

FDTD法に基づく声道音響解析におけるGPGPU導入の効果

☆福庭隆弘 (甲南大学大学院自然科学) 北村達也(甲南大学知能情報)

1.背景

MRI撮像技術向上

発話中の声道形状の
精密な観測が可能に

MRI装置(ATR-P)

FDTD法(時間領域差分法)

単純な計算
で解析可能メモリの
大容量化計算機の
高速化3次元空間の解析も**現実的に**

Takemoto et al. (2010)

FDTD法により十分な精度で
声道音響解析が可能なることを実証声道形状と音声の対応づけ、
音声生成メカニズムの解明へ

2.目的

FDTD法に基づく3次元声道音響シミュレータの開発

FDTD法(3次元空間)は計算量が多く、処理に時間がかかる

高速化が必要マルチコアCPUやGPUを用い
て並列計算により計算時間
の短縮を目指す

様々な声道や条件での解析

声道形状による音声合成
に向けて

声道

3.FDTD法

波動方程式

p : 音圧
 u_x, u_y, u_z : 粒子速度
 t : 時間
 ρ : 媒質の密度
 κ : 弾性体積率

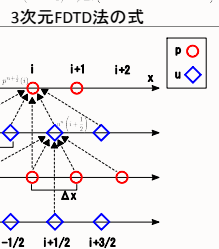
差分化
(時間的,空間的)空間軸を
4次精度で差分化

$$p^{n+1}(i,j,k) = p^n(i,j,k) - \frac{\Delta t}{\Delta x} \left\{ u_x^n \left(i+\frac{1}{2}, j, k \right) - u_x^n \left(i-\frac{1}{2}, j, k \right) \right\} \\ - \frac{\Delta t}{\Delta y} \left\{ u_y^n \left(i, j+\frac{1}{2}, k \right) - u_y^n \left(i, j-\frac{1}{2}, k \right) \right\} \\ - \frac{\Delta t}{\Delta z} \left\{ u_z^n \left(i, j, k+\frac{1}{2} \right) - u_z^n \left(i, j, k-\frac{1}{2} \right) \right\}$$

$$u_x^{n+1} \left(i+\frac{1}{2}, j, k \right) = u_x^n \left(i+\frac{1}{2}, j, k \right) - \frac{1}{\rho \Delta x} \left\{ p^{n+1}(i,j+1,k) - p^{n+1}(i,j,k) \right\} \\ + \frac{1}{\rho \Delta x} \left\{ p^n \left(i+\frac{1}{2}, j, k \right) - p^n \left(i-\frac{1}{2}, j, k \right) \right\}$$

$$u_y^{n+1} \left(i, j+\frac{1}{2}, k \right) = u_y^n \left(i, j+\frac{1}{2}, k \right) - \frac{1}{\rho \Delta y} \left\{ p^{n+1}(i,j,k+1) - p^{n+1}(i,j,k) \right\} \\ + \frac{1}{\rho \Delta y} \left\{ p^n \left(i, j+\frac{1}{2}, k \right) - p^n \left(i, j-\frac{1}{2}, k \right) \right\}$$

$$u_z^{n+1} \left(i, j, k+\frac{1}{2} \right) = u_z^n \left(i, j, k+\frac{1}{2} \right) - \frac{1}{\rho \Delta z} \left\{ p^{n+1}(i,j,k+1) - p^{n+1}(i,j,k) \right\} \\ + \frac{1}{\rho \Delta z} \left\{ p^n \left(i, j, k+\frac{1}{2} \right) - p^n \left(i, j, k-\frac{1}{2} \right) \right\}$$

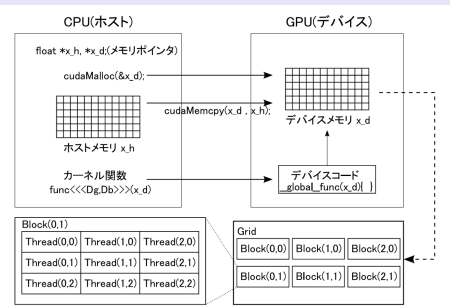


FDTD(2,4)法の空間配置と時間配置

3次元FDTD(2,4)法の式

5.シミュレータの高速化(GPU + NVIDIA CUDA)

- 多数のスレッドを生成(最大65535 × 65535 × 512個)
- 多数のSP(最小演算ユニット)で並列演算



基本的な配列計算

処理手順

- 1.CPU側, GPU側でメモリ確保
- 2.CPU側からGPU側にメモリコピー(初期値)
- 3.カーネル関数起動(スレッド数を指定)
- 4.配列のインデックスに対応させて計算

CUDAによる高速化のポイント

- 3次元空間を1次元配列で表現
- 声道周辺の計算しない領域もメモリ確保

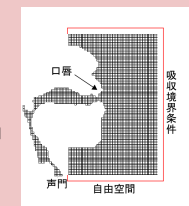
前処理を行い、条件分岐を行わずに解析領域のみのメモリアクセスを行う

- ホスト(CPU)とデバイス(GPU)間のメモリコピーはできるだけ抑える
- プロファイラを用いて各処理(カーネル関数)の占める計算時間を調べ、ボトルネックとなっている処理を探す
- スレッド数等はGPUに最適な値に設定する(様々な値で試す)

4.シミュレーション

MR画像
(ATR母音発話データベース)

二値化画像



3次元解析空間

- 画像の解像度をダウンサンプリング
- 目視によって声道を抽出
- 声門を水平にするために回転

自由空間を追加

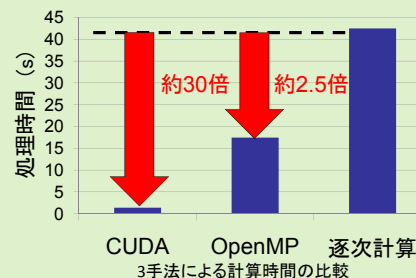
吸収境界条件

シミュレーション条件

離散空間間隔 $\Delta x = \Delta y = \Delta z$	2.00 mm
離散時間間隔 Δt	2.0×10^{-8} s
解析空間	110 × 90 × 90 voxel
音速 c	352.9 m/s
媒質の密度 ρ	1.12 kg/m ³
吸収境界条件	Murの2次吸収境界条件
入力音源	ガウシアンパルス

- 日本人成人男性1名による日本語5母音(/a/, /i/, /u/, /e/, /o/)の発話
- 発声同期画像法により画像のぶれを抑え、S/N向上
- 歯列撮像法および歯列補填法により歯列のはめこみ

6.評価実験



まとめ

- 3次元声道音響シミュレータ開発
- GPUを利用した並列演算により逐次計算の約30倍の高速化に成功
- PCクラスタ等よりも低コストかつ小規模な環境で実現



使用GPU(Tesla C1060)のスペック

シェーダクロック	1.3 GHz
メモリクロック	16 GHz
メモリバンド幅	102 GB/s
デバイスメモリ	4 GB
SP数	240 個
SM数	30 個



計算機のスペック

CPU	Intel Core i7 920
CPUクロック数	2.66 GHz
CPUのコア数	4 個
メインメモリ	3 GB

本研究は2010年度科研費(21300071)の支援を受けた

本研究で利用したMRIデータは、ATR人間情報科学研究所が独立行政法人情報通信研究機構からの研究委託「人間情報コミュニケーションの研究開発」に基づいて収録し、公表した「ATR母音発話MRIデータ」の一部です。本データの使用および成果の発表は、株式会社ATR-Promotionsとの使用許諾契約に基づいています。