caused by focused ultrasound. We discuss a model that consists of equations describing the propagation of the ultrasound wave and the heat diffusion in the tissue. Modern ultrasound transducers consist of several hundreds of Piezo elements, which can individually create pressure waves of different intensity and phase. Consequently, a numerical simulation of the acoustic field in the vicinity of the ultrasound transducer (nearfield) must superpose the single element contributions. Moreover, for the simulation of the acoustic wave in the farfield, i.e. in some distance from the transducer, there is a natural connection of the wave's frequency to the spatial resolution. Thus, for the megahertz acoustic ultrasound waves used in HIFU treatments, computational grids of very fine resolution are needed leading to an enormous computational demand. In this paper we use graphics hardware to tackle this computational demand. Our goal is to develop efficient simulation techniques that can be integrated into an interactive software assistant for the prediction of the focused ultrasound wave and the thermal tissue destruction. A screenshot of a first software prototype is shown in Figure 1.

1.1. GPU Architecture

Over the last years, graphics processing units (GPUs) have shown a significant increase in performance on intrinsically parallel computations. This is due to the fact that GPU design focuses on maximizing the throughput of all available parallel units. GPUs provide a high number of parallel units as well as efficient hardware-support to hold significantly more active threads than parallel units. Thereby, the GPU can effectively hide latencies, and thus GPUs often outper-

form parallel CPU-based implementations. Consequently, in the last years considerable effort has been spent on the efficient implementation of general numerical methods on programmable GPUs [OLG*05, HG07, OHL*08]. GPUs can be accessed by a number of high-level APIs. One of the most popular is the CUDA API [NVI08], which, however, is limited to NVIDIA architectures. On the other hand, OpenCL [Khr11] is a platform-independent alternative maintained by the Khronos group, which allows execution on GPUs of different vendors.

NVIDIA's recent Fermi architecture [NVI09] consists of 15 multiprocessors, and each multiprocessor has 32 scalar cores to execute integer or floating point operations. Currently, double precision floating point operations are supported at half the speed of single precision operations. On each multiprocessor several hundreds threads can exist co-residently, and hardware registers are shared among these threads. Memory can be accessed read-only through texture filtering units, which provide tri-linear interpolation in hardware as well as an on-chip texture cache, which can significantly improve the performance.

The threads on each multiprocessor are executed in so called warps, in which all of the 32 threads run in lock-step. Therefore, the GPU can only exploit its full potential if all threads within one warp have the same execution path. Although threads can follow different execution paths, in this case the executions are serialized thereby reducing the overall performance. By automatically scheduling warps on the multiprocessor at very low costs, latencies caused by memory access operations can be hidden effectively.

1.2. Contribution

To the best of our knowledge, this is the first approach that simulates focused ultrasound on a GPU. Especially, we propose effective parallelization approaches for state-of-the-art techniques to determine the ultrasound near- and farfield. This includes the heavy use of the computational units of the GPU. Furthermore, our approach significantly benefits from the texture filtering units which allow to perform tri-linear interpolation in a data set. For that purpose, we propose a scheme that stores the complex data values as magnitude and phase value rather than as real and imaginary part such that hardware interpolation of the sound waves is energy preserving even if a wave is only sampled at a few nodes. This is important to accurately determine the heat dissemination in the tissue, which is achieved by solving the bioheat transfer equation directly on the GPU, too.

2. Related Work

Numerical simulation of ultrasound propagation in biological tissue is a challenging task due to the fine-scale interaction of the short waves with complex interfaces at bones and tissues of varying acoustic properties. For focused waves of

high amplitude as in cancer treatment with HIFU, additional effects of non-linear wave propagation may become important [HBtH02, EH04, Lei07].

The classical linear model for time-harmonic acoustic wave propagation in inhomogeneous tissue is the complex-valued Helmholtz equation. Different numerical methods for this model have been suggested for the simulation of therapeutic ultrasound like finite differences [ATP*03], pseudo-spectral finite elements [LHHH10, WFZ09] and the ultra-weak variational formulation [HMK*05]. All of these methods, however, involve expensive computations on a three-dimensional globally fine resolved mesh and suffer from ill-conditioned system matrices. In non-linear models suggested for focused ultrasound, the computational effort is even increased [CMK*00, MCtH00, LLZG07].

In a homogeneous medium, the pressure field generated by a single transducer element can be described by the analytical Rayleigh-Sommerfeld integral, which is an approximation of the well-known Kirchhoff's integral formula for the Helmholtz equation. In order to reduce the computational costs for computing the pressure field in 3D domains, an already computed axial slice can be advanced in the propagation direction with the help of a Fourier transformation in the plane directions (angular spectrum method). For details, see [HBtH02, ZM08, ZM09] and references therein. However, these methods can only be applied within a homogeneous medium, for instance in the water bath surrounding the transducer. In order to deal with media interfaces, some extensions of the angular spectrum method have been suggested. In [CH02], the ultrasound waves are focused through the human skull using a combined angular spectrum and ray-tracing simulation approach. In [VC08], the angular spectrum method is extended by an additional phase correction in order to account for inhomogeneities.

Highly parallel distributed memory computations have been recently introduced for the optimization of antenna parameters in micro-wave induced hyperthermia [SMS*09]. Although conceptually similar, the physics is different in this application, since the acoustic wave solver needs to be replaced by an electromagnetic field solver.

Ultrasound simulation on the GPU based on ray-tracing algorithms has been applied before in the context of diagnostic ultrasound [RPAS09, KSN09]. In these works, the focus is either the real-time generation of physically plausible 2D ultrasound images for training simulators or on real-time registration of 2D ultrasound images.

