

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 25 年 5 月 23 日現在

機関番号：34310
 研究種目：基盤研究(C)
 研究期間：2010～2012
 課題番号：22560233
 研究課題名（和文） シリコンコンサートホールのための音響レンダリング手法の高精度・高速化
 研究課題名（英文） High accuracy and high speed sound field rendering for the silicon concert hall
 研究代表者
 土屋 隆生 (TSUCHIYA TAKAO)
 同志社大学・理工学部・教授
 研究者番号：20217334

研究成果の概要（和文）：本研究では、バーチャルにコンサートホールの音響を計算するシリコンコンサートホールのための音響レンダリング手法の高精度・高速化を検討している。レンダリング手法にコンパクト時間領域差分法を採用し、GPU クラスタに実装することで使用メモリや計算時間などの検討を行っている。その結果、IWB法が最も計算機資源が少なく、かつ精度も良いことが示している。さらに、レンダリング結果を 24ch スピーカアレイシステムに出力することで可聴化実験を行ったところ、自然な残響を伴った立体音場が再生可能なことを示している。

研究成果の概要（英文）：In this study, high accuracy and high speed numerical schemes for the sound field rendering of the silicon concert hall are examined. The compact explicit finite difference time domain (CE-FDTD) method is applied to large-scale 3D sound field analysis to reduce computer resources. The CE-FDTD method is implemented on a GPU (Graphics Processing Unit) cluster system. The calculation performance is evaluated for the 3D sound wave propagation. The most accurate scheme is the interpolated wideband (IWB) scheme in which the cut-off frequency is in agreement with Nyquist frequency. The rendering results are then reproduced by the 24ch loudspeaker array system. It is confirmed that the high quality and high definition auralization can be realized.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	700,000	210,000	910,000
2011年度	600,000	180,000	780,000
2012年度	700,000	210,000	910,000
総計	2,000,000	600,000	2,600,000

研究分野：工学

科研費の分科・細目：機械工学・機械力学・制御

キーワード：シリコンコンサートホール、音響レンダリング、CE-FDTD法、GPU、数値分散

1. 研究開始当初の背景

コンピュータグラフィックスの分野では、3次元形状データや色データを基にコンピュータ上に仮想空間を構築し、ディスプレイに画像を提示するレンダリング技術が確立され、広く普及している。これと同様に、形状データや音響特性データを基に3次元仮想音響空間をコンピュータ上に構築し、臨場感あ

ふれる音響情報を数値的に提供する音響レンダリング技術が提案されている。音響レンダリングは、時間領域の音場解析により仮想的な受聴者が受聴する音圧波形を数値的に計算する手法である。特に、コンサートホールのような大規模な音響空間をリアルタイムで音響レンダリングするシステムを"シリコンコンサートホール(SiCホール)"と呼び、

その実現のための基盤技術の開発が行われている。SiC ホールは、FPGA (Field Programmable Gate Array) や GPU (Graphics Processing Unit) などのハードウェアデバイスを用いて超高速に音響レンダリングを行うことで、家庭に居ながらにしてコンサートホールの音響を体感可能にするシステムを目指している。

SiC ホールでは、レンダリング手法に時間領域差分(FDTD)法系の計算手法を採用している。しかしながら、差分系の数値解析手法では伝搬波長が空間サンプリング周波数に近づくと従い、計算上の伝搬速度が低下するという数値分散誤差が顕著に現れることが知られている。この数値分散誤差により、大規模空間の解析において伝搬距離が増えるほど耳障りな高周波音が遅れて聞こえることになり、コンサート鑑賞どころではなくなる恐れがある。この誤差は、オーバーサンプリングとフィルタ処理により回避可能ではあるが、計算機資源が大きくなり過ぎ、実用的とは言えない。この問題を解決し、高精度なレンダリングが可能となれば、より少ない計算機資源で、より高音質なレンダリングが可能となり、音響レンダリング技術の研究を大きく前進させることが可能となる。

