

平成21年度戦略的基盤技術高度化支援事業

「電話音声の高域復元による聴認度改善技術の研究開発」

研究開発成果等報告書

平成22年 3月

委託者 近畿経済産業局

委託先 日本ロジックス株式会社

電話音声の高域復元による聴認度改善技術の研究開発

目 次

第1章 研究開発の概要	4
1-1 研究開発の背景・研究目的及び目標	4
1-1-1 研究の概要	4
1-1-2 実施内容	5
1-2 研究体制	7
1-2-1 研究組織及び管理体制	7
1-3 成果概要	8
1-3-1 高域復元アルゴリズムの研究	8
1-3-2 ソフトウェア研究開発	8
1-3-3 アルゴリズム構築検証用振幅スペクトラムサンプル	8
1-3-4 DSPソフトIPの研究開発	8
1-3-5 HDLソフトIP仕様研究開発	8
1-3-6 プロジェクトの管理・運営	9
1-4 当該研究開発の連絡窓口	9
第2章 本論	10
2-1 高域復元アルゴリズムの研究	10
(1) 既存技術の性能検証	10
(2) 非線形信号処理を発展させた新方式	10
(3) 統計学習を発展させた新方式	11
(4) コードブックを発展させた新方式	13
2-2 ソフトウェア研究開発	13
(1) 非線形信号処理を発展させた新方式の実装と実験	13
(2) 統計学習を発展させた新方式の実装と実験	13
(3) コードブックを発展させた新方式の実装と実験	14
2-3 アルゴリズム構築検証用振幅スペクトラムサンプル	15
(1) Off-line 検証データ作成	15
(2) 擬似実環境検証データ作成	15
(3) 携帯電話実環境評価の準備	15
2-4 DSPソフトIPの研究開発	15
(1) DSP 搭載 DSK への実装	15
DSP 搭載 DSK ボードは、動作クロック 225MHz の DSP を搭載するアプリケーション開発 ボードである。この DSK ボードを用いて、携帯電話での実時間動作検証が行える環境構築 を行った。開発工程を以下の) ~) に示す。	15
(2) DSP 内蔵 ARM 搭載 BeagleBoard への実装	15
2-5 HDLソフトIP仕様研究開発	16
(1) アルゴリズム HDL 化検討	16

(2) HDL ソフト IP 仕様作成.....	16
(3) HDL 用ライブラリ作成.....	16
(4) FPGA システムボード作成	17
2-6 プロジェクトの管理・運営	18
第3章 全体総括.....	19

第1章 研究開発の概要

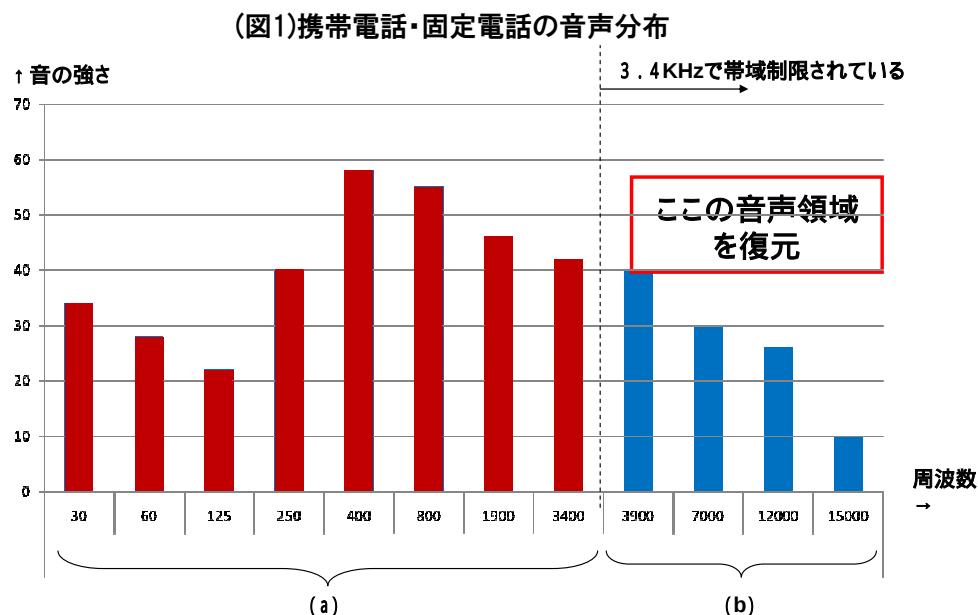
1-1 研究開発の背景・研究目的及び目標

1-1-1 研究の目的

人間の音声は図 1 の(a)+(b) のような分布になっている。携帯電話や固定電話の通信網では データ量を削減する為に高域信号(b)をカットして通信を行っている。

受信側では高域信号をカットした信号(a)を 再生するため 音声や音楽の品質として不満足を感じる場合が多い。本研究では高域信号(b)を 高精度な復元をし、クリアな音声や音楽の再生を実現する。

目標として従来技術(SINC 関数による補間)よりさらに 5db 以上の向上をめざす。



1-1-2 研究の概要

本研究では、3.4KHz に帯域制限された電話音声信号の高域を、高精度推定して復元する技術を開発する。具体的には、(1)Totalvariation 信号分離を行い、骨格信号と周期信号それぞれに新規の高解像度化技術を適用する手法と、(2)男女年齢等のグループ分けをしてそれぞれ最適な復元スペクトラムを生成する統計学習によるカテゴリー分類技術を適用する手法等を融合させる。また、このアルゴリズムを薄型カードまたは LSI に実装して製品化し、携帯電話と有線電話に導入する。高域復元にて電話音声が増えやすくなるので、高齢者や騒音環境下など市場は大きい。

