Published in final edited form as:
J Neurosci Methods. 2015 January 30; 240: 28–36. doi:10.1016/j.jneumeth.2014.10.018.

Towards Real-Time Detection of Seizures in Awake Rats with GPU-Accelerated Diffuse Optical Tomography

Tao Zhang[a], **Junli Zhou**[b], **Paul R. Carney**[a,b,c,d], and **Huabei Jiang**[a,*]

[a]J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611, USA

[b]Department of Pediatrics, University of Florida, Gainesville, FL 32611, USA

[c]Department of Neurology, University of Florida, Gainesville, FL 32611, USA

[d]Department of Neuroscience, University of Florida, Gainesville, FL 32611, USA

Keywords

Real-time diffuse optical tomography; Noninvasive seizure dynamic imaging; GPU computing

1. Introduction

Diffuse optical tomography (DOT) is an emerging imaging modality based on the scattering and absorption properties of nonionizing near-infrared light in biological tissue [1,2]. As an extension of the established functional brain imaging tool near infrared spectroscopy (NIRS), DOT not only shares the merits of providing continuous readings of cerebral oxygenation, but also has the ability to provide spatial resolution in the millimeter scale[3]. DOT has been successfully applied in the imaging of breast tumors[4], osteoarthritis[5] and cortex activations[6]. Recently DOT has been introduced in more dynamic neuroimaging studies of cerebral perfusion[7] and brain network analyses[8].

Despite pharmacological, surgical, and neuromodulation advances in the treatment of chronic epilepsy, seizures are often not controlled in as many as 25 percent of patients. A growing body of research indicates that preventing seizures may be possible if the earliest stages of seizure initiation could be reliably identified by some means[9,10]. To this end, DOT, which is capable of detecting hemodynamic responses spatially equivalent to fMRI[11] with

*Corresponding author at: J. Crayton Pruitt Family Department of Biomedical Engineering, Gainesville, FL, 32611, USA. Phone: +1(352)273-9336. Fax: +1(352)273-9221. hjiang@bme.ufl.edu.

Author contributions

T.Z. developed the imaging system and paralleled codes, designed and performed experiments, analyzed data and wrote the paper; J.Z. performed experiments; P.R.C. guided the experimental design and wrote the paper; H.J. directed the study and wrote the paper.

Competing financial interests

The authors declare no competing financial interests.

high sampling rate (>1Hz), may provide the necessary spatial and temporal resolution to identify impending seizures[12]. It is known that activity evoked optical spectroscopic changes measured in the brain are most likely generated by early changes in cerebral hemodynamics[13]. Therefore, such optical property changes are also the hallmarks of seizure initiation and propagation. A deeper understanding of hemodynamic changes at the beginning and throughout the evolution of epileptic seizures can help identify the most susceptible brain regions[14]. With such knowledge, preventative techniques could be developed and then coupled with treatment strategies in order to interrupt the process before seizure onset.

In our previous study with simultaneous DOT and IEEG in anesthetized rats, the pre-seizure state was investigated and uniformly identified[15]. We found that the onset of generalized seizures originates locally and early hemodynamic responses occur for several minutes before seizure onset. However, when introduced to the study of seizure prevention, the same challenge to immobilize the subject that affects other traditional neuroimaging tools like SPECT[16] and fMRI[17] arose. This brings the requirement to anesthetize animal subjects in order to achieve relatively long time stable recordings. Studies have demonstrated that hemodynamic responses elicited by seizure-induced drugs may differ between anesthetized and unanesthetized animals. This is due to the fact that anesthetics may act on various neurotransmitters and sensory feedback from the whole body contributes to the seizure threshold[18]. Ideally, experiments in non-paralyzed and fully conscious animals could overcome the profound impact of anesthetics on cerebral hemodynamics and better reflect neural activity in realistic situations.

Another challenge to applying DOT for early seizure detection is the requirement of having a timely response which refers to the real-time availability of the reconstructed images derived from hemodynamic changes. Finite-element based nonlinear iterative algorithm has been a powerful reconstruction method for obtaining high quality DOT images[19]. However, it is computationally demanding especially in three-dimensional (3D) cases with high density optode arrays. The calculation for one reconstruction requires several to tens of minutes depending on the imaging domain size, making it an unrealistic task in real-time applications and therefore limited to retrospective analysis.

The primary goals of this study were to confirm that pre-seizure state related hemodynamic changes can be detected in awake rats with DOT, and to perform real-time 3D monitoring with our reconstruction procedure. To this end, first we developed methods to conduct seizure experiments in fully awake rats using a subject-specific helmet built with 3D scanning and printing, and a restraining mechanism based on hanging wraps. Second, graphic processing units (GPU) based parallel code was implemented using a CUDA programming model (NVIDIA), which allows, in principle, for online reconstructions. As these proof-of-concept methods and results are extended, we expect that DOT can become an important tool in the study for the prevention of epileptic disorders.

2. Methods

2.1 Animals

Five adult male Sprague-Dawley rats (Harlan Labs, Indianapolis) weighing ~400 g were used for the *in vivo* experiments. Animals were housed in pairs in a controlled environment (12:12-h light/dark cycle; food and water ad lib). All the experimental protocols and procedures involving animal care were conducted in conformity with the standards of the NIH and IACUC committee at the University of Florida.

Prior to imaging, rats were anaesthetized under inhalational anesthesia maintained at 2% isoflurane mixed with 0.4 L/min oxygen during the experiment preparation, which included shaving the heads of the rats with hair removing lotion, placing them on the hammock, strapping on the helmet, and plugging in all the fiber bundles. Then the rats rested for one hour without anesthesia, allowing recovery to a fully conscious state.

The generalized tonic-clonic seizures were induced by an IP injection of 50 mg/kg of the GABA-antagonist pro-convulsant pentylenetetrazol (PTZ) (Sigma-Aldrich, Inc.)[20]. Time-locked EEG/DOT were performed continuously for 30 minutes surrounding each PTZ or saline (control experiments) injection, 5 minutes before injection for the resting state recording, and then 25 minutes after injection for the seizure recording. A total of 3 rats received PTZ, while 2 rats were used as age-matched sham controls.

2.2 Restraining mechanism and training procedure

We implemented a restraining mechanism that allowed for the highest level of comfort possible[21], since the rat stayed awake during the relatively long time recording in the seizure experiment. For restraint, we used both hammocks and straight-jacket type wraps to better fit our needs. The jacket can kept the animals warm and with some Velcro on it, allowed for a tight fit and improved immobility. The use of a hanging wrap (hammock) provided rats with more comfortable and relaxed conditions (Fig. 1d). The hammock, consisting of a sheet bottom with leg holes and a laced top, created a tight fit around the rats, keeping them from wiggling extensively. To further lessen the weight on the rat head, the helmet was tied to the supporting bars on both sides.

