On Decreasing the Calculation Time in Multi-Dimensional Acoustic Numerical Simulation by Multi-GPU Parallel Computing

Naoki Kawada (1), Kan Okubo (1), Norio Tagawa (1) and Takao Tsuchiya (2)

(1) Faculty of System Design, Tokyo Metropolitan University, 6-6 Asahigaoka, Hino-shi, Tokyo 191-0065, Japan
(2) Faculty of Science and Engineering, Doshisha University, 1-3 Tatara-Miyakodani, Kyotanabe-shi, Kyoto 610-0321, Japan

PACS: 43.58.Ta, 43.55.Ka

ABSTRACT

Numerical analysis for sound wave propagation in time domain has been investigated widely as a result of computer development. Now, the development of accurate numerical schemes in time domain is an important technical issue. When we analyze large-scale sound wave propagation, the reduction of the calculation time is a necessary requirement. Recently, GPU (Graphic Processing Unit) is used as an acceleration tool for the calculation in various study fields. This movement is called GPGPU (General Purpose computing on GPUs). In the last few years the performance of GPU keeps on improving rapidly. This study makes an examination on decreasing the calculation time in acoustic field numerical analysis using GPU. We implement time-domain acoustic simulation (FDTD method, CIP method, and GCIP method) on GPU by CUDA (Compute Unified Device Architecture).

INTRODUCTION

To date, numerical analysis for sound wave propagation in time domain has been investigated widely as a result of computer development. Acoustic simulation in time domain is an effective technique for the estimation of time-series sound pressure data (e.g., nonlinear acoustic propagation phenomenon, acoustical measurements and instrumentation, architectural acoustics). Now, the development of accurate numerical schemes in time domain is an important technical issue.

The finite difference time domain (FDTD) method [1-8] is the most popular scheme used in time-domain acoustic simulation. However, we know that, using Yee's leapfrog algorithm, finite difference approximation certainly causes error owing to numerical dispersion. In the past study, the authors have proposed an acoustic simulation technique using generalized constrained interpolation profile method (GCIP method) [9], which is an expanded CIP method [10-16]. It is a method of characteristic [17] with very high accuracy; i.e., it enables the calculation with less-numerical dispersion. However, this method requires more calculation time than the conventional dispersive schemes.

When we analyze large-scale sound wave propagation, the reduction of the calculation time is a necessary requirement. Of course, there are usually trade-off relationships between required calculation time and numerical accuracy. Generally, calculation time and computational cost are proportional to the number of grid points. Additionally, more accurate schemes also require more computational cost.

Recently, GPU (Graphic Processing Unit) is used as an acceleration tool for the calculation in various study fields [18]. This movement is called GPGPU (General Purpose computing on GPUs) [19]. In the last few years the performance of GPU keeps on improving rapidly. That is, a PC (personal computer) with GPUs might be a personal supercomputer. GPU computing gives us the high-performance computing environment at a lower cost than before. Therefore, the use of GPUs contributes to a significant reduction of the calculation time in large-scale sound wave propagation.

This study makes an examination on decreasing the calculation time in acoustic field numerical analysis using GPU. We implement time-domain acoustic simulation (FDTD method, CIP method, and GCIP method) on GPU by CUDA (Compute Unified Device Architecture) [20-21]. We examine suitable algorithm and efficient thread models of CUDA for single- and multi- GPU computing.

ACOUSTIC NUMERICAL SIMULATION

Governing equations for linear acoustic fields

The governing equations for linear acoustic fields are given in Eq. (1) and Eq. (2).

\[ \rho \frac{\partial \vec{v}}{\partial t} = -\nabla p, \]  

(1)

\[ \nabla \cdot \vec{v} = -\frac{1}{K} \frac{\partial p}{\partial t}. \]  

(2)

In those equations, \( \rho \) denotes the density of the medium, \( K \) is the bulk modulus, \( p \) is sound pressure and \( \vec{v} \) is the particle velocity. Here we assume that the calculation is for a lossless and homogeneous medium.
Moreover, for simplicity, assuming $\vec{v} = (v_x, 0, 0)$ in order to analyze one-dimensional (1-D) acoustic field propagation in the $x$-direction, we can obtain the following equations from Eq. (1) and Eq. (2).

