

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年 6月7日現在

機関番号：42718

研究種目：若手研究（B）

研究期間：2010～2011

課題番号：22700060

研究課題名（和文） マルチ GPU クラスタによるリアルタイム処理及び倍精度数値計算高速化に関する研究

研究課題名（英文） Studies of Real-time processing and high-speed double-precision computation using Multi-GPU cluster system

研究代表者

高田直樹（TAKADA NAOKI）

湘北短期大学・情報メディア学科・准教授

研究者番号：50290713

研究成果の概要（和文）：コンピュータグラフィックス処理を行う GPU ボードを3枚搭載した PC を4台用い、GPU クラスタシステムを構築した。本システムに究極の立体映像技術として知られている計算機合成ホログラムの計算を実装し、リアルタイム処理を実現した。また、電磁波のシミュレーションを実装した。なお、電磁波のシミュレーションでは高い精度を必要とするため、倍精度を用いた。本研究により、国際会議を依頼されており、高い評価を得た。

研究成果の概要（英文）：Multi-GPU cluster system composed of four PCs, each of which is equipped with three GPU boards, has been built. The computations of computer-generated hologram (CGH) and electromagnetic simulation were implemented in the multi-GPU cluster system. CGH is well-known as the technique of ultimate three-dimensional television. Electromagnetic simulation requires high-precision. The electromagnetic simulation using multi-GPU cluster system was performed by double-precision calculation. The result of this study was reported in the invited talk of international conference.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	3,600,000	1,080,000	4,680,000
2011年度	500,000	150,000	650,000
年度			
年度			
年度			
総計	4,100,000	1,230,000	5,330,000

研究分野：総合領域

科研費の分科・細目：情報学・計算機システム・ネットワーク

キーワード：Graphics Processing Unit, GPU, GPGPU, GPU コンピューティング, マルチ GPU クラスタ, 計算機合成ホログラム, CGH, 電磁界シミュレーション, FDTD 法

## 1. 研究開始当初の背景

GPU (Graphics Processing Unit) は、コンピュータグラフィックス(CG) 処理を行うための半導体チップである。PC内のグラフィックスボード上に搭載されており、一般的に利用されている。近年、ゲーム機器の急速な発展に伴い、3DCG表示に必要な計算処理を行うGPUの計算性能及びコストパフォーマンス

は著しく向上している。ゲームにおいて、より自然に近い感覚が求められるようになり物理シミュレーションが取り入れられるようになった。その演算を高速にするためGPUに単精度浮動小数点演算ユニットが多数搭載されるようになった。これに伴いGPUを用いて汎用的な数値計算を高速化する研究が盛んに行われるようになった。当初、GPU

を用いた汎用数値計算の高速化は、コンピュータグラフィックス(CG)処理を用いて行われていたことから、GPGPU (General purpose computation on GPU) と呼ばれていた。2007年にNVIDIA社からCUDAプログラム開発環境が発表された。CG処理に依存せずC言語に似た言語を用いてGPU用プログラムを開発することが可能となった。CG処理の専門知識を必要とせず、プログラムを開発できることから、GPUを用いた汎用的な数値計算高速化に関する研究は、GPUコンピューティングと呼ばれるようになった。

GPUは単精度浮動小数点演算性能のみならず、コストパフォーマンスにも優れていることから、ハイパフォーマンスコンピューティング(HPC)の分野においても重要な研究となっており、世界的に研究が盛んに行われている。国内において、東工大のスーパーコンピューティングシステムであるTSUBAME1.2にGPUが搭載され、汎用的な数値計算の高速化に関する研究が行われた。

研究開始当初、GPUで倍精度浮動小数点演算がサポートされるようになった。また、当時の次世代GPUとして、単精度のみならず倍精度浮動小数点演算性能も向上させた「Fermi」アーキテクチャを搭載したGPUが発売される時期と重なっていた。「Fermi」アーキテクチャを搭載したGPUのひとつであるNVIDIA GeForce GTX 480では、1.4GHzで動作する480個のコアからなり、メモリバンド幅も177.4GB/secと高性能であった。また、デバイスメモリから読み出す際、L1キャッシュを使用できるようになり、メモリアクセスが多い数値計算において効果が期待できるアーキテクチャとなっている。