2. 研究の目的

本研究は、上述の数値分散誤差を回避するために、高精度な波動音響解析手法を実装し、高音質な音響レンダリングを実現することを目的としている。計算アルゴリズムとしては、空間精度を向上させた CE-FDTD (Compact Explicit-Finite Difference Time Domain) 法を採用し、高精度化を実現する。本研究では、CE-FDTD 法により数値分散誤差を低減するとともに、計算時間も短縮できることを理論的に示す。また、CE-FDTD 法を GPU (Graphics Processing Unit) に実装した場合の計算性能等を評価する。さらに、本手法の効果を確認するために、24ch のスピーカレイシステムにレンダリング結果を出力し、可聴化による評価も行う。

3. 研究の方法

(1) CE-FDTD 法の適用

CE-FDTD 法では、音圧 p に関する 3 次元波動方程式を直接離散化する。コロケートグリッド上で図 1 のような $2D \times 2D \times 2D$ (Δ はグリッド間隔) の離散化領域(セル)を考え、セル内の 27 の格子点を用いて離散化する。ただし、空間は一様と考えるため x, y, z 方向のグリッド間隔を等しく Δ と考えている。CE-FDTD 法も本質的には中心差分であるが、通常の FDTD 法のように座標軸に沿った離散化ではなく、斜めや対角線方向も考慮して次式のように離散化する。

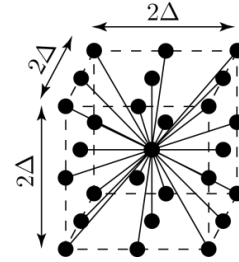


図 1 CE-FDTD 法のセル

$$p_{i,j,k}^{n+1} = d_1(p_{i+1,j,k}^n + p_{i-1,j,k}^n + p_{i,j+1,k}^n + p_{i,j-1,k}^n + p_{i,j,k+1}^n + p_{i,j,k-1}^n) + d_2(p_{i+1,j+1,k}^n + p_{i+1,j-1,k}^n + p_{i-1,j+1,k}^n + p_{i-1,j-1,k}^n + p_{i,j+1,k+1}^n + p_{i,j+1,k-1}^n + p_{i,j-1,k+1}^n + p_{i,j-1,k-1}^n + p_{i-1,j+1,k+1}^n + p_{i-1,j+1,k-1}^n + p_{i-1,j-1,k+1}^n + p_{i-1,j-1,k-1}^n) + d_3(p_{i+1,j+1,k+1}^n + p_{i+1,j+1,k-1}^n + p_{i+1,j-1,k+1}^n + p_{i+1,j-1,k-1}^n + p_{i-1,j+1,k+1}^n + p_{i-1,j+1,k-1}^n + p_{i-1,j-1,k+1}^n + p_{i-1,j-1,k-1}^n) + d_4(p_{i,j,k}^n - p_{i,j,k}^{n-1}) \quad (1)$$

ただし、位置 $(i\Delta, j\Delta, k\Delta)$ 、時刻 $t = n\Delta t$ における音圧を $p_{i,j,k}^n$ と表記し、 $d_1 \sim d_4$ は係数、 $\chi = c_0 \Delta t / \Delta$ は CFL 数である。 $d_1 \sim d_4$ のパラメータを調節することで、表 1 のような様々な手法が提案されている。ここで、 χ_m は CFL 数の上限値、 \bar{f}_c は規格化カットオフ周波数である。

表 1 CE-FDTD 法各パラメータ

手法	d_1	d_2	d_3	d_4	χ_m	\bar{f}_c
SLF	1/3	0	0	0	$1/\sqrt{3}$	0.196
CCP	0	1/4	0	-1	1	0.333
OCTA	0	0	1/4	0	1	0.25
IISO	1/4	1/8	0	-1	$\sqrt{3}/2$	0.333
IISO2	15/48	3/32	1/64	-9/8	$\sqrt{3}/2$	0.333
IWB	1/4	1/8	1/16	-3/2	1	0.5