$$\frac{\partial v_x}{\partial t} + \frac{1}{\rho} \frac{\partial p}{\partial x} = 0,$$  \hspace{1cm} (3)

$$\frac{\partial p}{\partial t} + K \frac{\partial v_x}{\partial x} = 0.$$  \hspace{1cm} (4)

**FDTD method**

We can obtain Eqs. (5) and (6) from Eqs. (3) and (4) by employing second-order central difference approximation on a staggered grid.

$$v_x^{n+1/2}(i+1/2,j) = v_x^{n-1/2}(i+1/2,j)$$

$$\frac{\Delta t}{\rho} \left( p^n(i+1,j) - p^n(i,j) \right),$$

$$p^{n+1}(i,j) = p^n(i,j)$$

$$v_x^{n+1/2}(i+1/2,j) - v_x^{n-1/2}(i-1/2,j)$$

$$- K\Delta t \frac{\Delta x}{\Delta x},$$

where $\Delta t$ is the timestep, and $\Delta x$ is the grid size. Figure 1 depicts the grid model used in three-dimensional (3-D) FDTD analysis.

**GCIP($l,m$) method**

By addition and subtraction of Eq. (3) and Eq. (4), we obtain

$$\frac{\partial (p \pm Zv_x)}{\partial t} \pm c \frac{\partial (p \pm Zv_x)}{\partial x} = 0.$$  \hspace{1cm} (7)

In those equations, $Z$ indicates the characteristic impedance (i.e. $Z = \sqrt{K/\rho}$) and $c$ represents the sound velocity in medium (i.e. $c = \sqrt{K/\rho}$).

In addition, through simple spatial differentiation of the equations, the derivatives are given as

$$\frac{\partial}{\partial x} (p \pm Zv_x)$$

$$\frac{\partial}{\partial t} (p \pm Zv_x).$$

Therein, $\partial_x = \partial / \partial x$. Figure 2 depicts the grid model used in GCIP analysis, in which both acoustic field components and the derivatives of fields are located on the same grid (i.e., collocated grid).

We show the means to calculate the fields of the $(n+1)$ time step from the fields of $n$ time step, applying the CIP method to discretized acoustic field components. We define $F_{x^+}, F_{x^-}, G_{x^+},$ and $G_{x^-}$ as follows:

$$F_{x^+} = p + Zv_x, $$

$$F_{x^-} = p - Zv_x, $$

$$G_{x^+} = \partial_x p + Z\partial_x v_x,$$

$$G_{x^-} = \partial_x p - Z\partial_x v_x.$$  \hspace{1cm} (8)

Consequently, these components on grid points ($x = i\Delta x$) at the time step $n$ are given $F_{x^+}, F_{x^-}, G_{x^+},$ and $G_{x^-}$ as

$$F_{x^+}^n(i) = p^n(i) \pm Zv_x^n(i),$$

$$G_{x^+}^n(i) = \partial_x p^n(i) \pm Z\partial_x v_x^n(i).$$  \hspace{1cm} (9)

Next, applying the Hermite interpolation ($H$ and $H'$) to $F_{x^+}^n(i)$ and $G_{x^+}^n(i)$ yields the following equations related to propagation to the $\pm x$ -direction.

$$F_{x^+}^{n+1}(i) \leftarrow H^{(l,m)}(F_{x^+}^n, G_{x^+}^n),$$

$$G_{x^+}^{n+1}(i) \leftarrow H^{(l,m)}(F_{x^+}^n, G_{x^+}^n).$$  \hspace{1cm} (10)

Here, $H^{(l,m)}$ represents the Hermite interpolation with the $l$-th order using additional $m$ moments. For example, $H^{(3,1)}$ expresses the Hermite interpolation with the 1st order using physical values and their derivatives. That is, Hermite inter
polation in case of the GCIP(3,1) method, represented as \( H^{(3,1)} \), is given as

\[
F_{xx}^{n+1}(i) = \sum_{k=1}^{l} h_k^{(0)}(x) F_{xx}^{n}(k) + \sum_{k=1}^{l} h_k^{(1)}(x) G_{xx}^{n}(k),
\]

(17)

\[
G_{xx}^{n+1}(i) = \sum_{k=1}^{l} h_k^{(0)}(x) F_{xx}^{n}(k) + \sum_{k=1}^{l} h_k^{(1)}(x) G_{xx}^{n}(k),
\]