3. Simulation of Focused Ultrasound

The focusing of ultrasound waves is typically achieved by a phased array system (ultrasound transducer) which consists of several (up to hundreds) transducer elements. These elements can either be arranged on a plane or on a sphere, and each transducer element can be controlled individually in its

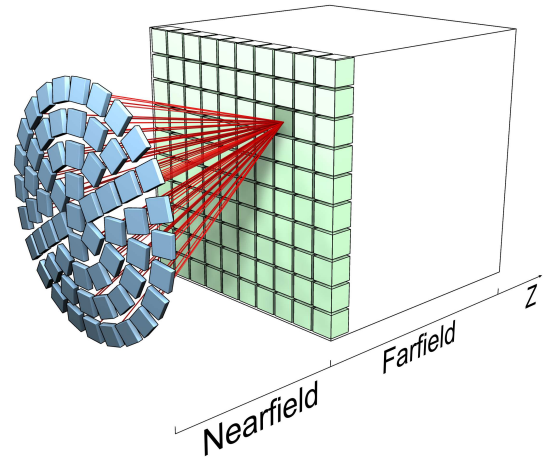


Figure 2: Overview of the simulation of focused ultrasound. The contribution of each single transducer element (blue) is determined at the interface (green) in the nearfield simulation. In the farfield, the pressure field can be determined by the angular spectrum approach thereby respecting material inhomogeneities.

intensity and phase shift. By adjusting the phase shift values accordingly, the interference of the waves emerging from the transducer elements can have a high intensity focus at a predefined position. The goal of the so-called nearfield simulation is to determine the contribution of each single transducer element assuming homogeneous material surrounding the transducer elements (water or ultrasound gel), and thus to simulate the focusing waves at an interface layer of the farfield. The nearfield computations are typically achieved by evaluating the Rayleigh-Sommerfeld integrals for each transducer element and computing the superposition of all of these waves at an interface layer of the farfield. In the farfield, one cannot assume homogeneous material, since we have to respect the transitions from the ultrasound gel to the skin and then through different types of tissue (eventually including bones). Hence, the goal of the farfield simulation is to take these inhomogeneities into account. The overall simulation approach described above is illustrated in Figure 2. Note, that we only compute the pressure values (and temperature values) in the farfield (including the interface layer), which we also refer to as the simulation domain. However, the transducer is positioned outside of this domain. Furthermore, we assume that the main propagation direction of the waves is along the z axis.

3.1. Nearfield

The nearfield simulation has to compute the pressure field as it is generated by the interference of all transducer elements. Thereby, one has to take into account individual transducer

parameters such as the phase shift and the intensity. For a single transducer element, the Rayleigh-Sommerfeld integral can be used to determine the pressure field. This is an approximation of the well-known Kirchhoff's integral formula for the Helmholtz equation. It is therefore only valid in a homogeneous medium. The complex sound pressure p at $\mathbf{x} = (x, y, z)$, $z > 0$ can then be described by

$$p(\mathbf{x}) = \frac{ik}{2\pi} \rho c e^{i\omega t} \int_{\xi \in \mathbf{S}} u(\xi) \frac{e^{-ik|\mathbf{x}-\xi|}}{|\mathbf{x}-\xi|} d\mathbf{S}. \quad (1)$$

Here, time t is regarded as a parameter, and ρ and c represent the density and the speed of sound, respectively. The parameter k is the acoustic wave number, which equals $k = 2\pi/\lambda = \omega/c$ with λ being the wavelength and ω the driving frequency, i.e. the physical frequency multiplied by 2π . Furthermore, $u(\xi)$ is the particle velocity at the transducer element at the given transducer point ξ , and \mathbf{S} is the transducer surface area. Note that the particle velocity is proportional to the acoustic intensity emitted by the transducer. For details, see for instance the book of Hill et al. [HBtH02]. If \mathbf{S} describes a rectangular surface area centered at the origin, and the transducer has a uniform particle velocity profile u_0 , the integral can be formulated as a sum of four one-dimensional integrals [ZM09]. The pressure is then

$$p_0(\mathbf{x}) = -\frac{u_0}{2\pi} \rho c e^{i\omega t} \left(-s_1 \int_{l_1}^{l_2} \frac{e^{-ik\sqrt{z^2+\sigma^2+s_1^2}} - e^{-ikz}}{\sigma^2 + s_1^2} d\sigma \right. \\ + s_2 \int_{l_1}^{l_2} \frac{e^{-ik\sqrt{z^2+\sigma^2+s_2^2}} - e^{-ikz}}{\sigma^2 + s_2^2} d\sigma \\ - l_1 \int_{s_1}^{s_2} \frac{e^{-ik\sqrt{z^2+\sigma^2+l_1^2}} - e^{-ikz}}{\sigma^2 + l_1^2} d\sigma \\ \left. + l_2 \int_{s_1}^{s_2} \frac{e^{-ik\sqrt{z^2+\sigma^2+l_2^2}} - e^{-ikz}}{\sigma^2 + l_2^2} d\sigma \right) \quad (2)$$

with $\mathbf{x} = (x, y, z)$ and $s_1 = x - w/2$, $s_2 = x + w/2$, $l_1 = y - h/2$, $l_2 = y + h/2$ with w and h being the width and height of a transducer element facing in positive z direction and centered at zero.

We can integrate damping into (1) and (2) by replacing k with $\tilde{k} = k - i\alpha$. Here, α is a damping coefficient with the unit $1/\text{m}$. Since this would require several additional numerical instructions in each of the single integrals and the damping effect in our application is not significant close to a single transducer element, our implementation approximates damping by pulling it out of the integral, thereby uniformly damping the contributions of a single transducer element:

$$\tilde{p}_0(\mathbf{x}) = p_0(\mathbf{x}) \cdot e^{-\alpha|\mathbf{x}|}.$$

Assuming a transducer that is set up of the same kind of transducer elements, e.g. a planar phased array system, we can pre-compute a complex pressure volume $p_0(\mathbf{x})$ for a single transducer element. This volume can then be used to determine the superposition of all transducer elements by

applying the inverse transformation matrix describing the transducer element's position and orientation. The total pressure in the nearfield is then

$$p(\mathbf{x}) = \sum_{n=1}^N p_0(T_n^{-1}(\mathbf{x})) A_n e^{i\varphi_n} e^{-\alpha|T_n^{-1}(\mathbf{x})|},$$

where N is the number of transducer elements, φ_n describes the phase shift, and T_n the transformation matrix of each element. The contribution of each single transducer element to the pressure field can be weighted using the so-called apodization weight A_n . Given the fact that we only want to compute the nearfield solution for the interface, we first determine the smallest possible bounding box applicable for the pressure volume p_0 such that for each sample point \mathbf{x} in the interface $T_n^{-1}(\mathbf{x})$ is in the bounding box for each transducer element n . Note that this step is especially important for non-planar phased array system, e.g. if the transducer elements are arranged on a sphere.