(2) GPU クラスタへの実装

本課題では図 2 のような 8 ノードのコンピュータから構成される GPU クラスタシステムを開発した。1 つのノードには GPU を 4 基搭載している。したがって、システム全体で 32 基の GPU を並列動作可能である。1 ノード内の各 GPU 間の通信は、PCI Express × 16 を介して行い、各ノード間の通信は、InfiniBand を介して行う。

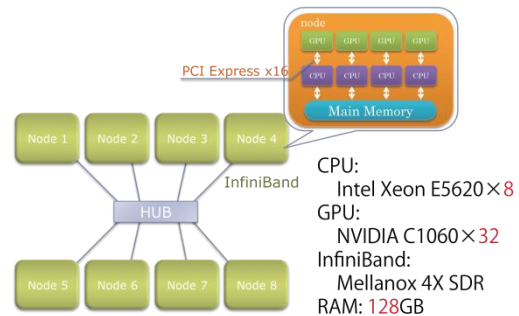


図 2 GPU クラスタ

使用する GPU は、NVIDIA 社製の Tesla M2075 (6GB RAM, 1.5GHz, 448 CUDA core) である。システム全体では 192GB のメモリが使用可能である。マザーボードは、SUPERMICRO 社製の X8DTG-QF を使用し、CPU は、Xeon E5620 (4 Core, 2.4GHz) を搭載している。このマザーボードには、PCI express×16 が 4 基搭載されている。InfiniBand ボードは、Mellanox 社製の MHES18- XTC QDR (10Gbit/s) を用いている。搭載 OS は、CentOS 5.5 である。使用言語は、CUDA 4.1, OpenMPI 1.4.2, OpenMP 2.5 である。

(3) 24ch スピーカアレイシステムによる可聴化
レンダリング結果は、図 3 に示す 24ch の半球ドーム型スピーカアレイへ出力し、可聴化による評価を実施する。このシステムは、大きさ 4.91×4.00×2.50 m の防音室内に 24 個のスピーカをドーム状に配置している。スピーカアレイに対応した半球状の仮想境界面をレンダリングモデルに想定し、24 の仮想受音点位置で 2 秒間のインパルス応答を計算する。ただし、計算点では単一指向性を仮想境界面の外向きに想定し、法線方向から内部へ入射する音波を記録する。

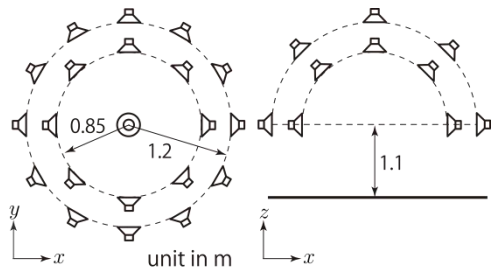


図 3 21ch スピーカアレイシステム

可聴化では法線入射インパルス応答に、無響室録音されたオーケストラの各パートの楽曲を畳み込む。パートごとに 24ch 分の法線入射インパルス応答の畳み込みを行った後、ミックスしてオーケストラ楽曲を再現する。レンダリングを実施した曲は、ベートーヴェンの交響曲第 7 番第 1 楽章、マーラーの交響曲第 1 番第 4 楽章、モーツァルトの歌劇ドン・ジョヴァンニからドンナ・エルヴィラのアリアである。