(18)

On the other hand, Hermite interpolation \( H^{(7,1)} \) used in the GCIP(7,1) method is given as

\[
F_{xx}^{n+1}(i) = \sum_{k=1}^{l} h_k^{(0)}(x) F_{xx}^{n}(k) + \sum_{k=1}^{l} h_k^{(1)}(x) G_{xx}^{n}(k),
\]

(19)

\[
G_{xx}^{n+1}(i) = \sum_{k=1}^{l} h_k^{(0)}(x) F_{xx}^{n}(k) + \sum_{k=1}^{l} h_k^{(1)}(x) G_{xx}^{n}(k),
\]

(20)

where \( h_k^{(m)}(x) \) is assumed Hermite interpolation function.

That is, the present method involves the use of the values of the acoustic field and their spatial derivatives at two or four grid points.

Moreover, using the following Eqs. (21) to (24), one can obtain acoustic field components \( p \) and \( v_x \) of time step \( (n+1) \).

\[
p^{n+1}(i) \leftarrow \frac{F_{xx}^{n+1}(i) + F_{xx}^{n+1}(i)}{2}
\]

(21)

\[
v_x^{n+1}(i) \leftarrow \frac{F_{xx}^{n+1}(i) - F_{xx}^{n+1}(i)}{2Z}
\]

(22)

\[
\partial_x p^{n+1}(i) \leftarrow \frac{G_{xx}^{n+1}(i) + G_{xx}^{n+1}(i)}{2}
\]

(23)

\[
\partial_x v_x^{n+1}(i) \leftarrow \frac{G_{xx}^{n+1}(i) - G_{xx}^{n+1}(i)}{2Z}
\]

(24)

**GPU PROGRAMMING BY CUDA**

Recently, GPU has evolved into a highly parallel, multi-threaded, many core processor with great computational ability and very high memory bandwidth owing to the huge computing demand for real-time and high-definition 3D graphics [4]. Figure 3 shows the performance comparison between GPUs and typical CPUs.

In November 2006, NVIDIA introduced CUDA, a general purpose parallel computing architecture. CUDA comes with a software environment that allows developers to use C as a high-level programming language. GPU should be used as an acceleration tool for the calculation in computational acoustics.

Figure 4 depicts CUDA’s grid model and block model. The blocks of the kernel grid, which is invoked by a CUDA program on the host CPU, are enumerated and distributed to multiprocessors with available execution capacity. The threads of a thread block execute concurrently on one Streaming Multiprocessor (SMP). As thread blocks terminate, new blocks are launched on the vacated SMP.

A SMP consists of eight Streaming Processor (SP) cores, a multithreaded instruction unit, on-chip constant cache, texture cache and shared memory. GPU have several SMPs; for example, GeForce GTX 285 owns 30 SMPs, GeForce GTX 295 owns \( 2 \times 30 \) SMPs. That is, GeForce GTX 285 has 240 parallel cores. Figure 5 shows the Hardware Model of NVIDIA GPU (Compute Capability 1.0 – 1.3).

We can use multiple GPUs programming as CUDA devices, if the computer system loads same type GPUs. Multiple GPUs computation needs to transfer boundary data between neighboring nodes across the PCI Express bus. Therefore, in multiple GPUs programming, we must consider the overhead introduced by the transference.

![Figure 3. Performance comparisons between GPUs and typical CPUs](Source: ([21], 2010))

![Figure 4. Thread Hierarchy (Grid of Thread Blocks)](Source: ([21], 2010))
RESULTS

NVIDIA’s CUDA programming environment is used with NVIDIA GeForce GTX 285 and GTX 295 GPUs. For comparison, we employ the PC with Intel Core i7 920 2.67GHz. This processor has eight hyperthreaded cores, or effectively scales 8 threads. We use FLOPS (Floating point operations per Second) and FUPS (Fields Update Per Second) for performance evaluation of GPU calculation. 1 FUPS means the capability of updating acoustic fields point once per second. In this study, we use single precision floating-point numbers.

We adopt a simple uniform media (i.e., air) as a calculation model. Uniform square grids are employed for CPU and GPU calculation. We calculate the pressure distribution at each time. Moreover, in FDTD analysis we don’t consider absorbing boundary condition, whereas GCIP analysis automatically sets this condition.