1-1-3 実施内容

高域復元アルゴリズムの研究(学校法人立命館 立命館大学)

高域復元は5つの基本手法を考えている。

第一は、低域信号の全周波数に一樣となる非線形処理を行って、低域に比例した倍音を生成する手法であり、解析関数の絶対値距離またはユークリッド距離などを利用する。一部製品化の実績もある。

第二は、サンプリング点を倍に増やして DFT の様な直交変換を用いて、実空間と周波数空間を交互に移動して収束させていく手法である。画像高解像で実績があるが処理量が多くなる欠点がある。

第三は、実空間で SINC 関数の畳込みを行い、周波数空間にて LOW パスフィルタ帯域を広げるのに相当する処理を行う手法であり、我々はこれを改良した技術にて画像高解像度化を行っている。本課題では、この技術を音声高域復元に応用する。

第四は、音声波形を分離して、それぞれの分離波形に適切な高域復元処理を行って合成する手法である。完全再合成フィルタを使ってサブバンドに分け、その低域と高域との相関から復元高域を予測する手法と、積分型目的関数の最小化手法(Totalvariation 信号分離技術)でピッチ成分とランダム成分に分離し、それぞれに適切な高域復元を行う手法である。いずれも本課題で新たに提案している手法である。

第五は、男女年齢等によりグループ分けをして、それぞれ最適な復元スペクトラムを生成する手法である。グループ分けは、ケプストラム等から抽出した特徴量を KL 変換して直交サブ空間にて行う(統計学習によるカテゴリー分類技術)。またグループ毎に、高域復元手法のチューニングを行う。これも本課題で新たに提案している手法である。

本研究では、上記5つの基本手法を MATLAB にてプログラミングし、大まかなアルゴリズム検証を行う。また、これの基本手法を越える新たな手法の発見にも挑戦する。

ソフトウェア研究開発(TakumiVision株式会社)

学校法人立命館 立命館大学が提案している5つの基本手法に対して、以下の3つのステップにて開発を行う。

第一は、C++言語にてプログラミングし、それぞれの手法の性能とそれらを組み合わせた場合の性能とを精密に評価し、最適な組み合わせアルゴリズムを見出す。

第二は、C++でのシミュレーションにて、それぞれの実行ステップ数を計測し、DSP と FPGA とによる実時間システム実現の可能性を定量的に評価する。

第三は、実装困難なアルゴリズムに対しては、実装に適した代替アルゴリズムを見出す。また、実装コストを減らす点から、実現可能なアルゴリズムにおいても、実行ステップ数削減の改良を行う。

上記ステップにて、最終的に C++言語にてアルゴリズムを実装する。また評価サンプルを用いて検証を行う。

これらの開発では、実際の音声サンプルを使った大量の性能評価が重要である。この性能評価の方式策定と評価作業は、「日本ロジックス株式会社」、および「株式会社システムアドバンス」と協調して行う。

アルゴリズム構築検証用振幅スペクトラムサンプル(日本ロジックス株式会社)

アルゴリズム構築検証用振幅スペクトラムのサンプル収集及び作成、分析を行う。

上記 の統計学習によるカテゴリー分類をおこなう為に男女、年齢等ごとに統計的に振幅スペクトラムを収集し、分析及びグルーピングを行う。

また上記 の研究開発の為に音声サンプルを使った性能評価を行う。

DSPソフトIPの研究開発(株式会社システムアドバンス)

目的:

- (1) 学校法人立命館 立命館大学によって研究開発される、TotalVariation 信号分離技術と画像の高解像度化手法を融合させたアルゴリズムである画像高解像度技術及び統計学習によるカテゴリー分類技術をもとに、従来の音声の高域復元性能を大幅に向上させるDSP信号処理技術を用いた高性能高域復元手法を研究開発する。
- (2) 「目的:1」によって研究開発した成果の信号処理アルゴリズムをDSPヘインプリメントして実時間処理によるアルゴリズムの動作検証及び性能評価を行い、その過程で生じた高速化手法による影響度の計測や改善研究を行う。

研究テーマ:

1. 高調波倍音による高域復元技術のモデル化
電話音声の 3.4KHz 以下の帯域には、ある周波数成分をもつ波動から発生した高調波が含まれている。この高調波により高次の周波数成分を予測する理論を研究し、その予測理論の高域復元技術への応用研究を行う。
2. ケプストラムによる信号分離技術のモデル化
TotalVariation 信号分離技術を音声に適応するため、ケプストラムを用いた音声信号解析を行い、音声の信号分離研究を行う。
- (3) 浮動小数点DSPを用いた信号処理アルゴリズムの検証
本研究で導き出される以下に示す(1)(2)の高域復元アルゴリズムを浮動小数点DSPヘインプリメントし、その評価・検証を行う。また、研究テーマ1.2.で得た成果を組み合わせることにより高性能な高域復元アルゴリズムの開発を目指す。浮動小数点型での高域復元性能の目標値を、Lanczos フィルタにて達成できる精度の 5dB 以上とする。
 - (1) 画像高解像度技術のアルゴリズム検証
 - (2) 統計学習によるカテゴリー分類技術のアルゴリズム検証

HDLソフトIP仕様研究開発(日本ロジックス株式会社)