本研究を計画している時期に、Fermiアーキテクチャを搭載した当時の次世代GPUが話題となっていた。研究開始時期には、FermiアーキテクチャのGPUが発売されることが予想できたため、本研究で使用することを想定した。また、1台のPCに複数のGPUを搭載したマルチGPU環境のPCによる計算高速化の研究が始まりつつあり、実際に検証し実用的であることを確認していた。これらのことから、Fermiアーキテクチャを搭載した当時の次世代GPUを用いてマルチGPUクラスタを構築し、本研究を行うこととした。

## 2. 研究の目的

本研究では、当時、未開拓の研究であったマルチGPUクラスタシステムを用いたリアルタイム処理、及び、倍精度数値計算の高速化を目的とした。リアルタイム処理の課題として、計算機合成ホログラム(CGH)を扱った。倍精度数値計算として、FDTD法による電磁界シミュレーションを扱った。

最近、テレビ、映画、ゲーム機など、三次元(3D)立体映像がブームとなっている。しかし、一般的に、右目用と左目用の二次元映像を用いた3D表示が中心となっている。一方、ホログラフィは、三次元物体による物体光を干渉縞として媒体に記録、また、記録された媒体に再生光を照射することで物体光を生成し忠実に三次元物体を再現できる唯一知られた技術である。眼鏡を必要とせず、視覚疲労がないため究極の立体映像技術として期待されている。三次元物体情報を記録した干渉縞をホログラムと呼び、コンピュータによって作られたホログラムを計算機合成ホログラム(CGH: Computer Generated Hologram)という。CGHは、仮想的な三次元物体によるホログラムを作成することができる。CGHを液晶ディスプレイ(LCD: Liquid Crystal Display)などの電子デバイスに表示し、再生光を照射することにより三次元物体を再生することができる(電子ホログラフィ)。しかし、CGHは計算量が膨大であり、未だ実用化されていない。GPUは本来コンピュータグラフィックス用のプロセッサである。GPUで計算した結果をCPU(Central Processing Unit)を介さず、直接、LCDに表示することができる。CGH計算は並列計算に向いており、さらに計算結果であるCGHを画像としてLCDに表示する。このことから、電子ホログラフィの研究にGPUが向いていることがわかる。CGH計算において使用するデータ量は、演算量に比べ少ない。つまり、マルチGPUクラスタシステムにおいて、PC内に搭載されたGPUボード間、そして、PC間の通信がボトルネックにはならない。そのため、マルチGPUクラスタにおいて理想的な計算高速化を実現できる可能性が高い。

FDTD (Finite-Difference Time-Domain) 法は、マクスウェル方程式を、直接、時間と空間に対して離散化し、中心差分により得られた式を用いて現在から未来へ計算する手法である。境界条件の取り扱いが容易であることから、アンテナ解析など電磁界問題の解析に幅広く使用されている。実用的に使用されている電磁界シミュレーション手法の一つである。しかし、実際の空間を、電界及び磁界を、直接、離散化して計算するため、計算領域が大規模となる場合、または、解析モデルが複雑である場合には、使用する計算機メモリ量及び計算時間は膨大となる。FDTD法は、計算量よりも計算に使用するデータ量が多く、つまり、メモリアクセスが多い数値計算である。FDTD法の計算高速化を実現させるためには、メモリバンド幅が広くならない。GPUは、演算性能が優れているだけではない。GPUボードにはデバイスメモリが搭載されており、GPU-デバイスメモリ間のメモリバンド幅は、CPU-メインメモリ間のメモリバンド幅に比べ広がっている。よって、GPUによる計算

高速化が期待できる。マルチ GPU クラスタでは、GPU ボード間、PC 間の通信量がボトルネックとなる。また、FDTD 法では、電磁波の計算を行うため倍精度を必要とする。倍精度数値計算問題として FDTD 法を扱った。

