

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 9 日現在

機関番号：13301

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24560446

研究課題名(和文)高音質・高速収束ステレオエコーキャンセラに関する研究

研究課題名(英文)Studies on high-quality and fast-convergence stereophonic echo cancellation

研究代表者

平野 晃宏(Hirano, Akihiro)

金沢大学・電子情報学系・講師

研究者番号：70303261

交付決定額(研究期間全体)：(直接経費) 4,200,000円

研究成果の概要(和文)：ステレオエコーキャンセラの学習速度を改善するための手法として、複数の全域通過フィルタ(APF)を縦続接続した前処理方式と、複数のFIRフィルタを切り替える方式を検討した。APFを用いる方式では、極の角周波数を周期的に変化させる区間と極半径を周期的に変化させる区間を交互に繰り返す方式を開発した。複数のFIRフィルタを用いる方式では、1サンプル未満の遅延変化を与えることで収束速度が最大になることを示した。また、非線形フィルタを組み合わせることも検討した。主観評価により、音質の改善を確認した。CUDAやOpenCLを用いたエコーキャンセラの並列実現も検討した。

研究成果の概要(英文)：As fast convergence algorithms for stereophonic acoustic echo cancellation, two pre-processing filters are studied. In pre-processing filters using multiple all-pass filters (APFs), the angular frequency and the radius of the pole periodically controlled one after the other. Pre-processing filters using multiple FIR filters are also studied. Sub-sample inter-channel delay difference maximizes the convergence speed. Combination with non-linear pre-processors are also examined. Subjective test results show the improved sound quality. Parallel and effective implementation of echo cancellers by using CUDA or OpenCL is also studied.

研究分野：工学

キーワード：信号処理 エコーキャンセラ 前処理 並列処理

1. 研究開始当初の背景

情報ネットワークの普及やマルチメディア利用技術の進展により、映像、音声、データなどを統合したマルチメディア遠隔会議システム需要が高まっている。通信容量の増加にともない、臨場感のあるシステムが強く求められており、音声の多チャンネル化が不可欠となっている。そのため、遠隔会議システムが多チャンネル化が検討されている。

遠隔会議においては快適な会話の妨げになる音響エコーが問題であるため、多チャンネル音声に対応したステレオエコーキャンセラ (SEC) の研究が進められている。SECにおける最大の問題は、係数不確定性の問題である。最適解以外に無数の不定解が存在している。最適解は全ての話者に対してエコーを除去できる状態である。不定解は特定の条件でのみエコーを除去できる状態であり、話者位置が変化すればエコーを除去できなくなる。SECにおいては無数の不定解が存在し、最適値に収束する保証がない。我々は、この問題を世界に先駆けて指摘し、SECの研究を活発化させる契機となった。

この問題を解決する手法としては、前処理を用いてスピーカから再生される音声を変形させる前処理方式が主流となっている。前処理方式はスピーカで再生する音声信号を前処理によって変化させるため、音声品質が劣化する。また、現在の技術では収束に数十秒を要するため、エコーを素早く除去して快適な会話を行なうには収束速度は十分とはいえない。したがって、音質劣化を最小限度に抑えながら収束速度を高める必要がある。

2. 研究の目的

本研究の目的は、前処理による音質劣化を抑え、かつ、フィルタ係数が最適値に高速に収束する SEC を開発することである。本研究では、従来はトレードオフの関係にあったエコー消去量と係数誤差の両方を同時に改善することを試みる。

また、SECはチャンネル数の二乗に比例する膨大な演算量を必要とし、また、チャンネル間相関の影響で収束が遅くなるために RLS などの高速収束アルゴリズムが求められることから、並列処理による効率の良い実現法も検討する。

3. 研究の方法

本研究では、下記について検討を行った。

(1) 複数の APF を用いた前処理フィルタ

図1に示す、中心周波数の異なる複数の2次 APF (All Path Filter, APF) を縦続接続したものが検討されている。各帯域の遅延を独立に制御することで、遅延変化による音像の移動を知覚されにくくする。