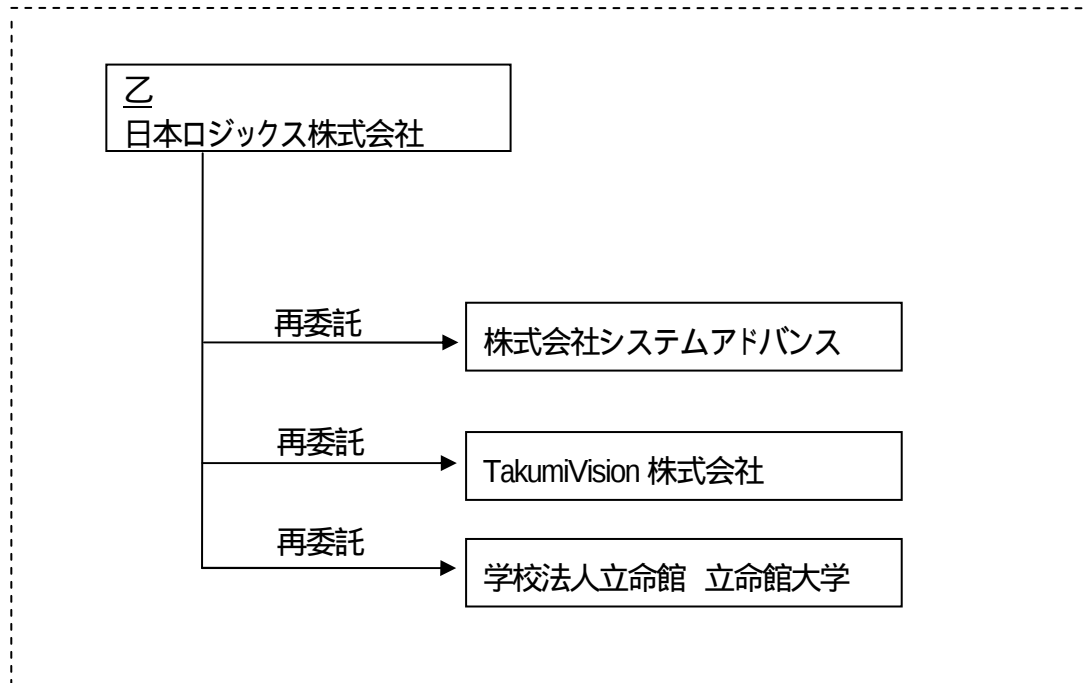
- ・ アルゴリズムHDL化検討
- ・ HDLソフトIP仕様作成
- ・ HDL用ライブラリ作成(複素演算子、DFT演算器等)
- ・ FPGAシステムボード作成

と の研究開発によって導き出されたアルゴリズムと の研究開発のDSPソフトよりHDLヘインプリメントをしWAIT時間をもたないリアルタイム処理を実現し、アルゴリズムの動作検証及び性能評価を行うため本契約期間ではアルゴリズムのHDL化検討及び、HDLソフトIPの仕様作成を行う、またHDL化の為にライブラリ作成(複素演算子、DFT演算器等)および、FPGAインプリメント後の実機評価の為にFPGAシステムボード作成を行う

1-2 研究体制

1-2-1 研究組織及び管理体制

1-2-1-1 研究組織(全体)



総括研究代表者（PL）

学校法人立命館 立命館大学
理工学部教授/MSI センター長
山内 寛紀

副総括研究代表者（SL）

日本ロジックス株式会社
代表取締役社長 千村 盛幸

1-3 成果概要

開発技術の市場競争力を高めるため、携帯電話以外の広い用途に適用できる方向に実施内容を拡大した。具体的には、(1)既存技術の限界の定量的把握、(2)新アルゴリズムの提案と実装、(3)実携帯電話環境での実証、(4)新応用分野の検討である。以下、分担の見直しを含めて、実施内容を以下の様に発展的に修正した。

1-3-1 高域復元アルゴリズムの研究

既存技術の限界を定量的に把握するため、「既存技術の性能検証」を研究項目に加えた。また、技術が多岐に渡り研究内容が膨れてきたので、「非線形信号処理での高域復元」をシステムアドバンスが担当することとした。実施内容は以下の4項目である。

- (1) 既存技術の性能検証
- (2) 非線形信号処理を発展させた新方式
- (3) 統計学習を発展させた新方式
- (4) コードブックを発展させた新方式

1-3-2 ソフトウェア研究開発

開発アルゴリズムをC++で実装し、音声データにて定量的評価することを進めた。ただし、項目 のアルゴリズム開発と平行して進めるため、アルゴリズム実装を分けて進めた。実施内容は以下の3項目である。

- (1) 非線形信号処理を発展させた新方式の実装と実験
- (2) 統計学習を発展させた新方式の実装と実験
- (3) コードブックを発展させた新方式の実装と実験

1-3-3 アルゴリズム構築検証用振幅スペクトラムサンプル

開発を進めて行く内に、実環境を模擬した検証と、実環境検証を早めに進めていくことが重要との認識に至り開発内容を発展させた。また分担の見直しも行った。実施内容は以下の3項目である。

- (1) Off-line 検証データ作成
- (2) 擬似実環境検証データベース作成
- (3) 携帯電話実環境評価の準備

1-3-4 DSPソフトIPの研究開発

用途の多様性に対応するため、実環境検証のターゲットシステムを、初期の1種類から2種類に拡大した。それに伴って分担の見直しも行った。ひとつは、DSP 搭載開発ボードであり携帯電話環境での実時間動作を目指した。もうひとつは、小型ターゲットボードあり性能検証に重点を置いた。実施内容は以下の2項目である。