Rats were habituated to immobilized conditions in hammocks 2 weeks before the experiments. The aim of the training was to achieve long term immobilization while reducing the stress of the rats. By observing the behavioral responses and gradually increasing the restraint time (starting at 15 minutes and continuing up to 2 hours daily) as rats exhibited positive behaviors, long restraint times were achieved in order to evaluate how calm the animals remained. Only the best-behaved rats after training were selected for our experiments.

2.3 Head modeling and mesh generation

A handheld laser scanner (ZScanner 700) was used to acquire the accurate structural information of the shaved rat head. This 3D scanner has a high resolution of 0.1mm in xyz directions, and stores the surface information in STL format. Rhino 3D was used to load this

STL file and design the 3D helmet model. First, raw surface mesh (Fig. 1a) was repaired by removing overlapped surfaces, filling up holes, and smoothing (Fig. 1b). Then additional cylindrical holders for optical fiber bundles, holes for EEG electrodes, and fixing edge structures were added in the model (Fig. 1c). As shown in Fig. 1c, the helmet consists of two parts. The upper part has 16 holders and the lower part has 8 holders. The distance between two holders is ~9 mm. The two parts are connected by four screws on each side of the helmet. An anchoring screw is placed in the upper part of the helmet to prevent any sliding between the scalp and helmet. It also serves as a stereotaxic marker and EEG reference electrode. A half-cylindrical shaped slot is added in the lower part to relieve pressure on the vagus nerve. Otherwise the helmet wall would suppress breathing when placed on the rat. To ensure the helmet was tightly attached to each rat, two additional helmets of smaller size were derived from the original model, shrinking them by 1 mm and 2 mm. We chose the most suitable helmet for each animal according to their respective head sizes. The helmet was fabricated by a 3D printer (Objet Eden 260V) with a semi-transparent, light weight resin material, offering highly refined detail and model durability. The 2 mm thin-walled shell of the helmet allows some flexibility while retaining sufficient strength. The whole helmet is ~20 grams, causing little burden on the rat.

The structural information was also used to generate a 3D mesh using Netgen for the finite element method based reconstruction (Fig. 3a). The mesh consists of 9972 elements of 1.8 mm, and 2457 nodes. The positions of the 24 holders for optodes are determined by comparing the actual position of each holder in the model with the nodes in the mesh.

2.4 Synchronized DOT and EEG system

We used a fast multispectral DOT system for optical signal recording, as described in detail previously[22] (Fig. 2). The system performs continuous-wave (CW) measurements using 72 light emitting diodes (LEDs) as the excitation sources through a time multiplexing approach and highly sensitive photodetectors working in parallel for light detection. Twenty four optical source/detector pairs were distributed along the helmet and attached firmly to the scalp, sequentially shining light upon the rat head and collecting the diffused scattered light. This DOT system allows for a full set of 3D data collection at 3 NIR wavelengths (660, 780, and 850 nm) for whole brain imaging in 14.4 Hz. The data acquisition for both DOT and EEG recording was synchronized through a controller board (PCI-7811R, NI).

Continuous time-locked intracranial electroencephalogram (IEEG) recordings were used to confirm the occurrence of seizures. To monitor the local field potentials, multisite IEEGs were recorded using two microwire recording electrodes (50 μm polyimide insulated tungsten microwire) placed in the bilateral frontal cortices (Bregma coordinates: AP: 3mm; ML: ±2.5mm; DV: 1.5mm) through holes drilled into the skull. EEG data was down sampled from 50 kHz to 600 Hz and then band-pass filtered from 2 to 200 Hz. An electrocardiogram was also utilized to monitor the heart rate of the rats. Seizure onset was identified by visual analysis of electrographic morphology and its corresponding PSD for each rat PTZ experiment. The beginning of continuous high-amplitude and high-frequency spikes observed in both IEEG channels was considered as the onset of generalized seizures.

The data acquisition for both DOT and EEG recording was synchronized through a controller board.

2.5 GPU-based paralleled image reconstruction

Here we discuss our 3D reconstruction algorithm and its parallelization on GPU.

2.5.1 Forward Model—Our DOT algorithm is based on the finite element solution to the following continuous-wave (CW) photon diffusion equation coupled with Robin boundary conditions[1]:

$$\nabla \cdot \mathrm{D(r)}\nabla\Phi(\mathrm{r})-\mu_{\mathrm{a}}(\mathrm{r})\Phi(\mathrm{r})=-\mathrm{S(r)} \quad (1)$$

$$-\mathrm{D(r)}\nabla\Phi_{\mathrm{n}}=\alpha\Phi \quad (2)$$

where Φ(r) is the photon density at position r, $D(r)=1/3(\mu_a+\mu_s^{'})$ is the diffusion coefficient, $\mu_a(r)$ and $\mu_s^{'}(r)$ are absorption and reduced scattering coefficients, S(r) is the source term, and α is a coefficient related to the internal reflection at the boundary. The finite element forward solver we employed is based on the Galerkin weak form of Eq. (1), where Φ(r), $D(r)$, and $\mu_a(r)$ are spatially discretized as $\Phi=\sum_{i=1}^{N}\Phi_i \mathrm{l}_i, \mathrm{D}=\sum_{i=1}^{N}\mathrm{D}_i \mathrm{l}_i, \mu_{\mathrm{a}}=\sum_{i=1}^{N}(\mu_{\mathrm{a}})_i \mathrm{l}_i$, l_i is a spatially varying Lagrangian basis function. After discretization, the forward model can be expressed as a system of linear algebraic equations[1,23]:

$$[A]\{\Phi\}=\{b\} \quad (3)$$

where the global matrix A is a N*N sparse matrix, N is the number of nodes in the finite element mesh, Φ is a column vector of size N, and *b* is a column vector of size N with a nonzero value only at the measurement node and zero at other nodes. Thus, the forward problem of solving for Φ becomes the solving of a sparse linear system. Note that *A* is highly sparse and typically sparse matrix solvers are used to obtain Φ. In our experiment, an iterative conjugate gradient solver was used with preconditioning, which can speed the convergence of the solution. The compressed sparse row (CSR) format was employed to reduce the memory requirement for the storage of *A*.