3.2. Farfield

For a homogeneous medium, the Rayleigh-Sommerfeld approach can also be used in the farfield. However, the computational costs are then in $\mathcal{O}(N)$ for each sample point, where N is the number of transducer elements. Therefore, more efficient techniques are preferable. Especially, the angular spectrum method has a computational complexity that does not depend on N in the farfield, and thus it is well-suited for our application. The core idea is to determine the pressure field in the frequency domain – more precisely to determine a slice of the pressure field in the frequency domain from the spectrum of the previous slice. The Fourier transform in x and y of the 3D Helmholtz equations yields a one-dimensional Helmholtz equation, for which a solution can be determined analytically [CH02]. As suggested by Zeng and McGough [ZM08, ZM09], for media with low absorption the angular spectrum approach can be implemented as follows: Based on the angular spectrum of the input slice $p(x, y, z)$ with $z = z_i$, denoted by $P(k_x, k_y, z_i)$, one can derive the so-called spectral propagator $H_p(k_x, k_y, \Delta z)$ to determine the angular spectrum of the output slice $z_o = z_i + \Delta z$ by

$$P(k_x, k_y, z_o) = P(k_x, k_y, z_i) \cdot H_p(k_x, k_y, \Delta z) e^{\frac{-\alpha k \Delta z}{\sqrt{k^2 - k_x^2 - k_y^2}}}, \quad (3)$$

where

$$H_p(k_x, k_y, \Delta z) = \begin{cases} e^{-i\Delta z \sqrt{k^2 - k_x^2 - k_y^2}}, & k_x^2 + k_y^2 \leq k^2 \\ e^{-\Delta z \sqrt{k_x^2 + k_y^2 - k^2}}, & k_x^2 + k_y^2 > k^2 \end{cases} \quad (4)$$

The different cases in (4) correspond to propagating and evanescent waves. A thorough discussion on these cases can be found in [HBtH02]. This yields to a final algorithm that starts at the interface layer and then determines successive slices in z direction one after another using the spectral propagator.

3.3. Inhomogeneous Media

Based on the work of Vyas and Christensen [VC08], the angular spectrum approach can be extended to account for inhomogeneous media. The idea is to combine the spectral propagation with a material correction which is consequently performed in the spatial domain, and which accounts for the material inhomogeneities from the input to the output slice. We first use the homogeneous spectral propagator to compute the output slice, where we choose average material parameters. Second, we correct the pressure field obtained in the spatial domain for the varying density, absorption, and speed of sound. Let ρ_h , c_h , and α_h denote the material parameters used in the spectral propagator, and let ρ_m , c_m , α_m denote the material at the current sample point (in the spatial domain). If Δz is the slice thickness, then the “homogeneous” pressure $p_h(x, y, z)$ at this point is corrected as follows:

$$p_m(x, y, z) = p_h(x, y, z) e^{i\Delta\varphi} \cdot \tau \cdot e^{(\alpha_h - \alpha_m)\Delta z}$$

with

$$\Delta\varphi = \left(\frac{c_h}{c_m} - 1.0 \right) \frac{-\omega \Delta z}{c_h}$$

$$\tau = 1.0 + \frac{\zeta_2 - \zeta_1}{\zeta_2 + \zeta_1},$$

where $\zeta_2 = \rho_m c_m$ is the acoustic impedance of the current material and $\zeta_1 = \rho'_m c'_m$ is the acoustic impedance of the sample point on the previous slice. The factor τ describes the fraction of the waves that are transmitted at the interface of media with impedance ζ_1 and ζ_2 [Lei07]. It is worth noting that the reflected waves are ignored at this time. Due to the relative high absorption in the tissue reflected waves are most relevant on bone structures close to the transducer. However, in this case reflected waves do not have a strong influence on the focus area, and thus they can be ignored. Nevertheless, we plan to extend the approach in future to incorporate reflecting waves in a spectral propagator that sweeps in opposite z direction.

3.4. Bioheat Transfer

The transfer of heat caused by the acoustic waves propagating through the tissue is modeled by the bioheat equation

$$\bar{c}\rho \frac{\partial}{\partial t} T(\mathbf{x}, t) - \nabla \cdot (\kappa \nabla T(\mathbf{x}, t)) + \nu_b (T(\mathbf{x}, t) - T_b) = f(\mathbf{x}).$$

Here, $T(\mathbf{x}, t)$ denotes the tissue temperature and T_b is the constant blood temperature. Moreover, \bar{c} and κ represent the specific heat capacity and the thermal conductivity of the tissue, respectively, and $\nu_b := w_b \bar{c}_b \rho_b$ is the product of the blood perfusion rate and the specific heat capacity as well as the density of blood. The heat source induced by the ultrasound waves is given by [ZM09]

$$f(\mathbf{x}) = \frac{\alpha}{\rho c} p(\mathbf{x}) p^*(\mathbf{x}).$$

If we consider inhomogeneous media, the set of material parameters depends on \mathbf{x} . Therefore, the bioheat transfer equation for (in)homogeneous media in three space dimensions reads

$$\bar{c}\rho T_t - (\kappa_x T_x + \kappa_y T_y + \kappa_z T_z) - \kappa(T_{xx} + T_{yy} + T_{zz}) + \nu_b(T - T_b) = f. \quad (5)$$