- (1) DSP 搭載 DSK への実装
- (2) DSP 内蔵 ARM 搭載 BeagleBoard への実装

1-3-5 HDLソフトIP仕様研究開発

複数の候補アルゴリズムに共通する基本ライブラリ(関数分離等)と、非線形信号処理による高域復元についての HDL 化を検討している段階である。具体的な実装は次年度

に行う。

1-3-6 プロジェクトの管理・運営

- ・ 再委託契約の締結
- ・ 国と再委託先との連絡及び調整
- ・ 報告書のまとめ
- ・ 委託費の執行及び管理
- ・ 携帯電話以外の広いマーケットへの投入検討

1-4 当該研究開発の連絡窓口

所属 : 日本ロジックス株式会社
氏名 : 西尾 嘉浩
電話 : 075 - 958 - 6131
FAX : 075 - 958 - 6130
E-mail : nishio@nlogics.co.jp

第2章 本論

2-1 高域復元アルゴリズムの研究

(1) 既存技術の性能検証

既存技術の限界を検証するため、その性能検証を行った。具体的には関連する特許・文献・製品のプレスレビューを精査し、その情報をもとに PC 上でのシミュレーションを行った。既存技術では、非線形処理方式とコードブック方式の2種類に大別できる。非線形処理は、狭帯域信号から何らかの方法で高域成分を抽出し、HPF を通した後、狭帯域信号と合成することにより高域復元音声を生成する手法である。コードブック方式は、原信号から生成した量子化ベクトルをもとにコードブックを作成し、そのコードブックもとに狭帯域信号から広帯域信号を生成する手法である。

コードブック方式は、生成するコードブックの精度によって性能が依存されるため、コードブックの作成方式開発を含めた研究が重要であり、数種の既存技術のコードブック方式を検証することはあまり意味を持たない。従ってここでは、狭帯域から高調波を生成して高域復元する非線形処理方式の以下について性能検証を行った。

ヒルベルト変換方式

半波整流方式

全波整流方式

ピークデテクト方式

ゼロ置換方式

符号反転方式

ゼロ挿入 + アップサンプリング方式

クリッピング方式

これら8つのアルゴリズムによる復元音声を主観評価した結果、調波構造が原音の波形に近く、雑音/歪み感が少ない全波整流方式とヒルベルト変換方式が、非線形処理では有力なアルゴリズムであると判断した。ただし、両方式とも、原音にはない雑音が発生する問題とともに、聴認度も不足しており、改善が必要である。

(2) 非線形信号処理を発展させた新方式

(1) の既存技術の性能検証を行った結果、既存技術による高域復元には聴認度の不足と雑音の重畳という問題があった。我々は本研究開発の一部として、これらの問題を解決する非線形信号処理を発展させた新方式の開発を行った。

まず、既存技術による復元波形をスペクトル解析し、高域減衰前の原音と比較することにより、聴認度の低下と雑音の重畳の原因を分析した。

一般に原音の音声には緩やかなスペクトル包絡成分があり、その強弱の存在が自然な聴認度につながる。しかし、既存技術では原音と比べて復元音のスペクトル包絡は一樣に平坦であり(図 1 参照)、復元対象の高域におけるスペクトル包絡成分の強弱が失われてしまうため、復元音の自然さが失われるという問題があったことがわかる。

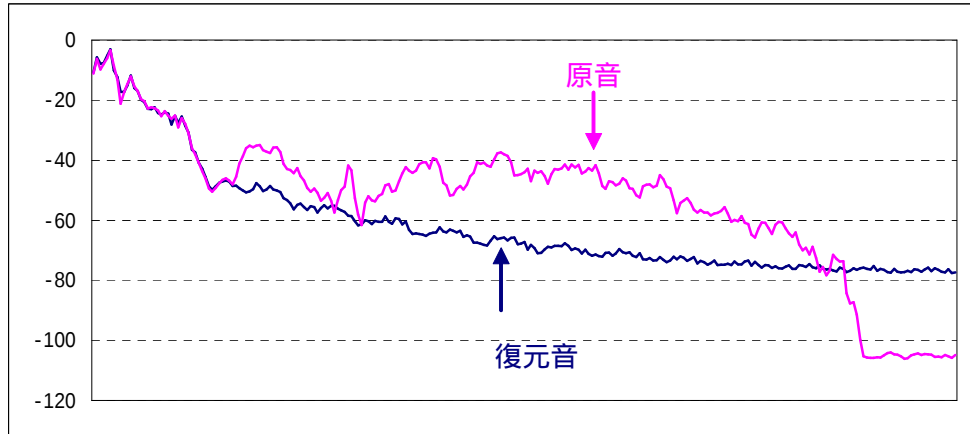


図 1 原音と復元音のスペクトル包絡波形

この問題を解決するために、新方式を用いて高域復元処理を施し合成する。これにより、従来技術では制御することが難しかった高域成分の包絡波形を制御することが容易となる。この結果、復元帯域における原音のスペクトル包絡の強弱が再現でき自然な聴認度を確保できる。

(3) 統計学習を発展させた新方式

1. 研究計画書段階で提案した3つの手法の総括

第1の提案手法である「SINC 関数の畳み込み技術の音声への適用」手法を種々検討したが、結果的には、画像の高域復元と同様、滑らかな補間の延長技術であり、高域信号の生成力が弱いことが明らかとなった。この手法よりも、第2節で述べた「非線形信号処理を発展させた新方式」の方が、高域復元性が優れているので、非線形信号処理で基本信号を発生させ、後述の技術でリファインしていくこととした。