2.5.2 Inverse Problem—In order to form an image, a regularized Newton's method is used to update an initial optical property distribution iteratively in order to minimize an objective function composed of a weighted sum of the squared difference between computed and measured optical data at the medium surface. The computed optical data is obtained by solving Eq. (4). The core procedure in our reconstruction algorithms is to iteratively solve the following regularized matrix equation:

$$\left(\mathrm{J^T J}+\lambda \mathrm{I}\right)\Delta \mathrm{q}=\mathrm{J^T}\left(\Phi^{(\mathrm{m})}-\Phi^{(\mathrm{c})}\right) \quad (4)$$

where I is the identity matrix, λ is the regularization parameter determined by combined Marquardt and Tikhonov regularization schemes, and $\Delta q = (\Delta D_1, \Delta D_2, \ldots, \Delta D_N, \Delta \mu_{a,1}, \Delta \mu_{a,2}, \ldots, \Delta \mu_{a,N})^T$ is the update vector for the optical property profiles. $\Phi_i^{(m)}$ and $\Phi_i^{(c)}$ are measured and calculated data for M boundary locations. J is the Jacobian matrix consisting of derivatives of Φ with respect to D or μ_a at each boundary observation node. For application with K sources, J is a (M*K, 2N) matrix, and $\Phi_i^{(m)}$ and $\Phi_i^{(c)}$ are vectors of size M*K. In DOT, the goal is to update μ_a and D distributions through the iterative solution of Eq. (4).

In practice, the Jacobian matrix is calculated using an efficient adjoint sensitivity method[24]. Basically, we need to solve an (N, M) adjoint sensitivity matrix Ψ, and the i^{th} column ψ_i of Ψ satisfies the following relationship:

$$[A]^T \{\psi_i\} = \{\Delta_d\} \quad (5)$$

where vector Δ_d has a unit value at the i^{th} measurement site and is zero at other nodes. Since Eq. (5) is also a sparse linear system, methods discussed to solve Eq. (3) in the forward model can be applied here. Then we can use Ψ to assemble the Jacobian matrix with the following equations[24]:

$$\{\partial \Phi / \partial D_k\} = [\Psi][\partial A / \partial D_k]\{\Phi\} \quad (6)$$

$$\{\partial \Phi / \partial \mu_k\} = [\Psi][\partial A / \partial \mu_k]\{\Phi\} \quad (7)$$

After reconstruction of absorption coefficients at the three wavelengths, concentrations of oxy-hemoglobin [HbO_2], deoxy-hemoglobin [Hb_R] and total hemoglobin [HbT] were obtained using the Beer-Lambert law[25], assuming HbO_2 and Hb_R are the two major chromophores in the rat brain.

2.5.3 Parallelization of the Reconstruction Algorithm—According to the discussion in the sections of the forward model and inverse problem, we notice that five modules require a large amount of calculations, including the sparse linear system solution for the forward model and the adjoint sensitivity method, assembly of the Jacobian matrix (J), matrix multiplication (J^TJ), and the dense linear system solution for Eq. (4). To identify and confirm the major computationally intensive modules in the reconstruction algorithm, the computation time in terms of percentage of total time taken per iteration using the mesh of rat head (2457 nodes, 24 sources/detectors) is indicated in Table 1.

In fact, four of these five modules, except for the assembly of the Jacobian matrix, are routine linear algebra operations. Several packages exist that provide options for carrying out such calculations on GPUs[26] utilizing the NVIDIA Common Unified Device Architecture (CUDA). In our study, we chose a GPU-accelerated linear algebra library CULA (http://www.culatools.com), featuring solvers for both dense and sparse problems. It consists of two toolboxes: “CULA dense” which provides accelerated implementations of the LAPACK and BLAS libraries for dense linear algebra and “CULA sparse” which

provides the tools necessary to rapidly solve large sparse systems using iterative methods. Interfaces of the device version were used whenever available to manually operate the data transfer between the CPU and GPU in order to avoid unnecessary data exchanges. Single precision solvers were also used due to their sufficiency for the diffuse optical tomographic image reconstruction[26], and ability to achieve more acceleration with less storage space, compared to the use of double precision solvers. The releases of CULA dense R16a and CULA sparse S4 based on CUDA 5.0 were used in this work. The GPU parallelization routine for each module is listed in Table 3.

To solve the sparse linear systems in the forward model and the adjoint method, a block conjugate gradient solver (culaDcsrCgBlockjacobi) with a block Jacobi preconditioner and CSR format was applied. We assembled a blocked version of Eq. (3) for all sources, in order to further take advantage of the parallel iterative solver and the blocked preconditioner:

$$\begin{bmatrix} A & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A \end{bmatrix} \begin{Bmatrix} \Phi_1 \\ \vdots \\ \Phi_K \end{Bmatrix} = \begin{Bmatrix} \mathrm{b}_1 \\ \vdots \\ \mathrm{b}_K \end{Bmatrix} \quad (6)$$

where K is the number of sources, and Φ_i and b_i are corresponding vectors in Eq. (3) for the i^{th} source. The same blocked version goes for Eq. (5) in the adjoint method. By this way, instead of solving for one right hand side (RHS) after another, we are able to solve simultaneously for all RHS.

For Eq. (4) in the inverse problem, a dense matrix multiplication solver (culaDeviceSgemm) was used for $\mathrm{J^T J}$ and a direct solver with Cholesky decomposition was applied to the dense linear system.

In reality, assembly of the Jacobian matrix is the most time consuming and memory intensive step in the current algorithm and there is no off-the-shelf library available. However, considering the high degree of locality for the calculation of $\partial\Phi/\partial D_k$ and $\partial \mathrm{A}/\mu_k$ and the high degree of sparsity of $\partial \mathrm{A}/\partial \mathrm{D}_k$ and $\partial \mathrm{A}/\partial \mu_k$, a significant speedup can be expected with a customized GPU-based code of massive parallelization. The calculation is divided into many parts and executed independently by a large number of threads in order to exploit data parallelism based on the two-level hierarchy of CUDA. The top level is a grid which consists of thread blocks. Thread blocks implement coarse-grained scalable data parallelism and are executed independently, which provides scalability to GPUs with a different number of processor cores. In the grid level, M*K blocks are generated, each calculating two rows of the Jacobian matrix ($\partial\left(\Phi_1^T,\ldots,\Phi_K^T\right)/\partial D_k$ and $\partial\left(\Phi_1^T,\ldots,\Phi_K^T\right)/\partial\mu_k$). In the lower block level, threads provide fine-grained data parallelism by synchronizing and using shared memory for communication with each other. Each block contains 32*16 threads (each thread block can consist of up to 1024 threads), where 32 corresponds to the maximum number of neighbors of each node, and 16 is the number of nodes calculated in this block. $\lceil N/16 \rceil$ kernels were carried out sequentially to determine all of the N nodes. Several optimization methods are used to achieve more acceleration, including storing variables in the fast shared memory to avoid accessing the high-latency global memory directly, pinning

memory to attain the higher bandwidth between the host and the device, and coalescing global memory access.