To shorten notation, the explicit time and space dependencies of the variables and parameters are omitted, and the derivatives of $\kappa(\mathbf{x})$ and $T(\mathbf{x}, t)$ with respect to x, y, z , and t are denoted by κ_x , T_x et cetera. To complete the formulation, we assume the initial temperature as well as the boundary temperature to be body temperature T_{body} , i.e.

$$T(\mathbf{x}, 0) = T_{body}, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^3 \quad (6)$$

$$T(\mathbf{x}, t) = T_{body}, \quad \mathbf{x} \in \partial\Omega. \quad (7)$$

Here, Ω refers to the farfield, and $\partial\Omega$ denotes all boundary values of the farfield, i.e. not only the interface layer.

We solve Equations (5)-(7) numerically by means of an explicit finite difference method, where we approximate the first and second spatial derivatives by means of the backward and the central difference quotient, respectively. In this way, we can directly solve Equation (5) for the unknown temperature at the current time step.

4. GPU Implementation

The GPU implementation performs four tasks. First, for a single transducer element we determine the pressure field by numerically evaluating the Rayleigh-Sommerfeld integrals (2). Second, we compute the superposition of all transducer elements at the initial (interface) slice of our simulation domain. Third, we use the hybrid angular spectrum approach to determine the pressure in the farfield. Last, we solve the dynamic bioheat equation with the source terms derived from the pressure field.

4.1. Phase/Magnitude Interpolation

Our goal is to store a pre-computed pressure field for a single transducer element. First, for a given transducer setup, we determine the minimal axis-aligned bounding box that we can use to pre-compute the pressure volume for the transducer elements. We determine this box by iterating over all transducer elements, and for each transducer element we enlarge the box such that the complete interface layer is in the interior of the bounding box. Note, that for homogeneous simulation using the Rayleigh-Sommerfeld approach in the farfield we use the whole simulation domain instead of only the interface at this point. For this bounding box, we pre-compute the pressure field p_0 using Equation (2), and we store the result in a 3D image. We use Gauss quadrature to numerically approximate the Rayleigh-Sommerfeld integral using 20 sample points as suggested by Zeng and McGough [ZM09].

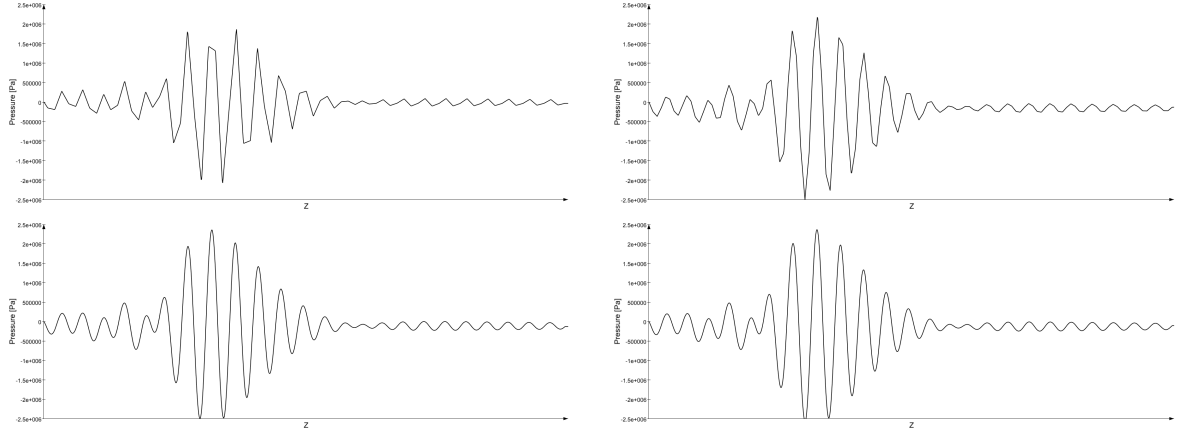


Figure 3: Analysis of the pressure profile on a ray in z direction through the focus area using different interpolation schemes to compute the interference of the waves. In the top image row, standard interpolation techniques (interpolation of real and imaginary part) are used while the bottom row shows our improved interpolation scheme (interpolation of phase and magnitude). The first column uses 3 samples per wavelength in the precomputed volume p_0 , while the second column uses 5. Both of these resolutions are not sufficient to compute the interference of the waves adequately using standard interpolation techniques; note the loss of acoustic energy in the top row compared to the bottom row. However, using our improved interpolation scheme already 3 samples per wavelength are sufficient to achieve nearly energy preserving superposition of the waves.

This pre-computation is performed for single slices parallel to the transducer element's plane with increasing distance. In that way, we can easily store the magnitude and phase value of the respective complex pressure rather than the imaginary and real parts, and we can ensure that wrap arounds in the phase value are avoided by guaranteeing that it is a strictly monotonic increasing function of the distance to the transducer. This is accomplished by inspecting the phase values at the previous slice, which can be read in the GPU kernel program, and by adding 2π in case that the phase value is smaller on the current slice. Hereby, we assume that we have consistent phase values in the previous slice, which can be accomplished by applying the same idea in the x and y direction at the first (initial) slice. In this way, we can use tri-linear hardware interpolation on the resulting pressure volume p_0 and directly obtain an interpolated magnitude and phase value which characterizes the pressure field at the given sample point. Figure 3 and 4 illustrate the improvements achieved when interpolating magnitude and phase value rather than real and imaginary parts. Especially, it is shown in Figure 3 that by using only 3 sample points per wavelength in p_0 we still can compute the interference of the waves accurately with the new interpolation scheme. Consequently, we can determine the resulting heat dissemination in the tissue from only a few samples per wavelength, which is in contrast to previous approaches which state that at least 10 samples per wavelength are required. However, it is worth noting that this interpolation approach fails if the distance of the sample points is larger than a wave length of the pressure field. However, in this case the grid resolution is anyway too coarse as to represent the pressure field appropriately.