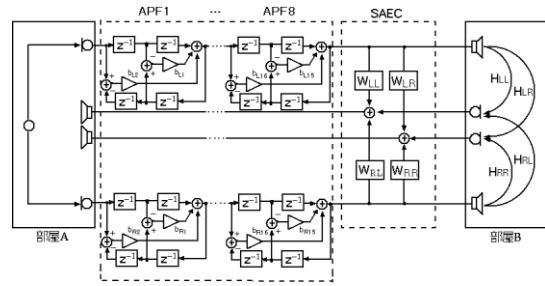


図1. 2次 APF 縦続前処理フィルタ

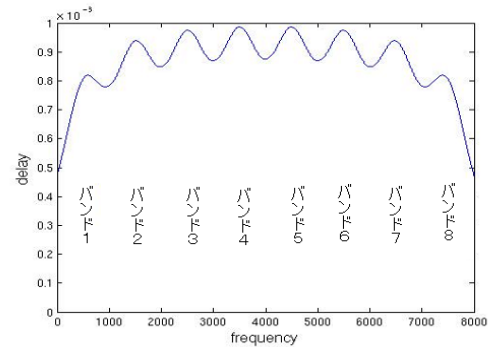


図2. 前処理ユニットの遅延特性

図2に遅延特性の一例を示す。極半径を周期的に変化させることで、遅延量を調整する。偶数バンドと奇数バンドで異なる極半径を使用する。

本研究では、まず、APF を縦続接続させる構成において、バンド数を変化させた際の音質や収束特性の変化、極の中心周波数を変化させた際の音質や収束特性の変化を調査する。この結果に基づいて、APF 数や極位置制御方法を最適化する。

(2) 複数の FIR フィルタを用いた前処理フィルタ

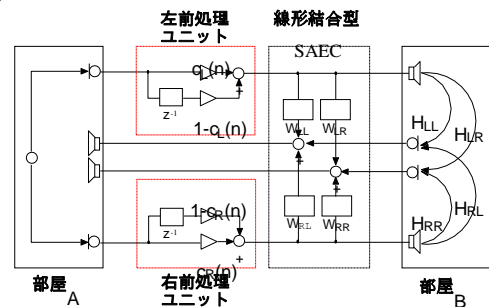


図3. 入力遅延法

図3に、現在最も良い性能を示している方法の一つである入力遅延法を用いた SAEC の構成を示す。左右両チャンネルに1サンプルの遅延を挿入・除去する操作を周期的に繰り返している。この手法の問題点は、遅延量がサンプリング周波数で規定されるために調整が困難であることである。

この問題を解決するために、FIR フィルタを用いて1サンプル未満の精度で遅延を調整

することを考える。遅延量を変化させ、最適な遅延変化を求める。また、非線形変換との併用も検討する。

(3) 音質の主観評価

隠れ基準付三刺激二重盲検法 (triple stimulus/hidden reference/double blind approach) を用いて主観評価を行い、前処理が音質に与える影響を評価する。

(4) 適応フィルタの高並列実現

近年、半導体技術の進歩により、複数の演算装置を有するプロセッサが通常のパーソナルコンピュータ (PC) にも採用されるようになってきている。中央演算処理装置 (Central Processing Unit, CPU) はベクトル演算が可能なコアを複数持つものが多い。また、グラフィックスプロセッサ (Graphics Processing Unit, GPU) は数百個から数千個の演算装置で並列処理とプログラミングが可能となっている。