4. 研究成果

(1) 音質を考慮したレンダリング条件

①理論性能の推定

レンダリングの音質は、規格化カットオフ周波数 f_c で制限される。ここで、再生帯域を f_b とすると、グリッド間隔 Δ は次式で決定される。

$$\Delta = \frac{c_0 \sin^{-1} \chi}{\pi \chi f_b} \quad (2)$$

例えば、 $f_b=20\text{kHz}$ の音質(帯域幅)を確保する場合を考えると、 χ と Δ の関係は図 4 のようになる ($c_0=340\text{ m/s}$)。このように、音質を考慮すると、 χ が大きいほど Δ を大きくとれることになり、レンダリングでは CFL 数は上限値 χ_m に設定するのが望ましいことがわかる。表 2 は 1m^3 の空間を $f_b=20\text{kHz}$ の音質でレンダリングする場合の総格子点数 N と 1s あたりの計算量 N_f である。メモリは N に比例するため、音質を考慮すれば IWB は SLF の 1/3 以下のメモリ量で計算できることが分かる。特に GPU 等の高速演算素子を使用した場合は、計算時間のほとんどがメモリ転送時間で消費されることになるため、メモリ使用量の少ない手法ほど計算時間も有利となる。

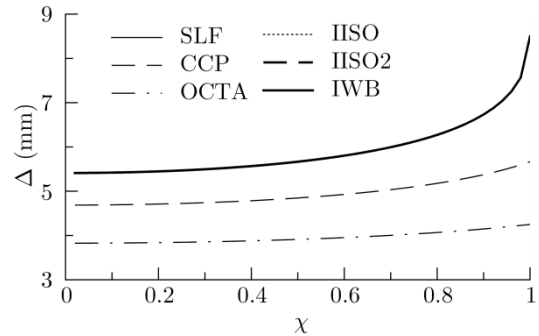


図 4 CFL 数とグリッド間隔の関係

表 2 1m^3 のレンダリングに必要な計算資源

手法	f_s (kHz)	Δ (mm)	$N \times 10^6$	$N_f \times 10^{12}$
SLF	102	5.77	5.20	4.78
CCP	60	5.66	5.51	4.97
OCTA	80	4.25	13.0	11.5
IISO	60	6.54	3.58	4.52
IISO2	60	6.54	3.58	6.67
IWB	40	8.50	1.63	2.02

②GPU による実測

各手法について、音質を考慮した場合の性能評価のために数値実験を行った。20kHz の音質を確保する場合について、各手法に必要な計算資源を計測した。計算モデルは、 1m^3 の立方体を想定し、1s 間のインパルス応答を計算するのに要する計算時間を各手法で比較した。まず、1つの GPU 上で実験を行った。手法による比較のために境界条件処理は行っていない。図 5 は、SLF の計算時間を 1 とした場合の実測結果である。全体として、GPU の共有メモリを使用することでさらに高速化できていることがわかる。特に、IWB は SLF に対して計算時間が 14% となっている。これ

は、GPU の場合、演算が高速であるためメモリ転送に多くの時間が費やされるが、GPU スレッド間で共有されるメモリを利用することでメモリ転送の重複が避けられ、効率的に演算を実行できたためである。

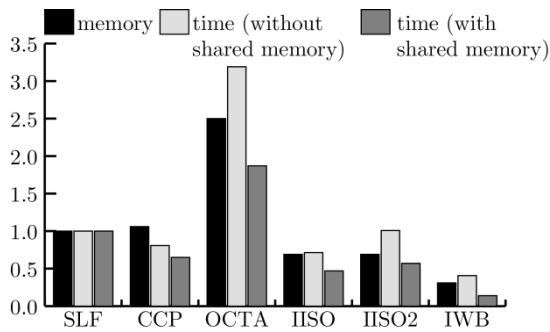


図5 音質を考慮した場合の計算資源比較 (シングル GPU)

次に GPU クラスタについて、計算モデルを 10m^3 の立方体と想定し、 0.1s 間のインパルス応答の計算に要する時間を各手法で比較した。図6はSLFの必要メモリ量および計算時間を1とした場合の実測結果である。IWBはSLFに対して、メモリ量で36%、計算時間は19%になるという結果が得られたが、シングルGPUの場合ほど高速化されていない。これは、GPU クラスタでは計算ノード間でのデータ交換のために通信を行う必要があり、この通信部分が手法により高速化できないためである。