第2の提案方法である「音声波形の分離」であるが、TotalVariation_ROF 方式と、Bi_Lateral_Filterer 方式が共に良い結果をもたらすことを実証した。特に、子音と母音の識別に有効であり、後の、CodeBook 作成において、子音と母音それぞれに最適な CodeBook を作成できるので、格段の改善効果があることが実証できた。

第3の提案方法である「統計学習によるカテゴリー分け」であるが、多方面から検討を進めた。第1段階は、LPC 分析後の PARCOR 係数による最小2乗フィルタによる一律な高域復元。第2段階は、同じく PARCOR 係数の主成分分析と最小2乗フィルタを組み合わせた一律な高域復元。第3段階は、第2段階の主成分分析後のベクトルを、フレーム単位でベクトル量子化して CodeBook を作成する高域復元。第4段階は、第3段階を発展させて、男女別それぞれにベクトル量子化して CodeBook を作成する高域復元である。より段階を得るごとに、復元精度が向上することを確認した。

2. LPC 分析後の PARCOR 係数による最小2乗フィルタによる一律な高域復元

1) 学習過程

提案する処理フローは、学習過程と復元過程とからなる。学習過程では、多くの学習サンプルを用いて、低域から高域への最適な変換行列を求めておく。復元過程では、その変換行列を用いて、入力音声から高域復元信号を生成する。

高域のない信号(ここでは 3.4KHz カット信号を用いる)を、高域信号(ここではテクスチャ信号と呼ぶ)と、低域信号(ここでは骨格信号と呼ぶ)とに分割する。基本的に、テクスチャ信号のみに高域復元処理を行う方式を用いる。すなわち、骨格信号は、高域復元されたテクスチャ信号にそのまま加算するので、学習もテクスチャ信号のみに行う。分離したテクスチャ信号に非線形処理を行って、高域信号を生成し(第1次高域復元信号と呼ぶ)、その信号に LPC 分析を行って PARCOR 係数を求める($\hat{a}_1 + \hat{a}_2$)。これを、復元 PARCOR 係数と呼ぶ。

一方、高域のある信号(ここでは原音信号を用いる)についても同様に、テクスチャ信号と骨格信号に分離する。そして、そのテクスチャ信号に LPC 分析を行って PARCOR 係数を求める($\hat{a}'_1 + \hat{a}'_2$)。これを、原音 PARCOR 係数と呼ぶ。

次に、多くのサンプルを集めて、復元 PARCOR 係数から原音 PARCOR 係数を生成する変換行列を最小2乗法にて求める($\hat{a}_1 + \hat{a}_2$)。すなわち、「復元 PARCOR 係数ベクトル」を線形変換して生成したベクトルと、「原音 PARCOR 係数ベクトル」との2乗距離が最小となる線形変換(変換行列)を求めることになる。

2) 復元過程

学習過程で求めた変換行列を用いた、復元過程は、学習過程と同様に、高域復元したい(ここでは 3.4KHz カット信号)からテクスチャ信号を求め、そのテクスチャ信号に非線形処理を行って第1次高域復元信号を生成し、その信号に LPC 分析を行って PARCOR 係数を求め、復元 PARCOR 係数と復元残差信号を得る($\hat{a}_1 + \hat{a}_2$)。

次に、復元 PARCOR 係数に学習過程で求めた変換行列をかけて「第2次の高域復元信号の PARCOR 係数」を求める。この係数と復元残差信号を用いて、「第2次の高域復元信号のテクスチャ成分」を生成する($\hat{a}_1 + \hat{a}_2$)。さらに、この復元高域信号のテクスチャ成分に骨格信号を加えて、最終的な高域復元音声を生成する($\hat{a}_1 + \hat{a}_2$)。

3) 評価

本手法の高域復元では、2つの選択が必要である。第1は、「第1次高域復元信号」生成手法である。種々の手法を検討した結果、Hilbert 変換方式と、全波整流方式が適切であることが分かった。

第2は、第1次高域復元信号から第2次高域復元信号を生成する「変換行列」を、どのような特徴ベクトルで作成するかである。Cepstrum と LPC 分析とを比較した結果、フレーム長は 512、テクスチャ信号生成は TotalVariation-ROF モ

デル、第1次高域復元信号生成は Hilbert 変換方式、テクスチャ信号の特長ベクトルは 20 次の PARCOR 係数と策定した。この結果を基本として、詳細を TakumiVision 社にて結めることとした。

(4) コードブックを発展させた新方式

2-1(3)2 節の「LPC 分析後の PARCOR 係数による最小2乗フィルタによる一律な高域復元」に、主成分分析を導入した方式をベースにして検討した(章(3)節の統計学習を発展させた新方式の実装と実験)。すなわち、PARCOR 係数空間に、学習で求めた KL 変換行列を操作した新たなベクトル空間にて、高域推定する変換行列を最小2乗法にて求めている。

この方式で不満足な点は2点であるが、これらの点を改良したアルゴリズムを提案し、そのプロトタイププログラムを作成した。

現在、評価を行いながら条件の最適化を行っている段階である。基本的には、CodeBook を持つことによる、単一の変換行列の場合より、確実に改良されている。ただし、有声音の多様なピッチに対するスケーリングでの対応が、まだ最適化されていない。継続して改良している。