4.2. Rayleigh-Sommerfeld Kernel

We use a kernel program to compute the superposition of the waves emitted by the transducer elements. This kernel is invoked for all voxels of the interface layer. We store the number of transducer elements as well as the transducers' transformation matrices in an OpenCL buffer object. Then, the kernel program iterates over all transducer elements, determines the position of the current interface voxel in local transducer coordinates, and looks up the pressure magnitude and phase value, which are then used to superpose the waves. This lookup benefits from the use of tri-linear hardware interpolation in the 3D image. These memory operations are local, i.e. the voxels of the interface that are processed on the same multiprocessor are likely to access the pre-computed pressure volume at nearby voxel positions, in which case the values required for tri-linear interpolation can likely be served from the GPU's texture cache.

4.3. Angular Spectrum Kernel

Since the angular spectrum approach depends on an input slice, we build kernels that perform the computations for one of these slices, which leads to the invocation of several kernels for the respective slice. Since the quality of the angular spectrum method depends on the discretization of the spectrum, we perform all computations on slices with double size, i.e. for a 128^3 simulation domain we process 128 slices of size 256^2 . From an implementation point of view, one important aspect here is the implementation of the fast fourier transform (FFT). For the timings we have measured, we have used a radix-2 implementation in OpenCL [Bai10], although slightly faster approaches exist; we plan to include these in

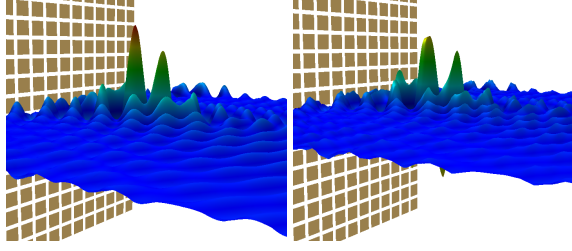


Figure 4: By storing magnitude and phase shift (left) rather than real and imaginary part (right) in the pre-computed pressure field, the tri-linear hardware interpolation on the GPU gives significantly better results for the superposed pressure field. Thereby, artifacts as they occur in the right image can be avoided by our approach (left).

future. The implementation is slice-wise, i.e. the slices are computed consecutively starting with the slice next to the interface, and each computation includes the forward FFT, the spectral propagator as denoted in Equation (4), the inverse FFT, and the material correction in the spatial domain as introduced in Section 3.3. Throughout the implementation of all these kernels, we use two 2D image objects in a ping-pong fashion, where one image is used as write-target, while the other one is used as input parameter. Moreover, the kernel used to compute the material correction in the spatial domain requires two slices including the material parameters – the current slice as well as the previous slice in order to account for the transmission factor τ .

4.4. Bioheat Equation Kernel

We solve the dynamic bioheat equation using an explicit Euler time integration scheme; therefore, we can update all temperature values in the simulation grid simultaneously. First, we determine the heat sources in an initial kernel pass from the pressure volume computed previously. Then, we invoke the bioheat kernel for the whole simulation domain, and we use two copies of the temperature volume which are used in a ping-pong fashion as read- or write-target, alternately. Material parameters are read in this kernel using a volume containing the material classification in combination with a lookup table for the material parameters for each of the material IDs. To respect the boundary conditions, we update only non-boundary voxels in the simulation domain.

5. Results

We analyze the performance of our implementation on a standard desktop PC (see Table 1), equipped with an Intel Core i7 3.0 GHz processor, 12 GB of RAM, and an NVIDIA GeForce GTX 480 graphics card with 1536 MB of local video memory. We use a planar transducer consisting of $17 \times 17 = 289$ quadratic transducer elements with a width of

Resolution	Intel Core i7		NVIDIA GTX 480		
	128 ³	256 ³	128 ³	256 ³	512 ³
p_0	11.3	87.4	0.082	0.469	3.24
RS	133	1043	0.180	1.44	11.5
AS	Init	1.07	4.10	0.003	0.007
	Prop.	1.18	9.71	0.175	0.876
Bioheat	37.3	304	0.167	1.28	10.6

Table 1: Timing statistics for the simulation of ultrasound propagation and heating in tissue (in s). Here, p_0 denotes the pre-computation of a pressure volume for the Rayleigh-Sommerfeld approach. RS denotes the homogeneous simulation of the pressure in the farfield by using the Rayleigh-Sommerfeld approach. AS denotes the heterogeneous simulation in the farfield using the angular spectrum method. Bioheat refers to the solution of the dynamic bioheat equation using 200 time steps.

1.8 mm, a physical frequency of 0.5 MHz and an initial particle speed $u_0 = 0.2 \frac{\text{m}}{\text{s}}$. We simulate the pressure field and temperature field based on a segmented MR data set of the human abdomen. The specific material parameters involved in the simulation are listed in Table 2. The results of the simulation using a 512³ grid are shown in Figure 6.

Our results are summarized in Table 1. For different resolutions of the simulation domain, we have measured the time required by the single steps of our approach using an optimized single-core CPU implementation as well as the parallel GPU implementation. First, we have measured the time to determine the transducer pressure volume p_0 . For simplicity we assume that this initial volume has the same resolution as the simulation grid. However, especially for planar phased array systems far less samples are required in the z direction if the angular spectrum method is used, since only the interface voxels have to be covered by the initial pressure volume p_0 . Then, we list the time required by the homogeneous Rayleigh-Sommerfeld method in the farfield. Next, we show the time required by the angular spectrum method assuming inhomogeneous media, where we split the timings into the computation of the pressure at the interface (Init) and the slice propagation (Prop.). Note, that this is our intended approach for the farfield, and the Rayleigh-Sommerfeld method in the farfield is shown for comparison reasons. The last row in the table contains the time required to solve the bioheat transfer equation at the same resolution as the simulation domain (farfield). We choose 200 time steps of 0.1 s, i.e. we simulate a treatment of 20 s, yielding a temperature increase of about 25 degree Celsius in the focus point for our parameter settings.