これらの装置を用いて、エコーキャンセル処理を並列実行する手法を検討する。

4. 研究成果

(1) 複数の APF を用いた前処理フィルタ

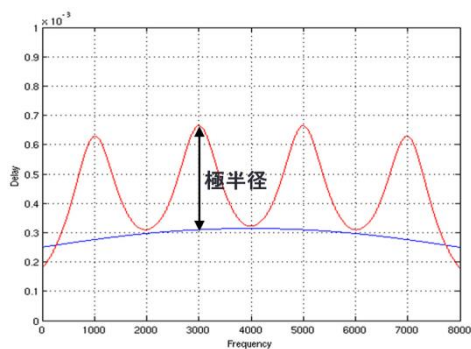


図 4. 4 個の APF による遅延特性

まず、前処理に用いる APF の個数を変化させ、収束特性の変化を調査した。図 4 に、APF を 4 個用いた場合の遅延特性を示す。従来と同様、極半径のみを周期的に変化させた。中心周波数を等間隔とし、4, 8, 16 個の APF を用いた場合を比較した。また、中心周波数を不等間隔とし、12 個の APF を用いた場合も比較した。

音質の評価結果より、APF 数を増加させると、極半径を小さく設定しないと音質が劣化することが確認できた。従来の 8 個を用いた場合と同等以上の音質を保った場合、12 個や 16 個を用いる構成では収束速度が低下した。APF を 4 個用いる構成では、音質の改善が可能な極位置制御もみられる。しかし、収束速度は低下した。この結果から、APF 数は 8 個が最適であると判断し、APF 数が 8 個の場合について、極位置制御方法を検討した。

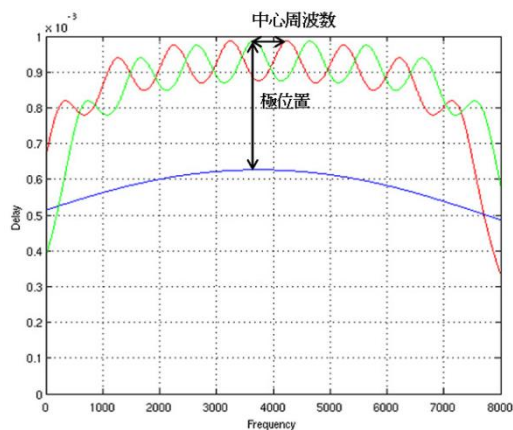


図 5. 極半径と中心周波数の制御

従来は極半径のみを周期的に変化させていたが、本研究では、極半径と中心周波数の両方を変化させる方法を開発した。極半径と中心周波数を同時に変化させると音質劣化が激しくなるため、極半径のみを変化させる期間と中心周波数のみを変化させる期間を繰り返すこととした。図 5 に極位置変化による遅延特性の変化を示す。

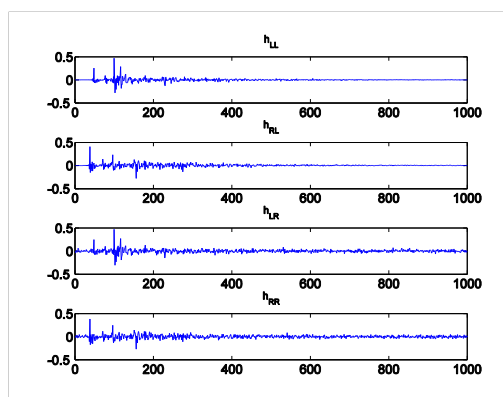


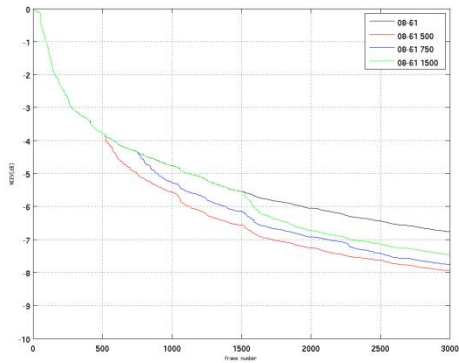
図 6. エコーパス

今回は、以下の条件で実験を行った。

- 入力信号: 音声 10 種類 (男声 5 種類, 女声 5 種類)
- サンプル周波数: 16kHz
- エコーパスのタップ数: 1000
- 適応フィルタのタップ数: 1000
- 適応アルゴリズム: NLMS アルゴリズム
- ステップサイズ: 0.5
- 遅延制御周期 Q: 4000
- 遷移域 L: 400
- 背景雑音レベル: -40dB