To achieve further acceleration, other miscellaneous tasks containing large loops, which did not take a significant amount of calculation time (Table 1), are also parallelized with customized codes using similar optimization as discussed above. Such tasks include the assembly of the global matrix A, assembly of the RHS in Eq. (4), CSR format transformation, and regularization related operations.

3. Results

3.1 GPU-Accelerated Simulations

We conducted simulations using the GPU-based parallel code with three different mesh sizes and compared them with the results using the CPU-based code. All simulations had the same number of 24 sources and 24 detectors with identical targets. Two targets of 3:1 absorption and 2:1 reduced scattering contrasts were placed in a homogeneous background (μ_a is 0.007 mm^{-1}, and μ_s' is 1 mm^{-1}), at the positions of 5 mm (close to the mouth) and 8mm below the top surface. The GPU and CPU code were written in C++ and both were compiled with Microsoft Visual Studio 2008 running on a Windows 7 64 bits system and 3.4GHz Intel CPU i7-2600. In order to test the computational performance of our proposed parallel method, the computational speed was benchmarked with three other types of GPUs, including a low-end GT 430 card with 96 cores, high-end GTX 580 card with 512 cores, and two multi-GPU GTX 590 cards (each containing two GTX 580 cards, but with slightly lower operating frequencies). In the multi-GPU application with four GPU cards installed on the same computer, we generated four instances from the GPU-based code and ran them simultaneously. According to our previous phantom and rat experiments[15,22], the 3D DOT reconstruction typically required 6 iterations to converge into a solution.

Fig. 3. gives the reconstructed absorption images from the two-target experiment with mesh 1 (Table 2). From the result, we can see that the targets can be clearly detected with a correct position, size, and optical property. The optical image obtained with the GPU-based code was visually identical to the image obtained using the CPU-based implementation.

A numerical comparison between the GPU-based and CPU-based results was conducted with the following equation and the relative differences of E listed in Table 2:

$$E=\frac{1}{N}\sum_{i=1}^{N}\frac{|R_{gpu,i}-R_{cpu,i}|}{R_{cpu,i}}*100\% \quad (7)$$

The relative difference (E) in all cases was less than 0.1%, asserting the fact that our single precision GPU-based code can carry out DOT image reconstruction with sufficient accuracy.

Table 2 listed the time taken on CPU and GPUs with different mesh sizes for the reconstruction of 6 iterations, and corresponding speedups, which is defined by the ratio between the time cost of the CPU and GPU. The acquisition time for one set of data is 0.07s,

which is not included in the time cost for image reconstruction. It is clear that GPU computing offers a significant speedup in comparison to CPU computing. For a smaller mesh, it only takes 2.4 s with a single GPU GTX 580 or 0.7 s on average when using multi-GPU GTX 590s to generate one image. The acceleration ratio increases with increasing mesh size. In fact, the speedup using a GTX 580 increases from 108x to more than 300x as the mesh size changes from 2457 nodes to 5851 nodes. The parallel computing capacity of GPU possesses an advantage when dealing with larger scale problems.

On the other hand, further acceleration has been achieved on GPUs with more cores, which demonstrates the scalability of our GPU-based reconstruction. Scalability is critical when developing parallel GPU applications. An adequately designed CUDA program will automatically utilize more parallelism on GPUs that have more processor cores. Thus, our application can automatically achieve more acceleration with the evolution of GPU, without any modifications of existing codes.

3.2 In vivo seizure experiments

Generalized tonic-clonic seizures and corresponding hemodynamic changes were observed in 3/3 seizure acute rat experiments. Similarly to our previous results in anesthetized rats, we found that significant hemodynamic responses were elicited during the generalized seizures, and preceded seizure onsets by tens of seconds in all rats. In the control experiments, no detectable IEEG spikes or significant hemodynamic changes were observed.

A representative result is shown in Fig. 4. The IEEG seizure onset recorded at the left primary motor (M1), featuring continuous spikes, began at about 70 seconds after PTZ injection (Fig. 4b). Sixteen positions located at regions (motor and somatosensory cortex, cortex, temporal lobe, and thalamus) that are closely related to the origin and spread of seizures[27] were selected from four coronal slices (Fig. 4a) ranging from −4mm to 4mm relative to bregma to show the hemodynamic changes (HbT, HbO_2, and Hb_R). The template MRI images were obtained from the Karolinska rat atlas (http://expmr.ki.se/research/ratatlas.jsp) for anatomical guidance. The significant hemodynamic responses were comprised of a prominent change of [HbO_2] and a smaller change of [Hb_R], which were also observed in other studies28,29. Significant hemodynamic changes started locally at the thalamus region (Fig. 4c, position 2 and 8) right after the PTZ injection and led to changes in other regions. Heterogeneous patterns of hemodynamic changes were observed in our experiments. In the left temporal lobe (Fig. 4c, positions 6, 10), a significant [HbT] decrease was observed, whereas an [HbT] increase appeared in other regions (Fig. 4c, positions 12, 16). Hemodynamic activities in the bilateral motor cortex (Fig. 4c, positions 1,3,5,7,9,11) were relatively stable compared to other regions in terms of absolute amplitude change. Depending on the interplay between hemodynamics and metabolism, increases of [HbT] could arise when the relative increase in CBF is more than the actual need for metabolism. On the other hand, a decrease of [HbT] suggests an exceedingly high metabolic demand without an appropriate compensatory blood flow especially around the seizure focus, rather than a decrease in neuronal activity[30].

4. Discussion

In this study, a subject-specific head interface and restraining mechanism was designed and successfully applied to the study of seizure activities in unanesthetized rats using DOT. A method for fast DOT image reconstruction during rapidly changing neural dynamics was also achieved. Results suggest that early hemodynamic changes could be observed tens of seconds prior to seizure onset in awake rats, and confirms previous results in anesthetized animals[15]. Finally, a GPU-based image reconstruction algorithm was developed and tested with significant acceleration that allows real-time online monitoring of the pre-seizure state.