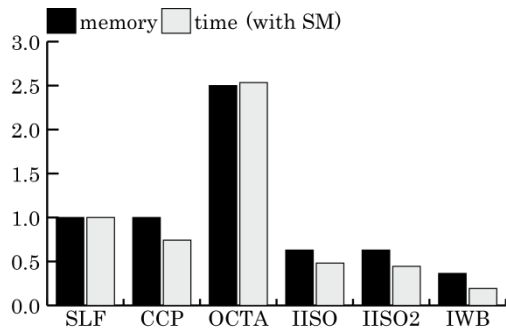


図6 音質を考慮した場合の計算時間比較 (GPU クラスタ)

(2) インパルス応答

つぎに、精度確認のために立方体モデルについてインパルス応答を計算した。1辺を256分割し、サンプリング周波数を40kHzとした。CFL数は各手法の上限値 χ_m に設定し、グリッド間隔はSLF: $\Delta = 5.77\text{mm}$ 、IWB: $\Delta = 8.5\text{mm}$ とした。音源は1/8セル(8点)に同時にインパルス音圧を与え、室中央から放射した。領域境界はすべて反射率を0に設定している。各手法とも $8000\Delta t = 0.2\text{s}$ までの応答を計算する。位置 $(64\Delta, 128\Delta, 128\Delta)$ で計算したインパルス応答とそのスペクトログラムを図7に示す。各手法とも、壁面の反射率が

0であるにも関わらず、残響のようなリングングが生じている。これは、数値分散により高周波成分に遅れが生じていることが原因である。さらに、各応答ともスペクトログラムに顕著なピークが存在するのが分かる。これらは、カットオフ周波数に対応しており、その周波数も理論的に予想された値にほぼ一致しているのが分かる。このように、IWBはSLFに比べ帯域幅が広いことが分かる。

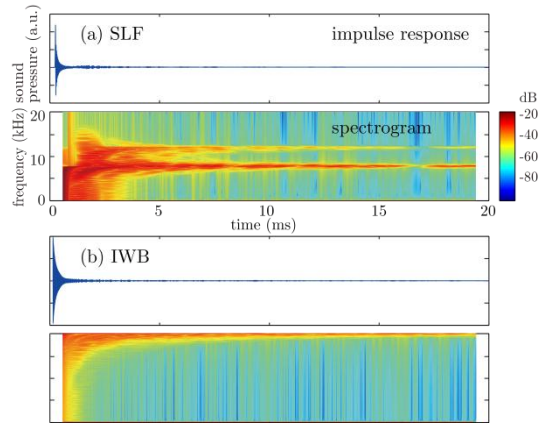


図7 インパルス応答とそのスペクトログラム

(3) 可聴化実験

最後に、 $18.5 \times 25.2 \times 9.8\text{m}$ の直方体室($4,531\text{m}^3$)を想定し、可聴化実験を行った。レンダリングはIWBを用いて実施する。 $f_b = 20\text{kHz}$ の場合、格子点間隔は約 8.5mm となるため、領域を $2176 \times 2944 \times 1152$ に分割することになる。領域境界はすべて反射率0.92に設定する。

レンダリングは、パート別に無響室録音されたオーケストラ音源に対して行う。パートごとに点音源を想定し(計18点)、インパルス応答を計算する。レンダリング結果は、24chスピーカアレイシステムへ出力する。レンダリングモデルの $(9, 5, 1)\text{m}$ の位置にスピーカアレイに対応した仮想境界面を想定し、24の受信点位置で2秒間のインパルス応答を計算する。ただし、計算点では単一指向性を仮想境界面の外向きに想定し、法線方向から内部へ入射する音波を記録する。

図8(a)は法線入射インパルス応答の計算例とそれに対応する残響曲線である。計算時間は1パート約12時間、計216時間であった。モデルが矩形鏡面反射場であるため、多重反射の存在も確認できる。この曲線から残響時間は 2.25s と計算された。また、波線で示される理論とも良く一致し、自然な残響場が形成されているのが分かる。