図 6 にエコーパス 4 個のインパルス応答を示す。

図 7 に係数誤差の変化を示す。黒線が従来法で、極半径のみを制御している。赤・青・緑線は、切り替え周期を 500, 750, 1500 サンプルとした場合の提案法である。極半径と中心周波数を交互に変化させることによって、収束特性を改善できることがわかる。



(2) 複数の FIR フィルタを用いた前処理フィルタ

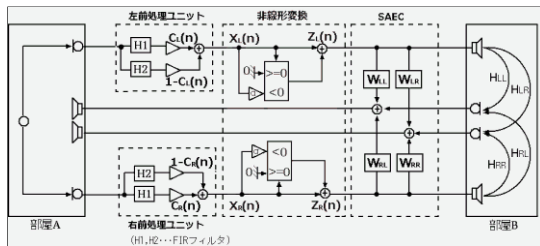


図 8. 2 個の FIR フィルタを用いた前処理

図 8 に、2 個の FIR フィルタを用いた前処理と非線形処理を組み合わせた場合を示す。FIR フィルタとしては、周波数領域で

$$H_i(\omega) = A_i(\omega)e^{-j\omega(N/2+t_i)}$$

を設計した。 $A_i(\omega)$ は振幅特性であり、 t_i は全帯域で共通の平坦遅延特性である。これを逆フーリエ変換してフィルタ係数を求める。

遅延変化と収束特性の関係を求めるため、まず、 $A_i(\omega) = 1$ として、 t_i のみを変化させた。2 個の FIR フィルタ $H_i(\omega)$ の遅延時間差 $t_2 - t_1$ の関係を数値的に解析した。

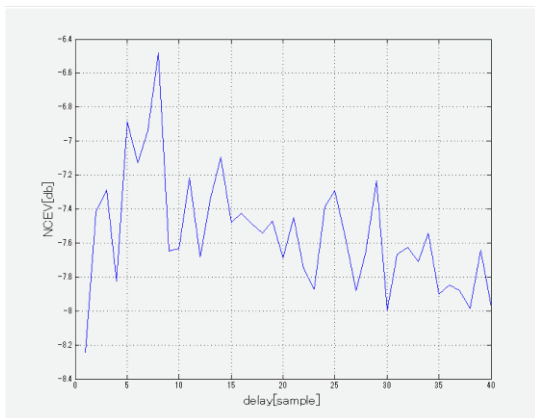


図 9. 遅延と係数誤差の関係 (遅延 1~40)

図 9~10 に遅延時間差と係数誤差の関係を示す。縦軸は学習後の係数誤差を示し、低いほど精度が高い。図 9 から、遅延変化を大きくすればするほど収束速度が速くなるわけではないことがわかる。

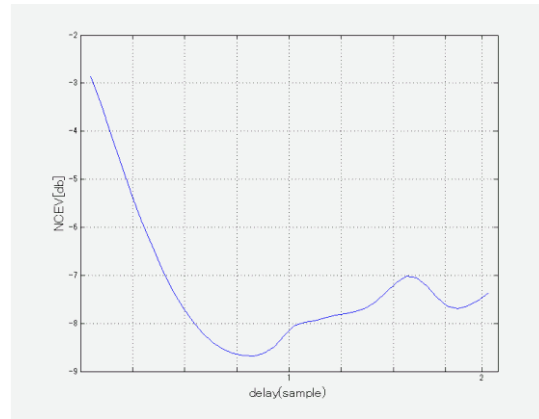


図 10. 遅延と係数誤差の関係 (遅延 0~2)

そこで、少ない遅延差の場合に係数誤差が少ないことに注目し、2 サンプル以下の実数値の遅延差を与えたものが図 10 である。この結果から、1 サンプル以下の遅延差を与えた時に収束速度が最も早くなっていることがわかる。