For patients without complete seizure control, this sudden and unforeseen occurrence represents one of the most debilitating aspects of epilepsy. If it were possible to reliably predict seizure occurrences, prevention techniques could then be coupled with treatment strategies targeted at interrupting the process before seizures evolve. For patients dependent on daily anticonvulsant treatments, this would be a significant breakthrough as treatment would conceivably occur only when needed: on-demand and in advance of an impending seizure[31]. The existence of a pre-seizure state, supported with laboratorial and clinical findings of changes in cerebral blood flow and oxygen availability before seizure occurrence[32], could be used as a precursor to develop seizure prediction models. However, there still exist barriers in identifying the pre-seizure state and understanding the regional interactions during the seizure evolution in awake animals, especially with imaging tools of sufficient spatiotemporal resolution like DOT.

The first challenge of creating the experimental setting is to ensure that the animals remain well restrained and calm during the long term experiment. Unlike other methods that tend to introduce stressful conditions that result in discomfort over long periods of time, the restraining mechanism using a hammock and jacket allows the rats to make slight adjustments in body position if necessary for comfort while providing adequate restraint still. Many studies[21] involving awake, nonanethetized, and head-restrained rats have applied this approach and proved that the rats remained calm and exhibited fewer bad behaviors during the prolonged experiment.

Another challenge addressed was the act of designing an interface tightly connecting a number of fiber bundles to a rat head. One of the major technical advances in the past few years is 3D printing and medical research has already benefited tremendously from it. The precision, speed, and accuracy offered by 3D printing are the main reasons that it is being implemented widely in the fields of dentistry, surgery, and prosthetic development[33]. 3D printed interface has been used in optical imaging to study neurovascular coupling in awake rats[34]. However, surgery was needed to implant the flat plate on the skull, and only 2 sources and 8 detectors can be placed on this interface. In another study[35], using a precision machined cap with tether, simultaneous optical and electrical recordings were conducted on freely moving rats. The cap with 4 sources and 16 detectors still needed to be surgically attached to the skull, and only allowed DOT in reflection mode. In our study, a subject-specific 3D printing model acquired from 3D scanner is able to capture accurate structural details. The helmet produced with enhanced customization allows a better match between the helmet and head to provide more comfort, and tight attachment without the need of

surgery. The location and quantitative accuracy of reconstructed images can be improved with a large number of sources and detector fibers covering the whole head[36]. Also, the head-matched interface allows for faster image reconstruction due to the reduced imaging domain size, compared to the cylindrical shape interface used in our previous study[15].

Over the last couple of years, the rapid advance in the development of graphics processing units (GPU), originally designed for computer video cards, have emerged as the most powerful chips for high-performance computing. NVIDIA's Common Unified Device Architecture (CUDA) offers a unified hardware and software solution for parallel computing on GPUs with high performance computing numerical libraries. In the field of medical imaging, which often faces heavy computing burdens, GPU has been successfully applied in techniques like magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound imaging for image rendering, construction, and analysis[37]. In the area of optical imaging, GPU has been used in Monte Carlo simulation of photon migration[38], fluorescence molecular tomography[39], frequency-domain DOT[26], frequency-domain optical coherence tomography[40,41], and recently in photoacoustic tomography[42].

We developed a GPU-accelerated implementation of our CW DOT algorithm, achieving significant improvements when compared with previous CPU-based codes. When running on multi-GPUs, it allows for generating the first image within a latency of 3 s and then continuously generating images at the rate of 1.4 frame/s. Considering the window between the early hemodynamic changes and first EEG spikes, our implementation provides the ability for on-line monitoring of the pre-seizure state. Current limitations to implementing our method in human brain imaging, which typically uses mesh consisting of tens of thousands of nodes, involve the limited onboard memory and greatly increasing computation time. These limitations restrict the reconstruction of a large set of optical parameters and make the time unacceptable for online monitoring of brain activities. However, GPU programs written using CUDA can scale transparently from an entry level GPU to the state-of-the-art GPUs. A promising solution is to scale the single-GPU implementation to multiple GPUs. With the rapid development and maturity of toolsets supporting multi-GPU and clusters, such a computationally intractable application should be enabled soon[43].

This work represents a technological solution of real-time diffuse optical tomographic imaging of seizures in awake rats. These approaches can be applied to freely moving rats to facilitate the study of chronic epilepsy with spontaneous seizures and its interruption in our future research endeavors.

Acknowledgments

This research was supported in part by an NIH grant (R01 NS069848), the J. Crayton Pruitt Family and the B.J. and Eve Wilder endowment funds.