### 3. 研究の方法

本研究で使用するマルチ GPU クラスタは、図 1 に示すように 4 つのノードから構成される。1 ノードに 3 枚の GPU ボードを搭載する。

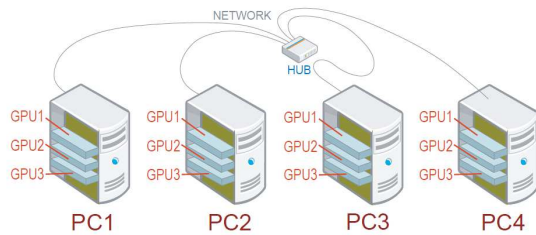


図 1 本研究で使用したマルチ GPU クラスタシステム

GPU には、当時の次世代 GPU である NVIDIA Geforce 480 GTX を用いた。CGH 計算で使用するメモリ量は計算量に比べ少ない。ノード間通信も物体点の座標データのみであるため、高速なネットワークを必要としない。よって、1Gbps のギガビット・イーサネットをネットワークに使用した。各ノードから、物体点の座標データが取得できるように Network File System (NFS) を使用した。

図 1 の各 GPU に電子ホログラフィ用高精細 LCD パネルを接続したマルチ GPU クラスタ電子ホログラフィシステム (図 2) を想定している。

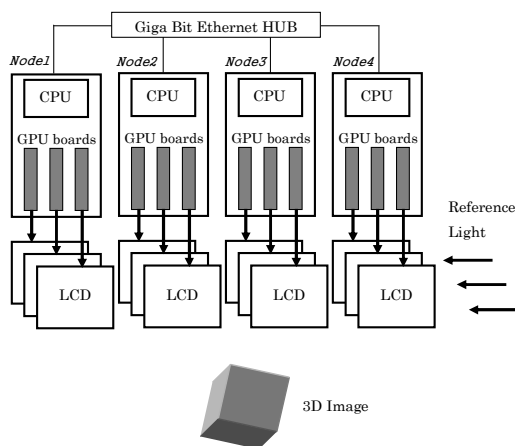


図 2 マルチ GPU クラスタ

FDTD 法においては、マルチ GPU クラスタにおいて、通信量がボトルネックとなるため、ノード間通信に Infiniband を用いた。

### 4. 研究成果

FDTD 法の計算において、最初に 1 枚の GPU (NVIDIA Geforce GTX 480) のスレッドブロック数と計算速度について調査し、最適なスレッドブロック数を割り出した。倍精度浮動小数点の場合の計算時間は、単精度浮動小数点の場合の計算時間の約 2 倍となることが確認され、最適なスレッドブロック数を選べば妥当な計算速度が得られることが確認された。また、Fermi アーキテクチャの GPU では、デバイスメモリから読み出す際、L1 キャッシュを使用できるようになった。GPU のプログラムであるカーネルにおいて、共有メモリを用いてコーディングする必要がなくなった。実際に共有メモリを用いた場合と、共有メモリを用いずにコーディングした場合は、ほとんど計算速度の違いが確認されなかった。本研究で構築したマルチ GPU クラスタに FDTD 法を実装し、計算高速化について検討した。

次に、CGH 計算についてであるが、各 GPU で解像度  $1,600 \times 1,024$  の CGH を作成した。本研究で使用したマルチ GPU クラスタには、全部で 12 枚の GPU が搭載されており、全部で解像度  $6,400 \times 3,072$  の CGH を作成した。計算で作成した CGH を図 3 に示す。次に、比較のため解像度  $1,980 \times 1,024$  の CGH を図 4 に示す。図 4 の CGH は図 3 に比べ、縞模様が粗くなっているのが確認できる。