2-2 ソフトウェア研究開発

(1) 非線形信号処理を発展させた新方式の実装と実験

前項 2-1(2) で研究を行った非線形信号処理を発展させた新方式のアルゴリズムを PC 上にシミュレーション環境(表 1を参照)を構築した。

表 1 シミュレーション環境

OS	WindowsXP Professional
開発言語	Visual C++
Bit 精度	最大倍精度 float
帯域制限入力音声	300Hz ~ 3.4kHz
復元音声帯域幅	3.4kHz ~ 7.8kHz

PC 上に構築したシミュレーション環境の性能評価をするために、高域復元音の周波数成分の検証と、主観評価を行った。

(2) 統計学習を発展させた新方式の実装と実験

1. 前検討

高域復元方式として2つの手法を検討した。第1は、第1次の高域復元信号を生成する手法である。立命館大学にて、「全波整流と Hilbert 変換」を検討し、TakumiVision 社にて、「Lanczos フィルタと Gabor フィルタ」を検討した。性能比較をした結果、歪の点では、「Lanczos フィルタ方式が優れているが、数倍の高調波を生成できる点で、全波整流および Hilbert 変換のような非線形変換を行う方式の方が優れていることが判明した。この結果、将来の可能性の点から、非線形変換を行う方式を採用することとした。

次に、音声信号を、テクスチャ信号と骨格信号とに分離する方式を定量比較した。1 つは Bilateral フィルタであり、もうひとつは TotalVariation-ROF モデルである。これらのプログラムは、TakumiVision にて作成した。この結果、パラメータによって大きく左右されるが、パラメータを選別さえすれば、いずれの場合も音声波形分離に耐えうる性能であることが分かった。ただし、処理時間比較を行うと、Bilateral フィルタの方が有利である。

2. LPC 分析と主成分分析を用いた高域復元

立命館大学にて検討した高域復元方式(高域復元アルゴリズムの研究、(2)統計学習を発展させた新方式)を発展させ、上記方式に主成分分析を導入した。同じように、学習過程と復元過程とからなる。学習過程では、多くの学習サンプルをも用いて、低域から高域への最適な変換行列を求めておく。復元過程では、その変換行列を用いて、入力音声から高域復元信号を生成する。

1) 学習過程

テクスチャ信号の PARCOR 係数(20 次)を求めるまでは同じである。これを、高域のない信号(ここでは 3.4KHz カット信号を用いる)の第 1 次高域復元信号について処理する。一方、高域のある信号(ここでは原音信号を用いる)にいて処理する。

この 20 次の PARCOR 係数について主成分分析を行う。フレーム長 512 として、男女の音声から 2,000 フレームを学習データとして選別して統計解析を行い、KL 変換行列を作成した。次に、高域のない信号の KL 変換行列処理後のベクトルと、高域のある信号の KL 変換行列処理後のベクトルの間で、最小 2 乗法処理を行って変換行列を求めた。

2) 復元過程

立命館大学にて検討した高域復元方式に対して、学習過程で求めた、KL 変換行列操作が加わっているのが特長である。この場合に変換行列は、KL 変換された PARCOR 係数間の変換行列であり、1 次高域復元信号から 2 次高域復元信号を生成する。

3) 評価

従来方式に比べて、原音に近いスペクトラムが得られていることが分かった。耳で聞いた主観評価においても、明らかに優れていることを確認できた。

(3) コードブックを発展させた新方式の実装と実験

立命館大学で開発した「統計学習を発展させた新方式」の改良を進めた。まず、主成分分析で重要な固有値問題の高速化を進めた。基本プログラムは、従来から大学で使用していた「特異値分解ライブラリ」を使用していたが、次元数がそれほど大きくないことから「ヤコビ法」に変更した。また、プログラムのモジュール化を進めた。ARM および DSP に実装する上で、モジュールごとの処理時間解析が行い易いようにすることが目的である。

次に、立命館大学でプロトタイプ開発として行った学習データを改良した。立命館大学では、学習データ数が 2,000 であったが、今回、学習データを 10,000 に増やして、より精密な変換行列を作成した。性能確認を行った結果、CodeBook 作成に伴う改善が、

主観評価ながら確認できた。しかしながら、ピッチによるクラス分類とスケーリングの効果は、それほど得られていないので、もう一段階の改良を継続して行う予定である。

2-3 アルゴリズム構築検証用振幅スペクトラムサンプル

(1) Off-line 検証データ作成

音声高域復元アルゴリズムの性能を評価する際に、理想環境での性能評価用音声データが必要と判断し、次に示す Off-line 検証データを作成した。

- ア) 男性__会話1
- イ) 男性__会話2
- ウ) 女性__会話1
- エ) 女性__会話2

この Off-line 検証データは、本研究の為に購入した音声データベースから、周波数帯域を300Hz から3400Hz にカットして作成した音声データを、携帯電話網通過前(送話)に (3) で後述する集音装置を用いて録音したデータである。

(2) 擬似実環境検証データ作成

音声高域復元アルゴリズムの性能を評価する際に、前述した (1) の方法で性能に問題が無いと判断できた場合は、次の段階としてノイズ成分が含まれた性能評価用音声データが必要と判断し、携帯電話網を通した次に示す2種類の擬似実環境検証データを作成した。

- (1) で作成した Off-line 検証データと、環境ノイズを合成したデータ
実環境で録音したデータ

(3) 携帯電話実環境評価の準備

前述した、2-3(2) で用いた集音装置について、携帯電話の公衆網通過前(送話)と公衆網通過後(受話)を IC レコーダに録音する装置を構築した。この集音装置を用いて録音した送話データと受話データを比較することで性能評価を行う。