From Table 1 it can be seen that the computation of p_0 is about two orders of magnitude faster than the respective

Material	$\rho \left[\frac{\text{kg}}{\text{m}^3} \right]$	$c \left[\frac{\text{m}}{\text{s}} \right]$	$\alpha \left[\frac{1}{\text{m}} \right]$	$\bar{c} \left[\frac{\text{J}}{\text{kg K}} \right]$	$\kappa \left[\frac{\text{J}}{\text{m s K}} \right]$	$\nu_b \left[\frac{\text{J}}{\text{m}^3 \text{s K}} \right]$
water	1,000	1,454	0.02	4,180	0.621	-
bone	1,900	3,800	50	1,370	0.435	0
liver	1,050	1,560	10	3,600	0.500	70,000

Table 2: Material parameters used for the simulation of ultrasound propagation and heat dissemination.

CPU variant using one CPU core. For the computation of the homogeneous Rayleigh-Sommerfeld in the farfield, the speedup is about 3 orders of magnitude due to the effective exploitation of the texture filtering units for the tri-linear interpolation. This reflects in the timings for the angular spectrum method, where the computations at the interface layer (Init) achieve the same speedups. However, due to the frequent memory accesses involved in the angular spectrum kernels the speedup for those is about 7 to 11 depending on the grid resolution. It is worth noting that the angular spectrum method is only about 10% faster for homogeneous media assumptions, in which case the material correction kernel can be omitted. Furthermore, for homogeneous media one observes that the angular spectrum method is about two orders of magnitude faster than the Rayleigh-Sommerfeld method on the CPU, while on the GPU a benefit of about a factor of 2 can only be observed for higher resolutions. For the bioheat transfer equations, we again achieve a speedup of about two orders of magnitude.

However, we want to mention that by exploiting multiple CPU cores a performance gain is achievable on the CPU. In our experiments, we observed speedups of 3 to 3.6 when exploiting all four CPU cores for the Rayleigh-Sommerfeld method as well as the solution of the bioheat equation. Still, the performance gain on the GPU is significant, which is due to the high number of cores available on the GPU as well as the texture filtering unit, which performs the tri-linear interpolation very efficiently in hardware.

5.1. Verification of Pressure Simulation

The accuracy of the implemented Rayleigh-Sommerfeld and angular spectrum methods in a homogeneous domain was verified by comparison with the analytical O'Neil solution for an ideal spherical fixed-focus transducer [O'N49]. The transducer had a focal length of 50mm and a surface area of 2102mm². It was approximated by 65 planar, quadratic elements of side length 5mm arranged on a spherical surface, where the smaller total transducer surface was accounted for by a factor 0.773 in the pressure amplitude. Results are shown for a frequency of 0.25MHz in Figure 5. A good correlation between the Rayleigh-Sommerfeld and angular spectrum approach can be found in both the axial and radial profiles. Moreover, both profiles coincide quite well with the profiles of the ideal spherical transducer, although we approximate the transducer with only 65 planar transducer elements.

5.2. Software Prototype

The methods described above have been integrated into a software prototype allowing for virtual treatment planning and simulation of focused ultrasound (see Figure 1). The software is able to load MRI data according to the DICOM standard. In order to identify the different material densities an initial segmentation of several anatomical structures can be performed. The resulting multi-valued image mask is used to parameterize the simulation with given densities and spatial information of inhomogeneous tissue.

For the simulation workflow, focus points can be positioned with a single mouse click. To achieve this focusing, we adapt the phase shift value of each transducer element such that at the focus point the waves emitted by the elements are in phase under homogeneous media assumption. Then, we simulate the resulting pressure and temperature fields taking inhomogeneous media into account, which are shown as overlays above the original MRI data. Additionally, a 3D volume rendering is shown with the given transducer configuration (see Figure 7).

6. Conclusion

In this work we have presented an efficient method for the simulation of focused ultrasound. The method employs the power of current graphics hardware to speed-up state-of-the-art methods. We use the Rayleigh-Sommerfeld approach as well as the angular spectrum method to determine the pressure field in heterogeneous tissue, and we solve the bioheat transfer equation to predict the heating in the farfield. This information can be used to decide on the destruction of the tumor cells as well as to predict potential risks for the patient, e.g. the heating of the ribs. By taking the simulation results into account in a pre-operative planning stage, an optimal treatment plan can be derived for the patient, and thus the treatment can be improved.

We have shown that by means of our approach simulation times can be significantly reduced while keeping a high resolution of the simulation grid, and thus pre-operative planning becomes more feasible. This includes the determination of a series of sonications that guarantees to destroy the tumor without damaging healthy tissue. Since this process requires an optimal placement of sonications, an optimization approach to reach this heavily depends on efficient simulation techniques to derive a treatment plan in reasonable time.