Abbreviations

DOT	diffuse optical tomography
GPU	graphic processing units

PTZ	pentylenetetrazol
IEEG	intracranial electroencephalogram

References

1. Paulsen KD, Jiang H. Spatially varying optical property reconstruction using a finite element diffusion equation approximation. Medical Physics-New York-Institute of Physics. 1995; 22:691–702.
2. Boas DA, et al. Imaging the body with diffuse optical tomography. Signal Processing Magazine, IEEE. 2001; 18:57–75.
3. Habermehl C, et al. Somatosensory activation of two fingers can be discriminated with ultrahigh-density diffuse optical tomography. Neuroimage. 2011
4. Dehghani H, Pogue BW, Poplack SP, Paulsen KD. Multiwavelength three-dimensional near-infrared tomography of the breast: initial simulation, phantom, and clinical results. Appl Opt. 2003; 42:135–145. [PubMed: 12518832]
5. Zhang Q, Jiang H. Three-dimensional diffuse optical imaging of hand joints: System description and phantom studies. Optics and Lasers in Engineering. 2005; 43:1237–1251.10.1016/j.optlaseng.2004.12.007
6. Zeff BW, White BR, Dehghani H, Schlaggar BL, Culver JP. Retinotopic mapping of adult human visual cortex with high-density diffuse optical tomography. Proc Natl Acad Sci U S A. 2007; 104:12169–12174.10.1073/pnas.0611266104 [PubMed: 17616584]
7. Habermehl C, Schmitz CH, Steinbrink J. Contrast enhanced high-resolution diffuse optical tomography of the human brain using ICG. Opt Express. 2011; 19:18636–18644. [PubMed: 21935232]
8. White BR, Liao SM, Ferradal SL, Inder TE, Culver JP. Bedside optical imaging of occipital resting-state functional connectivity in neonates. Neuroimage. 2012; 59:2529–2538.10.1016/j.neuroimage.2011.08.094 [PubMed: 21925609]
9. Fisher RS. Therapeutic devices for epilepsy. Ann Neurol. 2012; 71:157–168. [PubMed: 22367987]
10. Berényi A, Belluscio M, Mao D, Buzsáki G. Closed-loop control of epilepsy by transcranial electrical stimulation. Science. 2012; 337:735–737. [PubMed: 22879515]
11. Eggebrecht AT, et al. Mapping distributed brain function and networks with diffuse optical tomography. Nature Photonics. 2014
12. Obrig H. NIRS in Clinical Neurology–a 'Promising' Tool? Neuroimage. 2013
13. Grinvald A. Imaging input and output dynamics of neocortical networks in vivo: exciting times ahead. Proc Natl Acad Sci U S A. 2005; 102:14125–14126. [PubMed: 16189023]
14. Zaveri HP, Frei MG, Arthurs S, Osorio I. Seizure prediction: The Fourth International Workshop. Epilepsy Behav. 2010; 19:1–3. [PubMed: 20674508]
15. Zhang T, et al. Pre-seizure state identified by diffuse optical tomography. Scientific reports. 2014; 4
16. Hong SB, et al. Preictal versus ictal injection of radiotracer for SPECT study in partial epilepsy: SISCOM. Seizure. 2008; 17:383–386. [PubMed: 17768074]
17. Federico P, Abbott DF, Briellmann RS, Harvey AS, Jackson GD. Functional MRI of the pre-ictal state. Brain. 2005; 128:1811–1817. [PubMed: 15975948]
18. Sicard K, et al. Regional Cerebral Blood Flow and BOLD Responses in Conscious and Anesthetized Rats Under Basal and Hypercapnic Conditions: Implications for Functional MRI Studies. J Cereb Blood Flow Metab. 2003; 23:472–481. [PubMed: 12679724]
19. Jiang, H. Diffuse Optical Tomography: Principles and Applications. CRC Press; 2010.
20. André V, Pineau N, Motte JE, Marescaux C, Nehlig A. Mapping of neuronal networks underlying generalized seizures induced by increasing doses of pentylenetetrazol in the immature and adult rat: a c-Fos immunohistochemical study. Eur J Neurosci. 1998; 10:2094–2106. [PubMed: 9753096]

21. Topchiy IA, et al. Conditioned lick behavior and evoked responses using whisker twitches in head restrained rats. Behav Brain Res. 2009; 197:16–23. [PubMed: 18718491]
22. Yang J, Zhang T, Yang H, Jiang H. Fast multispectral diffuse optical tomography system for in vivo three-dimensional imaging of seizure dynamics. Appl Opt. 2012; 51:3461–3469. [PubMed: 22695584]
23. Dehghani H, et al. Near infrared optical tomography using NIRFAST: Algorithm for numerical model and image reconstruction. Communications in numerical methods in engineering. 2009; 25:711–732. [PubMed: 20182646]
24. Jiang H, Yuan Z, Gu X. Spatially varying optical and acoustic property reconstruction using finite-element-based photoacoustic tomography. J Opt Soc Am A Opt Image Sci Vis. 2006; 23:878–888. [PubMed: 16604770]
25. Bluestone AY, Stewart M, Lasker J, Abdoulaev GS, Hielscher AH. Three-dimensional optical tomographic brain imaging in small animals, part 1: hypercapnia. J Biomed Opt. 2004; 9:1046. [PubMed: 15447026]
26. Prakash J, Chandrasekharan V, Upendra V, Yalavarthy PK. Accelerating frequency-domain diffuse optical tomographic image reconstruction using graphics processing units. J Biomed Opt. 2010; 15:066009–066009. [PubMed: 21198183]
27. Bertram EH. Neuronal circuits in epilepsy: Do they matter? Exp Neurol. 2012
28. Gallagher A, et al. Non-invasive pre-surgical investigation of a 10 year-old epileptic boy using simultaneous EEG–NIRS. Seizure. 2008; 17:576–582. [PubMed: 18374608]
29. Cooper R, et al. Transient haemodynamic events in neurologically compromised infants: A simultaneous EEG and diffuse optical imaging study. Neuroimage. 2011; 55:1610–1616. [PubMed: 21255658]
30. Schridde U, et al. Negative BOLD with large increases in neuronal activity. Cereb Cortex. 2008; 18:1814–1827. [PubMed: 18063563]
31. Mormann F, Andrzejak RG, Elger CE, Lehnertz K. Seizure prediction: the long and winding road. Brain. 2007; 130:314–333.10.1093/brain/awl241 [PubMed: 17008335]
32. Schwartz TH, Hong SB, Bagshaw AP, Chauvel P, Bénar CG. Preictal changes in cerebral haemodynamics: review of findings and insights from intracerebral EEG. Epilepsy Res. 2011; 97:252–266. [PubMed: 21855297]
33. Rengier F, et al. 3D printing based on imaging data: review of medical applications. Int J Comput Assist Radiol Surg. 2010; 5:335–341. [PubMed: 20467825]
34. Sutin, J.; Wu, W.; Ruvinskaya, L.; Franceschini, MA. Biomedical Optics. Optical Society of America; Methods For Simultaneous Optical And Electrical Measurement Of Neurovascular Coupling in Awake Rats.
35. Holzer, M., et al. 4D functional imaging in the freely moving rat. Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE; IEEE; p. 29-32.
36. Habermehl C, et al. Somatosensory activation of two fingers can be discriminated with ultrahigh-density diffuse optical tomography. Neuroimage. 2012; 59:3201–3211. [PubMed: 22155031]
37. Eklund A, Dufort P, Forsberg D, LaConte SM. Medical Image Processing on the GPU-Past, Present and Future. Med Image Anal. 2013
38. Fang Q, Boas DA. Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units. Opt Express. 2009; 17:20178–20190. [PubMed: 19997242]
39. Freiberger M, Egger H, Liebmann M, Scharfetter H. High-performance image reconstruction in fluorescence tomography on desktop computers and graphics hardware. Biomedical optics express. 2011; 2:3207–3222. [PubMed: 22076279]
40. Zhang K, Kang JU. Real-time 4D signal processing and visualization using graphics processing unit on a regular nonlinear-k Fourier-domain OCT system. Opt Express. 2010; 18:11772–11784. [PubMed: 20589038]
41. Wang Y, et al. GPU accelerated real-time multi-functional spectral-domain optical coherence tomography system at 1300nm. Opt Express. 2012; 20:14797–14813. [PubMed: 22772175]

42. Yuan J, et al. Real-time photoacoustic and ultrasound dual-modality imaging system facilitated with graphics processing unit and code parallel optimization. J Biomed Opt. 2013; 18:086001–086001.