2-D Acoustic Field Calculation (using single GPU)

We estimated the calculation time required for a 2-D simple acoustic model except for the absorbing boundaries. First, we show calculation time using single GeForce GTX 285 GPU. Table 1, 2, and 3 show results of comparison of the calculation time between the GPU results and CPU results, where the analysis region is 1024 × 1024 cells and whole calculation time is divided into 1024 time steps. Here, all measurements are made on the above 285 GPU and Intel Core i7 processor machine with 12 GB of memory.

These tables also provide results of FLOPS and FUPS. FDTD results using GPU is ca. 9.5 times faster than 8-threads CPU calculation. The GCIP(3,1) method using GPU is ca. 28 times faster than 8-threads CPU calculation. Moreover, the GCIP(7,1) method using GPU is ca. 47.5 times faster than CPU results. Relative comparison of the run-time between the CIP result and FDTD result reveal the following: CIP analysis requires more run-time than FDTD analysis when equivalent discretization is used. However, by using GPU calculation, difference of required calculation time between FDTD result and CIP result became small.

2-D Acoustic Field Calculation (using multi-GPUs)

Next, we show results of acoustic field calculation using multiple GPUs. Table 4 shows the comparison results of FDTD calculation between the GPU results and CPU result, where the analysis region is 8192 × 8192 cells and whole calculation time is divided into 1024 time steps. All measurements are made on the above 295 GPU and Intel Core i7 processor machine. This table gives the calculation time in case of 1-GPU, 2-GPUs, 4-GPUs, and 8-GPUs. This result illustrated FDTD calculation on 8 GPUS is ca. 48 times faster than 8-thread CPU calculation.

Table 5 and 6 show the comparison results of GCIP calculation between the GPU results and CPU result, where the analysis region is 3072 × 3072 cells and whole calculation time is divided into 1024 time steps. All measurements are made on the above 295 GPU and Intel Core i7 processor machine. These tables also give the calculation time in case of 3-GPU, 2-GPUs, 4-GPUs, and 8-GPUs. This result reveals the followings: GCIP(3,1) calculation on 8 GPUs is ca. 117.5 times faster than 8-thread CPU calculation. On the other hand, GCIP(7,1) calculation on 8 GPUs is ca. 114.7 times faster.

Next, Fig. 6, 7, and 8 respectively show the calculation time by FDTD method and GCIP methods against the number of grids in the case of 1-GPU, 2-GPUs, 4-GPUs, and 8-GPUs computing. This indicates that the overhead introduced by the data transference between multi-GPUs caused the increase of calculation time when the number of grids is small. On the other hand, these overheads become small when the number of grids is large.
other hand, the calculation time almost linearly increases with the increase of the number of grids if it becomes larger. That is because operation time is dominant as compared to data transference time with the increase of simulation area.

We show results of calculation time in case of the large-scale analysis using multi-GPUs. Table 7 shows FDTD results, where the analysis region is $23552 \times 23552$ cells and whole calculation time is divided into 1024 time steps. FDTD calculation on 8 GPUs is 52.8 times faster than multi-thread calculation on CPU.

Table 8 and 9 show GCIP results, where the analysis region is $10240 \times 10240$ cells. GCIP(3,1) calculation on 8 GPUs is 153 times faster than 8-thread calculation on CPUs. On the other hand, GCIP(7,1) calculation on 8 GPUs is 179.6 times faster than 8-thread calculation on CPU. As a comparison of table 5 and 6 with table 8 and 9, we can see the multi-GPUs calculation by GCIP methods is effective for the analysis of the large-scale region.

### Table 7. Calculation Time of FDTD method

<table>
<thead>
<tr>
<th>Device</th>
<th>Calc Time [s]</th>
<th>GFLOPS</th>
<th>GFUPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core i7(8 threads)</td>
<td>1463.16</td>
<td>4.66</td>
<td>0.39</td>
</tr>
<tr>
<td>GTX 295 * 8</td>
<td>27.72</td>
<td>245.89</td>
<td>20.49</td>
</tr>
</tbody>
</table>