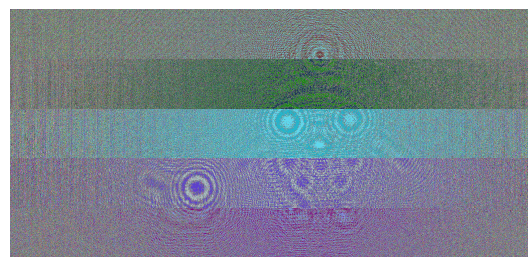


図 3 高解像度 ( $6,400 \times 3,072$ ) の CGH

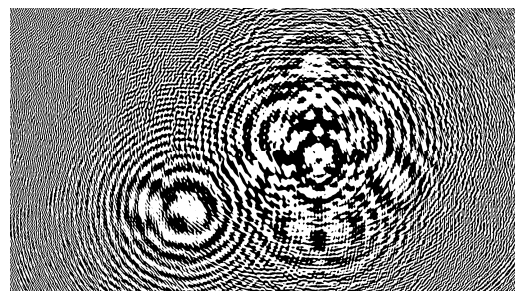


図 4 解像度 ( $1,980 \times 1,024$ ) の CGH

図3, 4のCGHによる3次元物体の再生像を数値シミュレーションにより, 得られた結果を図5, 6に示す. 明らかに高解像度(6,400×3,072)のCGHによる再生像のほうが鮮明であることが確認された



図5 図3のCGHによる再生像



図6 図4のCGHによる再生像

表1 CPUとマルチGPUクラスタシステムによるCGH計算時間の比較

物体 点数	CGH 計算時間 [ms]		高速化率
	A CPU (8 threads)	12 GPU	
512	19,764	21.64	913
1,024	39,534	32.65	1,211
2,048	78,982	54.59	1,447
3,072	118,455	76.61	1,546
4,096	157,935	98.56	1,602

CPU(Intel Core i7 930, 8スレッド使用)での計算時間と本研究で使用したマルチGPUクラスタによる計算時間を計測し, CPUに対する計算高速化率を求めた. その結果を表1に示す. 2,048点の三次元物体による約20MpixelのCGHを約20fpsで表示することができることが示された. さらに, 4,096点で構成された三次元物体による約20MpixelのCGH計算において, CPU(Intel Core i7 930, 8スレッド使用)に対し, 1,600倍の計算高速化を実現した.

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計4件)

- ① Tomoyoshi Shimobaba, Jiantong Weng, Takahiro Sakurai, Naohisa Okada, Takashi Nishitsuji, Naoki Takada, Atsushi Shiraki, Nobuyuki Masuda and Tomoyoshi Ito, "Computational wave optics library for C++: CWO++ library," Computer Physics Communications, 査読有, Vol. 183, Issue 5, May 2012, pp.1124-1138.  
DOI: 10.1016/j.cpc.2011.12.027
- ② 高田直樹, 下馬場朋禄, 中山弘敬, 老川稔, 増田信之, 伊藤智義, "マルチLCD用マルチGPUクラスタシステムによる計算機合成ホログラムの計算高速化," 第10回情報科学技術フォーラム(FIT2011)講演論文集 第1分冊, 査読有, 2011年9月, pp.117-122.
- ③ Naoki Takada, Tomoyoshi Shimobaba, Nobuyuki Masuda and Tomoyoshi Ito, "Improved Performance of FDTD Computation Using a Thread Block Constructed as a Two-Dimensional Array with CUDA," Applied Computational Electromagnetics Society Journal, 査読有, Vol. 25, No. 12, December, 2010, pp.1061-1069.  
<http://aces.ee.olemiss.edu/search.php?vol=25&no=12&type=2>
- ④ Hirotaka Nakayama, Naoki Takada, Yasuyuki Ichihashi, Shin Awazu, Tomoyoshi Shimobaba, Nobuyuki Masuda and Tomoyoshi Ito, "Real-time color electroholography using multi graphics processing units and multihigh-definition liquid-crystal display panels," Applied Optics, 査読有, Vol. 49, Issue 31, October 2010, pp. 5993-5996.  
DOI:10.1364/AO.49.005993

[学会発表] (計12件)