43. Schaetz, S.; Uecker, M. Algorithms and Architectures for Parallel Processing. Springer; 2012. p. 114-128.

Highlights

- Diffuse optical tomography can image hemodynamics in awake rats during evoked acute seizures;
- Massive parallelization with graphic processing units allows real-time DOT image reconstruction;
- Hemodynamic changes induced by generalized seizure onset precede EEG spikes;

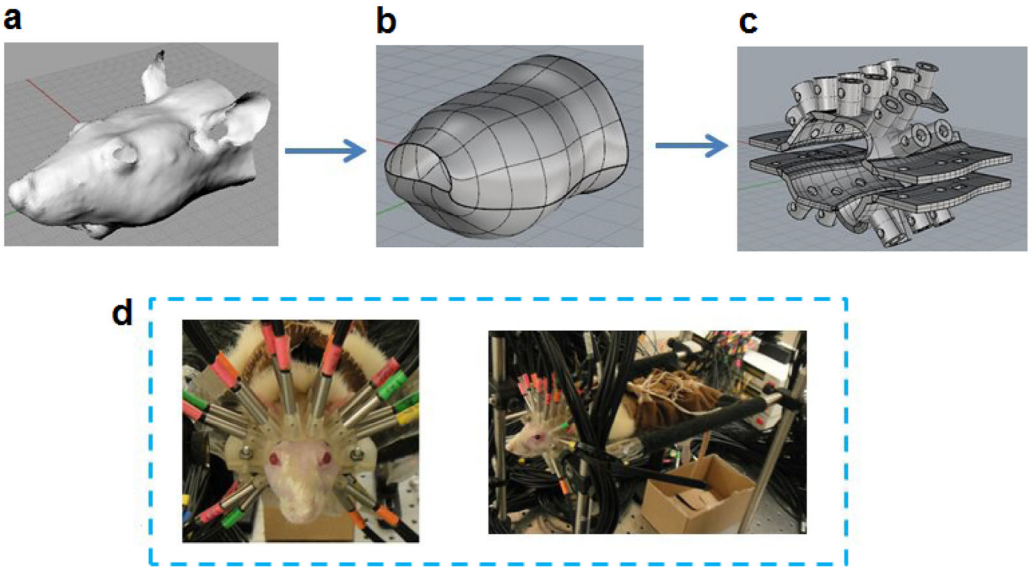

Fig. 1.

Illustration of 3D helmet model generation and restraining mechanisms. (a) head surface acquired with 3D scanner; (b) repaired surface model with Rhino 3D; (c) 3D helmet model; (d) rat wearing the helmet and jacket in the hammock during experiments.