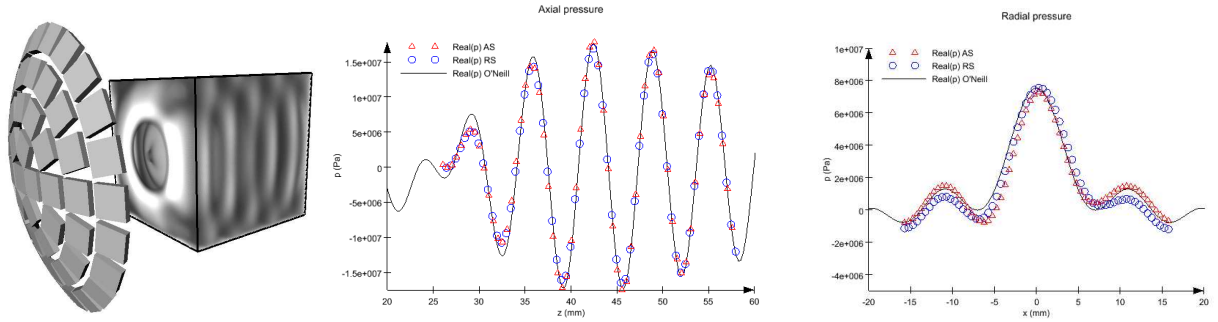


Figure 5: Axial (center) and radial (right) pressure profiles near the focus in a homogeneous medium compared to the analytical O'Neill solution for an ideal spherical transducer at frequency of 0.25MHz. We approximate this spherical transducer with 65 quadratic planar transducer elements as shown on the left.

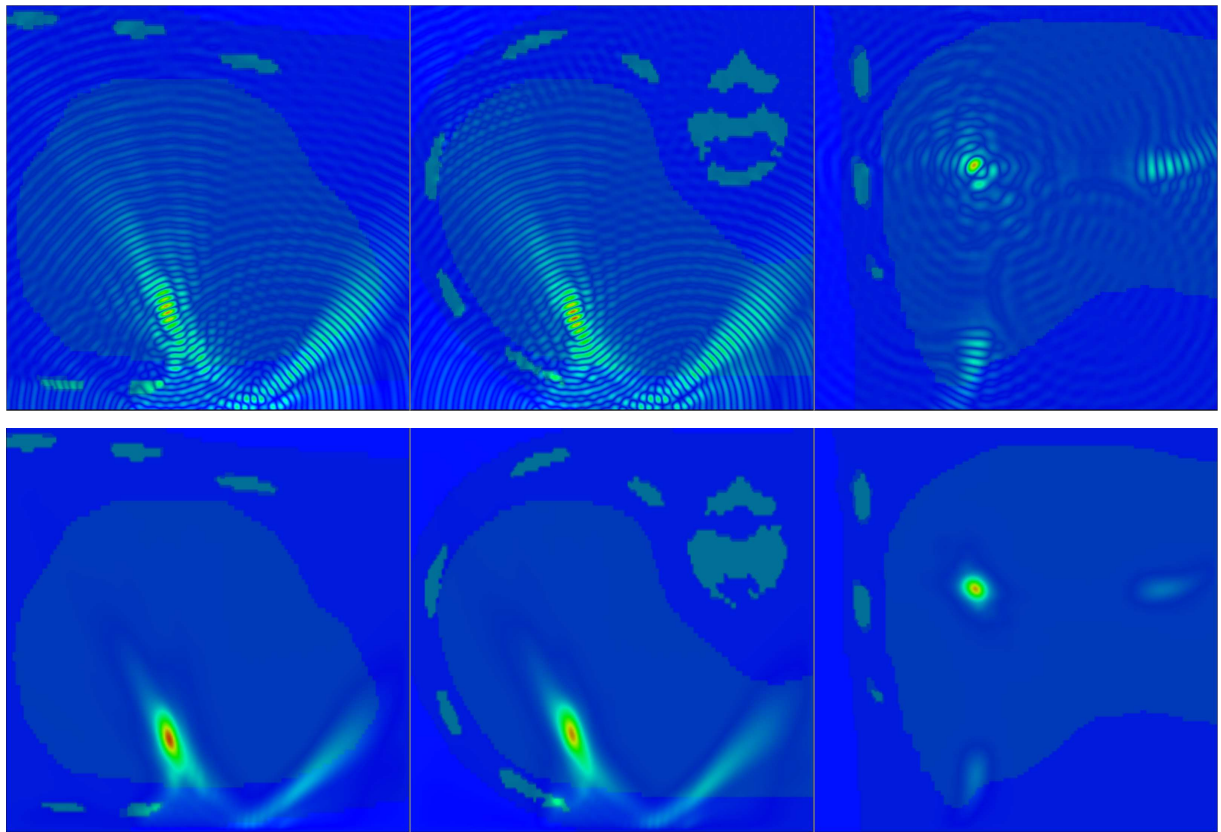


Figure 6: Visualization of the simulation results for a heterogeneous simulation using the proposed method. At the top, orthogonal slices through the pressure field are shown with a color map ranging from blue (0) to red ($2 \cdot 10^6$ Pa). At the bottom, the same orthogonal slices are shown in the temperature volume, where the color map shows temperature values from body temperature (blue) to 60°C (red). The segmented MR data set with bones and liver tissue is shown as overlay. Simulation results using a 512^3 grid have been received within 9 s (pressure) and 11 s (temperature). Note that heating occurs also close to the ribs due to the high absorption at the transition from tissue to bone.

Acknowledgements

This work was supported by the Fraunhofer Internal Programs under Grant No. MAVO 821 012.