2-4 DSPソフトIPの研究開発

(1) DSP 搭載 DSK への実装

DSP 搭載 DSK ボードは、動作クロック 225MHz の DSP を搭載するアプリケーション開発ボードである。この DSK ボードを用いて、携帯電話での実時間動作検証が行える環境構築を行った。開発工程を以下の) ~) に示す。

- i) 実時間動作のための仕様検討
- ii) インターフェースドライバソフトおよびライブラリソフトの組み込み
- iii) 実験1:新方式アルゴリズムの DSP ボードへの実装と検証
- iv) 実験2:携帯電話環境への適用

(2) DSP 内蔵 ARM 搭載 BeagleBoard への実装

計画書段階では、実時間実装は DSP 評価ボードのみを考えていた(株式会社システムアドバンス担当)が、最近、非常にコンパクトなターゲットボードが販売されるようになって来た。その代表が、米国 TI 社(Texas Instrument)が製品化した、Beagle Board である。安価で小型な利点を有する一方、処理能力では DSP 評価ボードに劣る。これらの点を考慮して、実装方式の比較検討ができるように、BeagleBoard への実装を進めた。

- 1) 開発環境の構築

2) ターゲットアルゴリズムの実装

2-5 HDLソフトIP仕様研究開発

(1) アルゴリズム HDL 化検討

2-1 で研究開発された、複数の候補アルゴリズムに共通する基本ライブラリ(関数分離等)と、非線形信号処理による音声高域復元についての HDL 化を検討した。その中から、今回の研究開発においては、基本ライブラリとして作成が必要か判断した。具体的な音声高域復元機能の HDL 実装は、次年度に行う。

(2) HDL ソフト IP 仕様作成

(1) で検討した内容を基に、基本ライブラリの HDL ソフト IP 仕様を作成した。

なお、仕様については以下の 図 2 仕様書例(S2P) に示す様式で、次の内容について記載している。

基本ライブラリの概要

入出力信号の名称

入出力信号の入出力方向

入出力信号のビット幅

入出力信号の極性

入出力信号の説明

動作概要(タイミングチャート)



図 2 仕様書例(S2P)

(3) HDL 用ライブラリ作成

(2) で作成した仕様を基に、基本ライブラリを HDL 言語で作成した。また、作成した基本ライブラリは、コンピュータ上で、シミュレーションを実施し、仕様通り信号処理動作をす

る事の確認を実施した。更に、(4) で作成した FPGA 評価デモボードに実装して仕様通り信号処理動作をする事の確認も実施した。なお、シミュレーションについては、以下の図 3 シミュレーション動作波形 に示す様な動作波形について確認し仕様書に記載されている動作概要通りである事を確認した。

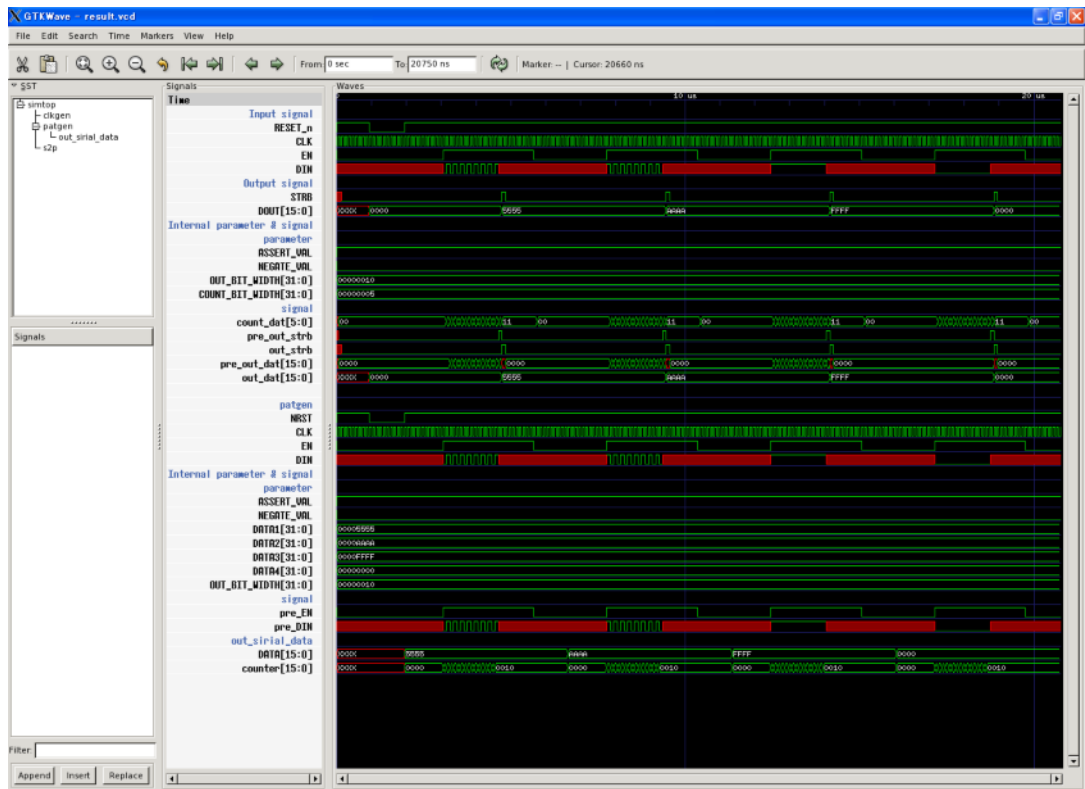


図 3 シミュレーション動作波形

(4) FPGA システムボード作成

当初は、DSP チップと FPGA チップが混載したカスタムボードを開発する計画だった。しかし、プロジェクト期間が、開発開始日程が遅延した事により、約1ヶ月短縮した事で、ボード仕様検討、設計、試作と試作ボード評価の期間が確保できないと判断した。そのことにより代替ボードの検討を実施し、以下の