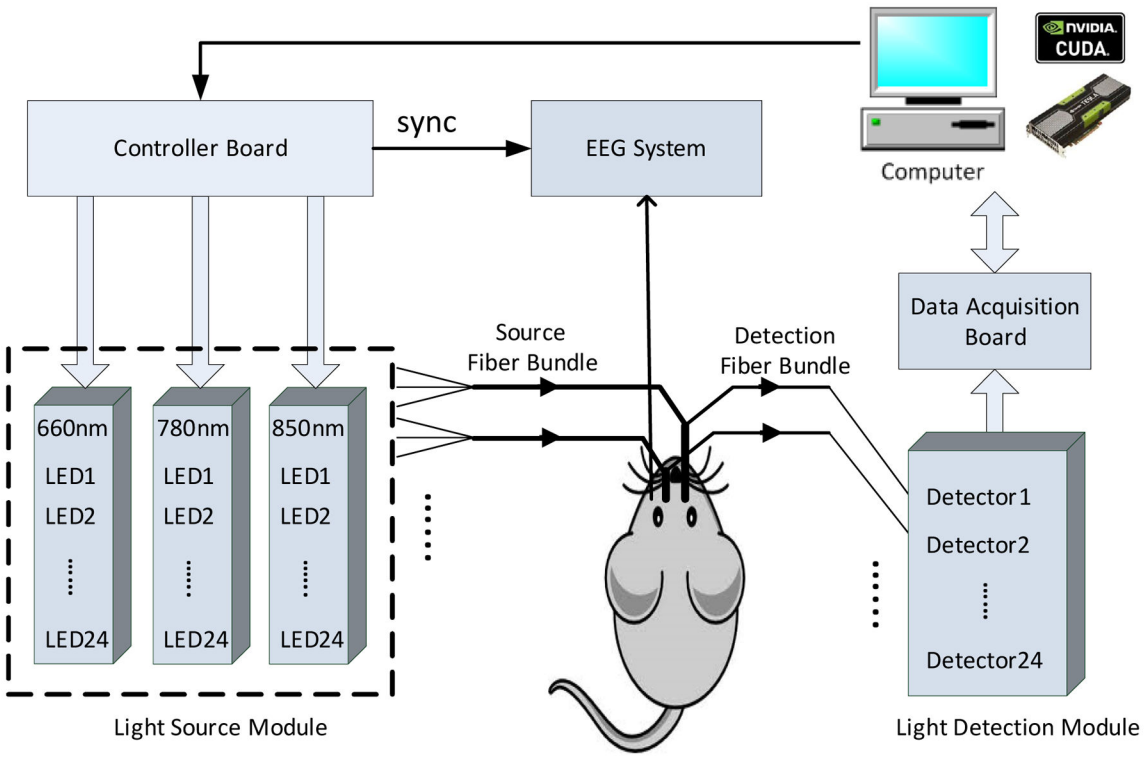

Fig. 2.
Schematic of the multispectral DOT system

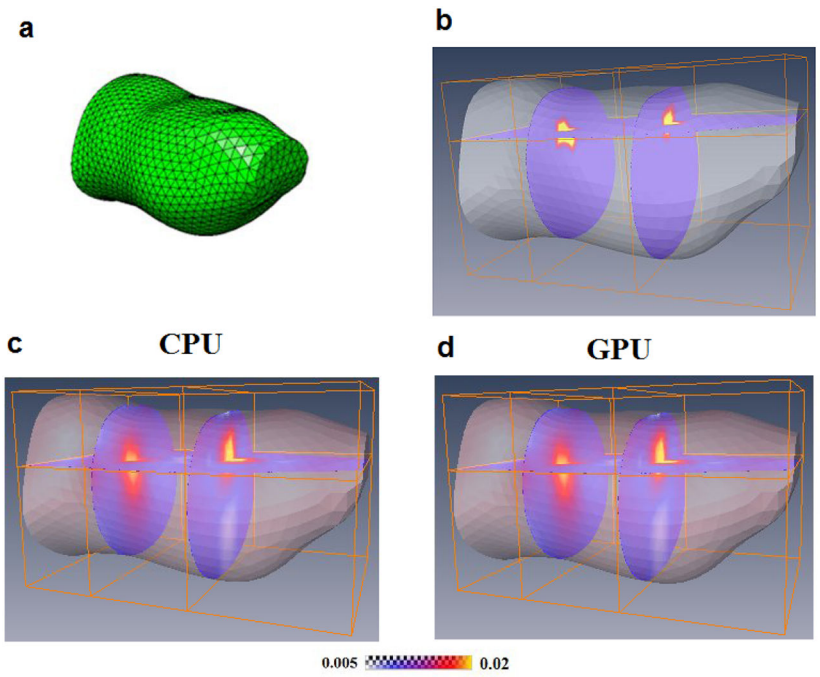

Fig. 3.

Comparison of simulation results from the CPU-based and GPU-based reconstructions. (a) 3D mesh of the rat head generated with Netgen; (b) Simulation setup with two targets; (c) CPU-based reconstructed μ_a image; (d) GPU-based reconstructed μ_a image.

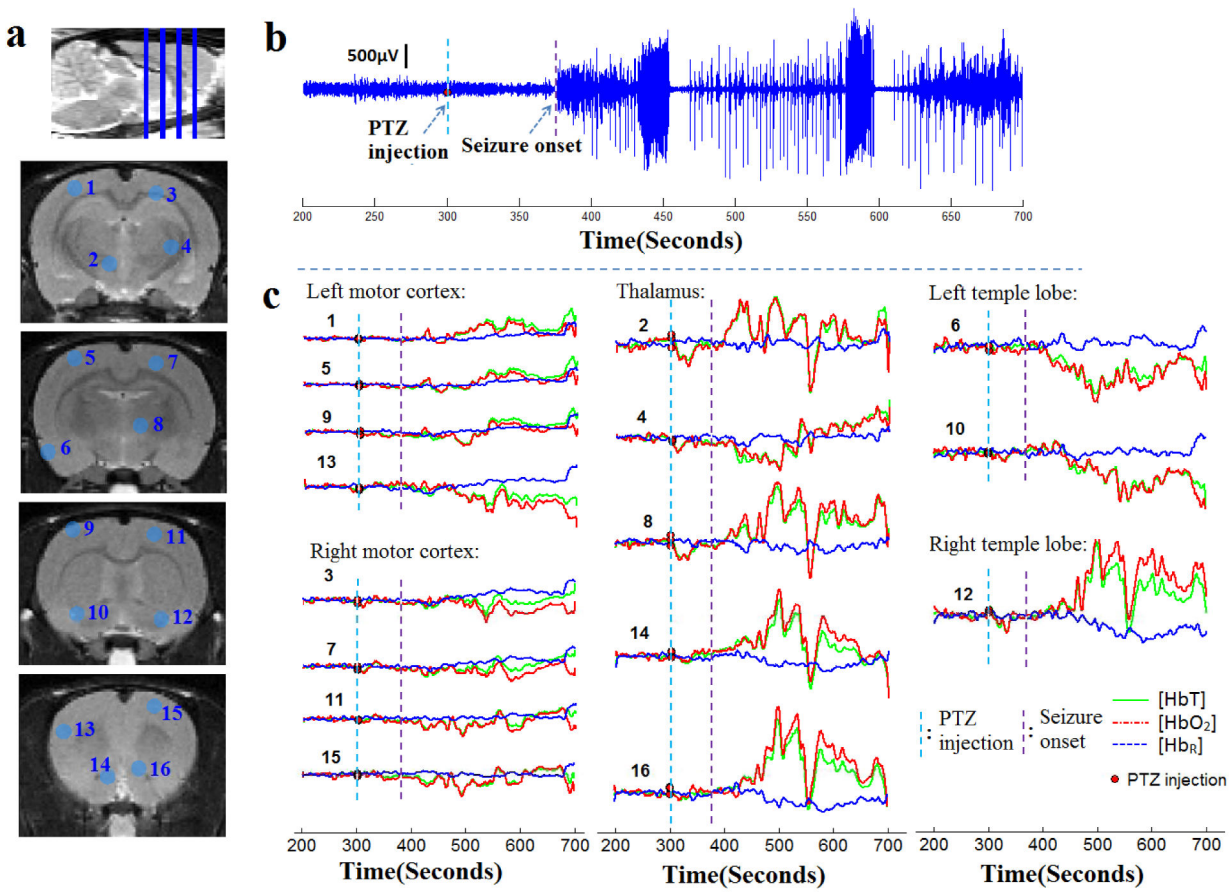

Fig. 4.

Simultaneous hemodynamic and neural changes during seizure onset. (a) Positions of four coronal slices and 16 regions of interest; (b) Seizure spikes were observed from IEEG recordings after PTZ injection; (c) Hemodynamic responses of [HbT], [HbO2], and [HbR] at 16 locations of interest from 4 coronal slices.

Table 1

Analysis of major modules in the reconstruction algorithm and their parallelization routines

Module	Calculation Time*	Parallelization routine
Sparse linear system in the forward model	14%	culaDcsrCgBlockjacobi (CULA sparse)
Sparse linear system in the adjoint method	14%	culaDcsrCgBlockjacobi (CULA sparse)
Assembly of Jacobian matrix	37%	Customized GPU-based code
Matrix multiplication	10%	culaDeviceSgemm (CULA dense)
Dense linear system solution	24%	culaDeviceSposv (CULA dense)
Miscellaneous	1%	Customized GPU-based code

*in percent of total calculation time per iteration running on CPU

Table 2

Performance comparisons between GPU-based and CPU-based reconstruction

Mesh	CPU		GPU		Speedup			Error(E)
	i7-2600	GT430	GTX580	GTX590	GT430	GTX580	GTX590	
mesh1: 2457 nodes	260s	8.1s	2.4s	2.8s	32x	108x	93x	<0.1%
mesh2: 3744 nodes	710s	15s	4.6s	5.0s	47x	154x	142x	<0.1%
mesh3: 5851 nodes	2250s	28s	7.1s	7.6s	80x	312x	296x	<0.1%