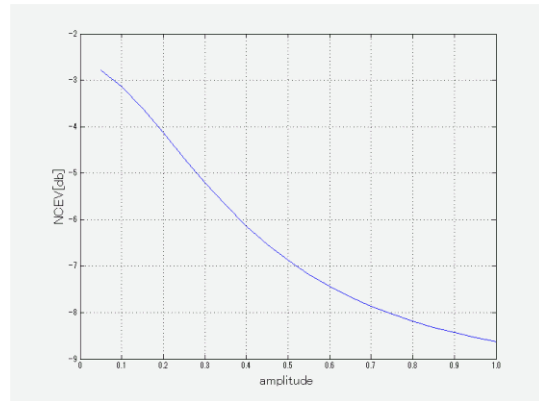


図 11. 振幅変化と係数誤差の関係

次に、遅延を変化させずに振幅変化を全域に一定の値で与えることで、振幅変化と係数収束の関係調べた。図 11 に振幅変化の大きさに対する NCEV の収束速度比較を示す。横軸は二つの前処理ユニットが入力信号に与える振幅の差を表しており、縦軸はそれぞれの値に対する係数誤差を表している。この結果から振幅変化は大きければ大きいほど収束速度が速くなっていることがわかる。

以上の結果から、振幅においては変化が大きければ大きいほど収束速度の高速化が見込めるので、音質劣化として知覚できないぎりぎりの大きさの変化幅を与え、遅延においては最も収束速度を高速化できる遅延差を全域に一定の割合で与えることを目標として、前処理フィルタを設計した。

前処理フィルタのタップ数が大きくなると遅延が大きくなる。また、演算量が増えるといった問題もあり、タップ数をできるだけ少なくするために、タップ数が少なくなった時にフィルタ設計にどのような影響があるか調査する必要がある。

前処理フィルタのタップ数は、設計を容易に

するために2のべき乗とすることにし、8タップから256タップまで調査した。32タップ以上は遅延特性が良好であり、64タップ以上では振幅特性も良好であることがわかった。そこで、前処理フィルタには64タップのFIRを用いることとした。

非線形変換とFIRフィルタを用いた方式を組み合わせ、その効果を調査することで非線形変換を加える量を決める。非線形処理前後の音声の波形を図12に、非線形成分を加える量を表すパラメータを $\alpha = 0.1 \sim 0.5$ まで変化させ、学習させた際の係数誤差を図13に示す。

図12から、 α を増やせば増やすほど収束速度の高速化が来ていることがわかる。係数誤差が-8dBまで下がった時間を比べてみると、非線形変換なしと比べて、 $\alpha = 0.1$ の時は2.8%、 $\alpha = 0.2$ の時は15.5%、 $\alpha = 0.3$ の時は26.5%、 $\alpha = 0.4$ の時は35.6%、 $\alpha = 0.5$ の時は41.4%収束時間を短縮している。

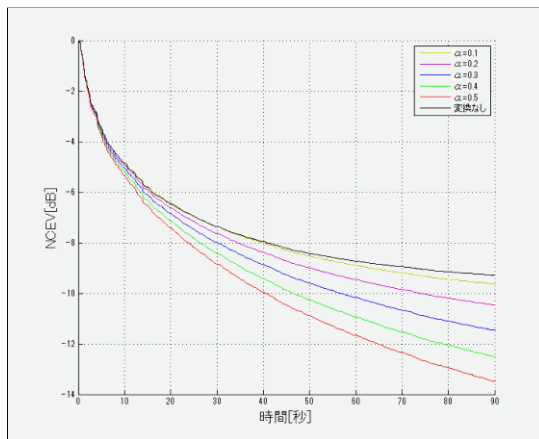


図12. 非線形変換の効果

次に音質の評価を行ったところ、 $\alpha = 0.3$ よりも大きいと音質の劣化を確認することが出来た。従来の研究においては、音質の劣化が許容できる範囲は $\alpha \leq 0.5$ とされていたが、提案法のSAECでは前処理ユニットで振幅の変化を与えるFIRフィルタを使用しているため、劣化が知覚されやすくなっていると考えられる。以上の結果から、 $\alpha = 0.3$ で非線形変換を行うこととした。