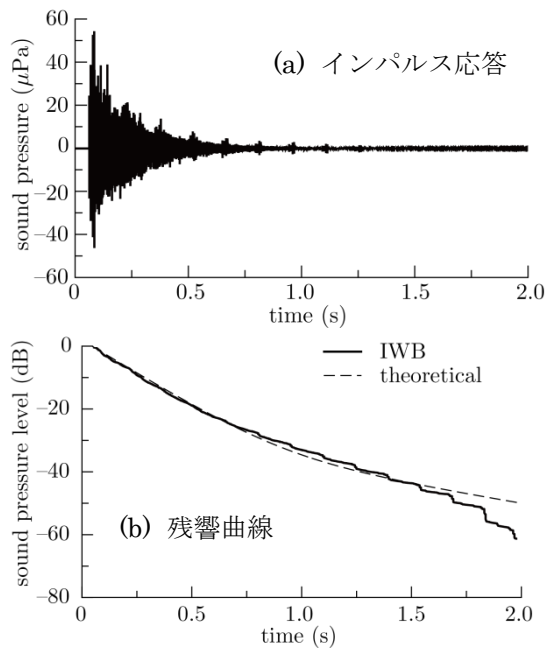


図8 インパルス応答と残響曲線

可聴化ではこの法線入射インパルス応答に、無響室録音されたオーケストラの各パートの楽曲を畳み込む。パートごとに24ch分の法線入射インパルス応答の畳み込みを行った後、ミックスしてオーケストラ楽曲を再現する。こうすることでパート位置を考慮した拡がりのある音場をスピーカアレイ内部に再現できる。レンダリングを実施した曲は、マーラーの交響曲第1番第4楽章等であるが、各曲とも音源を無指向性としたものと単一指向性とした場合についてそれぞれレンダリングした。

レンダリングの結果、自然な残響を伴った立体音場が再現できることが確認された。また、SLFよりもIWBの方が高域まで伸びがあることが確認された。アレイ内のどの位置においても音像が定位することも確認され、同時に複数人の聴取が可能であることが確認された。このようにIWB法を用いれば、通常のSLF法よりも高音質で実用的なレンダリングが行え、また24chのスピーカアレイシステムに出力することで高精細な可聴化が行えることが確認された。

このように、実用的な室容積およびサンプリング周波数でレンダリングを行った例は、これまで国内外において存在せず、現時点での世界最高水準である。今後は、この研究を突破口として、音響レンダリング技術が飛躍的に発展することを期待したい。

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2件)

- ① T. Ishii, T. Tsuchiya and K. Okubo, Three-dimensional sound field analysis using compact explicit-finite difference time domain method with graphics processing unit cluster system, Jpn., J. Appl. Phys., 査読有, 2013, 掲載予定.
- ② Y. Miyazaki and T. Tsuchiya, Perfectly Matched Layer for the Wave Equation Finite Difference Time Domain Method, Jpn., J. Appl. Phys., 査読有, Vol.51, 2012, pp.07GB02-1-5.

[学会発表] (計3件)

- ① 土屋隆生, 石井琢人, 大久保 寛, コンパクトで陽的な高精度 FDTD 法について, 日本音響学会 2012 年秋季研究発表会, 2012.9, 信州大学.
- ② 宮崎友孝, 土屋隆生, WE-FDTD法におけるPML境界について, 日本音響学会 2011 年秋季研究発表会, 2011.9, 島根大学.
- ③ T. Tsuchiya, A. Morikochi, and T. Ishii, A High Performance Numerical Simulation of Sound Field in Time Domain Using a GPU Cluster System, Proc. Forum Acusticum 2011, 2011.6, Aalborg, Denmark.

[図書] (計0件)

[その他]

特になし

6. 研究組織

(1) 研究代表者

土屋 隆生 (TSUCHIYA TAKAO)
同志社大学・理工学部・教授
研究者番号: 20217334

(2) 研究分担者

該当なし

(3) 連携研究者

該当なし