### Table 8. Calculation Time of GCIP(3,1) method

<table>
<thead>
<tr>
<th>Device</th>
<th>Calc Time [s]</th>
<th>GFLOPS</th>
<th>GFUPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core i7(8 threads)</td>
<td>4532.48</td>
<td>3.41</td>
<td>0.024</td>
</tr>
<tr>
<td>GTX 295 * 8</td>
<td>29.62</td>
<td>522.01</td>
<td>3.63</td>
</tr>
</tbody>
</table>

### Table 9. Calculation Time of GCIP(7,1) method

<table>
<thead>
<tr>
<th>Device</th>
<th>Calc Time [s]</th>
<th>GFLOPS</th>
<th>GFUPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core i7(8 threads)</td>
<td>6185.25</td>
<td>4.72</td>
<td>0.017</td>
</tr>
<tr>
<td>GTX 295 * 8</td>
<td>34.44</td>
<td>848.02</td>
<td>3.12</td>
</tr>
</tbody>
</table>

### 3-D Acoustic Field Calculation (using single GPU)

Table 10 shows results of comparison of the calculation time for 3-D acoustic simulation between the GPU and CPU results, where the analysis region is $256 \times 256 \times 256$ cells and whole calculation time is divided into 1024 time steps. Here, all measurements are made on the above 285 GPU and Intel Core i7 processor machine. It is clarified that 3-D FDTD acoustic calculation using GPU is ca.16.5 times faster than 8-threads CPU calculation. Therefore, in single GPU FDTD calculation, 3-D acoustic analysis is more accelerated than 2-D analysis.

Table 11 and 12 show the comparison results of GCIP calculation between the GPU results and CPU result, where the analysis region is $128 \times 128 \times 128$ cells and whole calculation time is divided into 1024 time steps. These tables indicate the followings; the 3-D GCIP(3,1) acoustic calculation using GPU is ca. 37 times faster than 8-threads CPU calculation. Moreover, the GCIP(7,1) method using GPU is ca. 53 times faster than CPU results. In single GPU GCIP calculation, 3-D acoustic analysis is more accelerated than 2-D analysis, similarly.

### Table 10. Calculation Time of FDTD method

<table>
<thead>
<tr>
<th>Device</th>
<th>Calc Time [s]</th>
<th>GFLOPS</th>
<th>GFUPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core i7(8 threads)</td>
<td>115.20</td>
<td>2.68</td>
<td>0.15</td>
</tr>
<tr>
<td>GTX 285</td>
<td>6.98</td>
<td>44.30</td>
<td>2.46</td>
</tr>
</tbody>
</table>

### 3-D Acoustic Field Calculation (using multi-GPUs)

Next, we show results of 3-D acoustic field calculation using multiple GPUs. Figure 9 depicts the present division model for parallel computing. We divide the analysis domain into multiple sub-domains along z-direction. The number of sub-domains is corresponding to that of multi-GPUs. Table 13 shows the comparison results of FDTD calculation between the GPU result and CPU result, where the analysis region is $512 \times 512 \times 512$ cells. All measurements are made on the above 295 GPU and Intel Core i7 processor machine.
This table gives the calculation time in case of 8-GPUs. This result illustrated FDTD calculation on 8 GPUs is ca. 59 times faster than 8-thread CPU calculation.

Figure 10 shows the calculation time by FDTD method against the number of grids in the case of 8-GPUs computing. In this figure, the results of following four cases are shown: (1) case X; analysis domain expands along x-direction, (2) case Y; analysis domain expands along y-direction, (3) case Z; analysis domain expands along x, z-direction, (4) case XYZ; analysis domain equally expands along x, y, and z-direction. These results illustrates that analysis domain should expand along the division direction.

CONCLUSIONS

This study made an examination on decreasing the calculation time in numerical analysis for sound wave propagation in time domain using single GPU and multi-GPUs computing. As a result, it is clarified that multi-GPU FDTD calculation is faster than multi-CPU FDTD calculation by 48 times at the maximum in 2-D acoustic analysis. On the other hand, multi-GPU GCIP calculation needs less calculation time than multi-CPU calculation by about 1/100 times. Therefore, the multi-GPUs calculation is very effective for GCIP methods as an acceleration tool.

Moreover, for large-scale acoustic simulation, these results show the feasibility of high-speed and high-precise simulation analysis by hardware acceleration using multi-GPU calculation. In near future we intend to examine more effective thread model for 3-D GPU calculation.

REFERENCES