References

- [ATP*03] AUBRY J.-F., TANTER M., PERNOT M., THOMAS J.-L., FINK M.: Experimental demonstration of noninvasive transskull adaptive focusing based on prior computed tomography scans. *Journal of the Acoustical Society of America* 113, 1 (2003), 84–93. [3](#)
- [Bai10] BAINVILLE E.: Opencl fast fourier transform, 2010. http://www.bealto.com/gpu-fft_intro.html. [6](#)
- [CH02] CLEMENT G. T., HYNYNEN K.: A non-invasive method for focusing ultrasound through the human skull. *Physics in medicine and biology* 47, 8 (2002), 1219–1236. [3](#), [4](#)
- [CMK*00] CURRA F. P., MOURAD P. D., KHOKHLOVA V. A., CLEVELAND R. O., CRUM L. A.: Numerical Simulations of Heating Patterns and Tissue Temperature Response due to High-Intensity Focused Ultrasound. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* 47, 4 (2000), 1077–1089. [3](#)
- [EH04] ENFLO B. E., HEDBERG C. M.: *Physical theory of non-linear acoustics*. Springer, 2004. [3](#)
- [HBtH02] HILL C. R., BAMBER J. C., TER HAAR G.: *Physical Principles of Medical Ultrasonics*. Wiley, 2002. [3](#), [4](#)
- [HG07] HOUSTON M., GOVINDARAJU N.: General-purpose computation on graphics hardware. In *SIGGRAPH 2007 GPGPU Course* (2007). [2](#)
- [HMK*05] HUTTUNEN T., MALINEN M., KAIPIO J. P., WHITE P. J., HYNYNEN K.: A full-wave helmholtz model for continuous-wave ultrasound transmission. *IEEE Trans Ultrason Ferroelectr Freq Control* 52, 3 (Mar 2005), 397–409. [3](#)
- [Jol09] JOLESZ F.: MRI-guided Focused Ultrasound Surgery. *Annu Rev Med* 60 (2009), 417–430. [1](#)
- [Khr11] KHRONOS OPENCL WORKING GROUP: The OpenCL Specification v1.1, 2011. <http://www.khronos.org/registry/cl/specs/opencl-1.1.pdf>. [2](#)
- [KSN09] KUTTER O., SHAMS R., NAVAB N.: Visualization and gpu-accelerated simulation of medical ultrasound from ct images. *Comput. Methods Prog. Biomed.* 94 (June 2009), 250–266. [3](#)
- [Lei07] LEIGHTON T. G.: What is ultrasound? *Progress in biophysics and molecular biology* 93, 1-3 (2007), 3–83. [3](#), [5](#)
- [LHHH10] LIU H.-L., HSU C.-L., HUANG S.-M., HSI Y.-W.: Focal beam distortion and treatment planning for transrib focused ultrasound thermal therapy: a feasibility study using a two-dimensional ultrasound phased array. *Med Phys* 37, 2 (Feb 2010), 848–860. [3](#)
- [LLZG07] LI J.-L., LIU X.-Z., ZHANG D., GONG X.-F.: Influence of ribs on the nonlinear sound field of therapeutic ultrasound. *Ultrasound Med Biol* 33, 9 (Sep 2007), 1413–1420. [3](#)
- [MCtH00] MEANEY P. M., CAHILL M. D., TER HAAR G. R.: The intensity dependence of lesion position shift during focused ultrasound surgery. *Ultrasound in medicine & biology* 26, 3 (2000), 441–450. [3](#)
- [NVI08] NVIDIA: NVIDIA CUDA Compute Unified Device Architecture, Programming Guide (v. 2.2), 2008. <http://www.nvidia.com/cuda>. [2](#)
- [NVI09] NVIDIA: NVIDIA's Next Generation CUDA Compute Architecture: Fermi, 2009. http://www.nvidia.com/content/PDF/fermi_white_papers/NVIDIA_Fermi_Compute_Architecture_Whitepaper.pdf. [2](#)
- [OHL*08] OWENS J., HOUSTON M., LUEBKE D., GREEN S., STONE J., PHILLIPS J.: Gpu computing. *Proceedings of the IEEE* 96, 5 (may 2008), 879–899. [2](#)
- [OLG*05] OWENS J. D., LUEBKE D., GOVINDARAJU N., HARRIS M., KRÜGER J., LEFOHN A. E., PURCELL T.: A survey of general-purpose computation on graphics hardware. In *Eurographics 2005* (2005). [2](#)
- [O'N49] O'NEIL H. T.: Theory of focusing radiators. *Journal of the Acoustical Society of America* 21, 5 (1949), 516–526. [8](#)
- [RPAS09] REICHL T., PASSENGER J., ACOSTA O., SALVADO O.: Ultrasound goes gpu: real-time simulation using cuda. *Proceedings of SPIE* 7261, 1 (2009), 726116–726116–10. [3](#)
- [SMS*09] SCHENK O., MANGUOGLU M., SAMEH A. H., CHRISTEN M., SATHE M.: Parallel scalable pde-constrained optimization: antenna identification in hyperthermia cancer treatment planning. *Computer Science - R&D* 23, 3-4 (2009), 177–183. [3](#)
- [VC08] VYAS U., CHRISTENSEN D.: Ultrasound beam propagation using the hybrid angular spectrum method. In *Proc. of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. (2008), pp. 2526–2529. [3](#), [5](#)
- [WFZ09] WANG S., FRENKEL V., ZDERIC V.: Preliminary optimization of non-destructive high intensity focused ultrasound exposures for hyperthermia applications. *Conf Proc IEEE Eng Med Biol Soc 2009* (2009), 3055–3059. [3](#)
- [ZM08] ZENG X., MCGOUGH R. J.: Evaluation of the angular spectrum approach for simulations of near-field pressures. *Journal of the Acoustical Society of America*, 1 (2008), 68–76. [3](#), [4](#)
- [ZM09] ZENG X., MCGOUGH R. J.: Optimal simulations of ultrasonic fields produced by large thermal therapy arrays using the angular spectrum approach. *The Journal of the Acoustical Society of America* 125, 5 (2009), 2967–2977. [3](#), [4](#), [5](#)

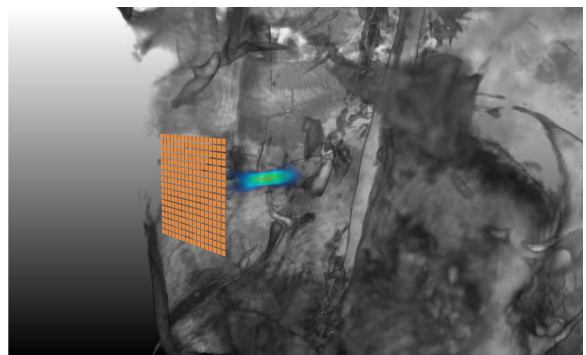


Figure 7: 3D visualization of a simulated sonication in the MRI data set.