提案したSAECの収束速度高速化の効果を確認するためシミュレーションを行った。今回は、以下の条件で実験を行った。

- 入力信号: サンプル周波数 16kHz の音声 10 種類 (男声 5 種類, 女声 5 種類)
- エコーパスのタップ数: 1000
- 適応フィルタのタップ数: 1000
- 適応アルゴリズム: NLMS アルゴリズム
- ステップサイズ: 0.5
- フィルタ制御周期 Q: 4000
- 遷移域 L: 400
- 背景雑音レベル: -40dB

本研究では、次に示す6つの構成について実

験を行った。

- 線形結合型
- 従来法1 (入力遅延法 (両チャンネル入力遅延法))
- 従来法2 (入力遅延法+非線形変換)
- 提案法1 (遅延変化のみ)
- 提案法2 (遅延変化+振幅変化)
- 提案法3 (遅延変化+振幅変化+非線形変換)

各構成において、係数不確定性の問題を回避して正しいエコー経路に収束しているか否かの評価指標として、係数誤差NCEVを用いた。

図14より、係数誤差が-8dBまで低下するまでに要した時間は、従来法1と比較して、提案法1では23%、提案法2では25%の短縮となっている。非線形変換を加えた提案法3では48%の収束時間短縮が来ている。また、非線形変換を行った従来法2と比較すると、提案法3は24%の収束時間高速化が来ている。

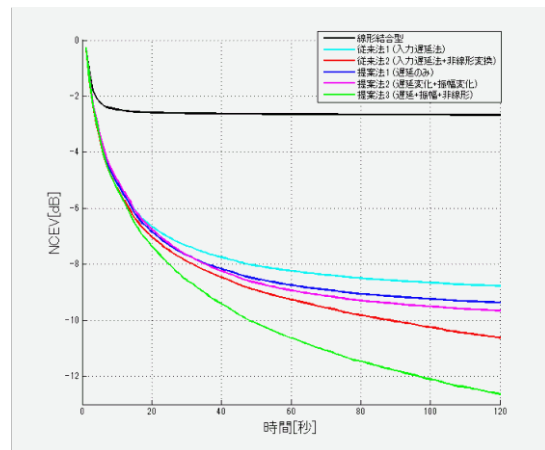


図13. 係数誤差の比較

120秒間学習させた後の係数誤差は、従来法1=-8.77dB、従来法2=-10.62dB、提案法1=-9.37dB、提案法2=-9.66dB、提案法3=-12.63dBとなった。

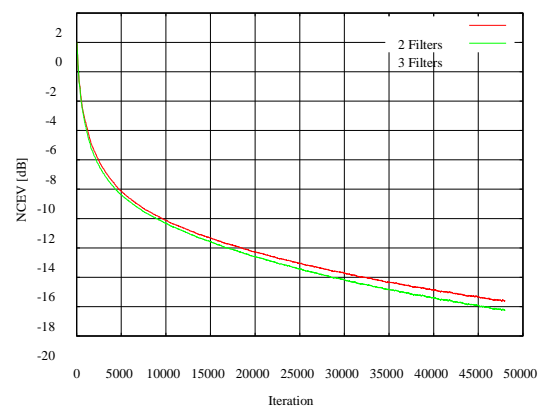


図14. 3個のFIRフィルタによる前処理

前処理に用いるFIRフィルタ数を2個から

3 個に増やした際の係数誤差の収束特性を図 14 に示す。方程式数を増やすことで精度を改善することを目指している。3 個のフィルタには、遅延が 1, 1.3, 1.6、振幅が 1, 1, 0.75 のものを用いた。係数誤差をわずかに改善しているが、大きな差は見られなかった。

(3) 音質の主観評価

前処理が音質に与える影響を確認するために主観評価を行った主観評価の手法として、隠れ基準付三刺激二重盲検法 (triple stimulus/hidden reference/double blind approach) を用いた。評価方法は ITU-R 5 段階劣化率に基づいている。隠れ参照音と処理音の提示順はランダムである。