- ① Takashi Nishitsuji, Tomoyoshi Shimobaba, Takahiro Sakurai, Naoki Takada, Nobuyuki Masuda and Tomoyoshi Ito, "Fast calculation of Fresnel diffraction calculation using AMD GPU and OpenCL," Korea-Japan Workshop for Information Photonics (DHIP2011), 2011年11月10日, ソウル国立大学(韓国).
- ② Naoki Takada, Tomoyoshi Shimobaba, Hirotaka Nakayama, Minoru Oikawa, Yasuyuki Ichihashi, Nobuyuki Masuda and Tomoyoshi Ito, "Fast Computation of Computer-Generated-Hologram Using Multi-GPU Cluster System," International Conference on 3D systems and Applications

- (3DSA 2011), 2011年6月22日, 韓国科学技術会館(韓国).
- ③ Takashi Nishitsuji, Tomoyoshi Shimobaba, Takahiro Sakurai, Naoki Takada, Nobuyuki Masuda and Tomoyoshi Ito, "Fast calculation of Fresnel diffraction calculation using AMD GPU and OpenCL," Digital Holography and Three-Dimensional Imaging(DH) (DH2011), 2011年5月11日, 東京大学(東京都).
  - ④ 西辻崇, 下馬場朋禄, 増田信之, 高田直樹, 伊藤智義, "AMD GPUとOpenCLを用いたフレネル回折計算の高速計算," 2011年電子情報通信学会総合大会, 2011年3月16日, 東京都市大学(東京都).
  - ⑤ Naoki Takada, Hirotaka Nakayama, Shin Awazu, Yasuyuki Ichihashi, Tomoyoshi Shimobaba, Nobuyuki Masuda and Tomoyoshi Ito, "A real-time color electroholographic reconstruction system using multi graphics processing unit," The 17<sup>th</sup> International Display Workshops, 2010年12月2日, 福岡国際会議場(福岡県).
  - ⑥ Tomoyoshi Shimobaba, Hirotaka Nakayama, Nobuyuki Masuda, Naoki Takada and Tomoyoshi Ito, "24 mega pixels phase electroholography," The 17<sup>th</sup> International Display Workshops, 2010年12月2日, 福岡国際会議場(福岡県).
  - ⑦ 高田直樹, 下馬場朋禄, 増田信之, 伊藤智義, "シングルCUDA 互換GPUを用いたFDTD法の計算高速化," 情報科学技術フォーラム(FIT2010), 2010年9月7日, 九州大学(福岡県).
  - ⑧ 下馬場朋禄, 市橋保之, 増田信之, 高田直樹, 伊藤智義, "CUDAとOpenCLを用いたNVIDIA及びAMD GPU上でのホログラム生成の性能比較," 3次元画像コンファレンス2010, 2010年7月10日, 東京大学(東京都).
  - ⑨ 下馬場朋禄, 三瓶卓方, 市橋保之, 増田信之, 高田直樹, 伊藤智義, "位相変調型LCDを用いた1800万画素電子ホログラフィシステム," 3次元画像コンファレンス2010, 2010年7月10日, 東京大学(東京都).
  - ⑩ Tomoyoshi Shimobaba, Tomoyoshi Ito, Nobuyuki Masuda, Yasuyuki Ichihashi and Naoki Takada, "Fast calculation of computer-generated- hologram using AMD GPU and OpenCL," International Conference on 3D Systems and Applications 2010, 2010年5月21日, 学術総合

センター(東京都).

- ⑪ [招待講演] Naoki Takada, Tomoyoshi Shimobaba, Nobuyuki Masuda and Tomoyoshi Ito, "The Performance Evaluation of CUDA Implementation for Fast FDTD Computation Using a Single GPU," The 26th International Review of Progress in Applied Computational Electromagnetics, 2010年4月27日, Tampere Hall (Finland).
- ⑫ Tomoyoshi Shimobaba, Naoki Takada, Yasuyuki Ichihashi, Nobuyuki Masuda and Tomoyoshi Ito, "Fast numerical wave-optics library using graphic processing unit: GWO library, and its applications to holography," Digital Holography and Three-Dimensional Imaging (DH2010), 2010年4月12日, The Deauville Beach Resort Hotel (U.S.A).

## 6. 研究組織

### (1) 研究代表者

高田 直樹 (TAKADA NAOKI)

湘北短期大学・情報メディア学科・准教授

研究者番号: 50290713