表 2 代替 FPGA・DSP システムボード に示すシステムボードを購入する判断を行った。今回の成果としては、DSP システムボードに、2-4(1) で研究開発され実装確認されたソフトウェアを同様に実装して同様に動作する事を確認した。また、2種類の FPGA 評価デモボードについても、(3) で作成した HDL 用ライブラリを実装して、仕様通り動作することの確認を完了している。

表 2 代替 FPGA・DSP システムボード

ボード名	メーカー / 型番	説明
DSP システムボード	Texas Instruments /TMDSDSK6713	2-4 で開発されたソフトウェアを実装することにより音声高域復元を実現し、お客様に対してデモや営業を実施する装置
FPGA 評価デモボード (Xilinx 版)	Xilinx /HW-V5-ML501-UNI-G	2-4 で開発されたソフトウェアを HDL 言語に置き換え、お客様が、Xilinx 製 FPGA チップに実装することを要望された時に音声高域復元を実現しデモや営業を実施する装置
FPGA 評価デモボード (ALTERA 版)	ALTERA /DK-N2EVAL-3C25N	上記 Xilinx 版ボードで実装した HDL 言語をお客様が、ALTERA 製 FPGA チップに実装することを要望された時に音声高域復元を実現しデモや営業を実施する装置

2-6 プロジェクトの管理・運営

- ・再委託契約の締結
- ・国と再委託先との連絡及び調整
- ・報告書のまとめ
- ・委託費の執行及び管理
- ・携帯電話以外の広いマーケットへの投入検討

開発技術の市場性を見直し、実施内容を拡大発展させながら進めてきた。その結果、計画段階での目標である「携帯電話への投入」という目的に対しては、ほぼ所望の成果を得、なお継続して改良できる余地を多く残している。この改良余地と隣接技術との融合を進めることで、携帯電話を含めた広いマーケットへの投入が可能になると判断している。その内容は以下の4点である。

- ◆ 極狭帯域信号(1KHz カット信号)の高域復元による、骨伝動通信の高品質化
- ◆ 高域強調技術との融合による、老人に優しい携帯電話
- ◆ 子音 / 母音分離と発話コードブック変換による聾啞者の発話サポート
- ◆ 音源分離と高域復元の融合によるテレビ会議音質の改善

これらの開発と実環境検証を次年度行い、マーケットに投入する。

3-1 本プロジェクトの総括

高域復元アルゴリズムの研究においては、既存技術の限界を検証するため、その性能検証を行い、非線形信号処理を発展させた新方式の研究開発、統計学習を発展させた新方式の研究開発、コードブックを発展させた新方式の研究開発を行った。次年度の研究開発に向けては、評価を行いながら条件の最適化を行っていく必要がある。

ソフトウェア研究開発については、PC 上にシミュレーション環境を構築し、高域復元アルゴリズムの研究で開発されたアルゴリズムを PC 上でシミュレーションが行え、評価が行う事ができた。これにより DSP ソフト IP の研究開発や、HDL ソフト IP の研究開発において、アルゴリズムに対するフィードバックが発生しても評価することが可能となり、次年度の研究開発に向けて研究開発環境が構築できた。

アルゴリズム構築検証用振幅スペクトラムサンプルについては、ソフトウェア研究開発や DSP ソフト IP の研究開発、HDL ソフト IP の研究開発において評価する際に用いる検証用データについて作成することができた。これにより評価する際に電話回線網にアクセスすることなく音声データによる評価が行える。また、高域復元された音声データについても原音データとの比較ができ正確な評価が行える。

DSP ソフト IP の研究開発については、実時間動作による評価が行える環境を構築した。これによりソフトウェア研究開発で評価を行ったアルゴリズムをシミュレーションのみではなく実時間動作による評価が可能となり、次年度の研究開発に向けてのソフトウェアは研究開発環境が構築できた。

HDL ソフト IP 仕様研究開発については、ハードウェアによる評価が行える環境を構築した。これによりソフトウェア研究開発や DSP ソフト IP の研究開発によって導き出されたアルゴリズムをハードウェアによる実時間動作の評価が行える環境を構築した。これにより、実製品に近い環境で評価が行えるようになり、次年度の研究開発に向けて事業化に向けた研究開発環境が構築できた。これらの成果により次年度の研究開発の基礎が構築でき、更には事業化に向けた市場に対する課題抽出が行えたことから本年度の研究開発は成功したと総括する。

参考文献

- [1] 中田和男, "改定 音声", 第4 - 5章, コロナ社 日本音響学会編(1998)
- [2] 三浦種敏ほか, "新版 聴覚と音声" 第3部, 電子情報通信学会(1980)
- [3] 大賀寿朗ほか, "音響システムとデジタル処理", 第1章, 電子情報通信学会(1995)
- [4] 古井貞熙: "デジタル音声処理", 第5章, 東海大学出版会
- [5] 守屋健弘: "音声符号化", 第2章, 電子情報通信学会(1998)
- [6] 佐藤寧: 「圧縮で失われた音の復元技術」, "電気学会誌 127 巻 7 号", 電気学会(2007)