今回は、次の 4 方式について評価を行った。

- 従来法 (入力遅延法 (両チャンネル入力遅延法))
- 提案法 1 (遅延変化のみ)
- 提案法 2 (遅延変化+振幅変化)
- 提案法 3 (遅延変化+振幅変化+非線形変換)

被験者は 15 人である。音声は、計算機実験で用いた 10 種類をそれぞれを約 10 秒切り出して使った。その他のパラメータは前章と同様である

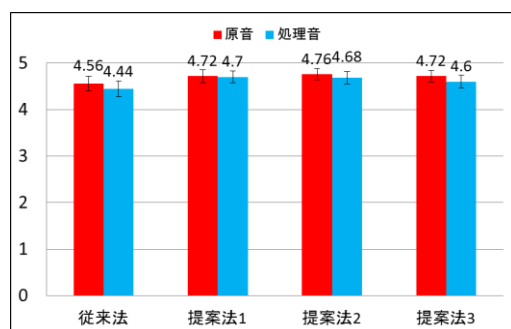


図 15. 主観評価結果

主観評価の結果を図 15 に示す。これは、すべての被験者と音声に対して平均した得点を棒グラフで示している。棒グラフの上端にある縦線の部分は、正規分布による 95%信頼性区間である。全体として、原音の得点より処理音の得点が低くなっており、どの前処理を用いても音質は劣化している。従来法の処理音の得点が 4.44 なのに対し、提案法 1 は 4.7、提案法 2 は 4.68、提案法 3 は 4.6 と、提案法のほうが得点の平均値が高い。

(4) 適応フィルタの高並列実現

本研究では、

- nVIDIA GeForce GPU と CUDA を用いた RLS アルゴリズムの高並列実現
 - OpenCL と各種 CPU と GPU を用いた NLMS アルゴリズムと SEC の並列実現
- を行った。シミュレーションの高速化や、PC ベースの装置試作に利用できる。

GeForce GPU と CUDA による RLS アルゴリズムの実装においては、従来は単一の SIMD アレイに限定していたものを、複数の SIMD アレイを用いるように改良した。8kHz のサンプリングで、512 タップの RLS アルゴリズムの実時間動作に成功した。

OpenCL を用いた NLMS 適応フィルタの並列実現においては、複数の CPU や GPU を用いた評価を行った。汎用 CPU に対しては、1 個のワークアイテムで 1 個のエコーキャンセラを実装するプログラムとし、ベクトル処理を用いた並列化のみを行った。ベクトル化によって最大 4 倍の性能向上がみられる。GPU を用いた実装では、多数のワークアイテムを用いた並列実現を行った。

SEC の実装においては、データ配置問題が発生しないベクトル化を行った。また、複数のワークグループを用いて、複数のシミュレーションを同時に実行する実験も行った。nVIDIA TESLA K20 を用いた場合、同時シミュレーション数 20 個以上まで、大きな計算時間の増加なしに並列化が可能であった。

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 0 件)

[学会発表] (計 3 件)

- ① Akihiro HIRANO, “Efficient implementation of Adaptive Filters Using OpenCL,” Proc. of 29th SIP Symposium, pp. 187-192, Nov. 13, 2014, 京都府京都市
- ② 宮城島良介, 平野晃宏, “前処理を用いたステレオエコーキャンセラの学習アルゴリズム,” 第 28 回信号処理シンポジウム講演論文集, pp. 647-652, Nov. 22, 2013, 山口県下関市
- ③ Akihiro HIRANO, Kenji Nakayama, “Efficient Implementation of RLS-Based Adaptive Filters on nVIDIA GeForce Graphics Processing Unit,” Proc. of 27th SIP Symposium, pp. 241-245, Nov. 29, 2012, 沖縄県石垣市

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

ホームページ等

<http://leo.ec.t.kanazawa-u.ac.jp/>

6. 研究組織

(1) 研究代表者

平野 晃宏 (HIRANO, Akihiro)

金沢大学・電子情報学系・講師

研究者番号：70303261

(2) 研究分担者

(3) 連携研究者