

平成 22 年度

学士学位論文

聴覚特性に基づくマスキングを用いた小型
スピーカによる音声強調

Speech enhancement by loudspeakers using auditory
property based masking

1110222 安藝 遼

指導教員 福本 昌弘

2011 年 3 月 1 日

高知工科大学 情報システム工学科

要 旨

聴覚特性に基づくマスキングを用いた小型スピーカによる音声 強調

安藝 遼

携帯電話で音楽を聞く場合、音が小さいため聞こえにくい、またはテレビやミュージックプレイヤーで聞く音声と比べて音質が異なって聞こえることがある。これらの原因として、携帯電話はスピーカのサイズが非常に小さいため定格入力の限界が低いことが挙げられる。したがって、携帯電話の音を大きくするためにアンプの出力を上げることができない。また、携帯電話で音楽を再生した場合、音が鳴りにくい周波数帯域があるため音質が異なることが挙げられる。そこで、従来再生される大きさより大きく聞こえるように対象の音声を聞きやすくする音声強調方法がある。しかし、この方法では携帯電話のアプリケーションによって有効でない場合がある。

本研究では、小型スピーカに適した音声強調を行い、対象のアプリケーションの音が従来再生される大きさより大きく聞こえるようにする。提案手法では、携帯のスピーカ特性から目的のアプリケーションごとに元の信号に対して、音がより大きく聞こえるような処理をする。

設計したデジタルフィルタが正常に処理できていることを計算機シミュレーションによって確認している。また、処理した音声を被験者実験により評価し、従来再生される音より大きく聞こえることが確認できたためシステムが有効であることを確認している。

キーワード 小型スピーカ, デジタルフィルタ, 聴覚特性, 音声強調

Abstract

Speech enhancement by loudspeakers using auditory property based masking

Ryo AKI

When listener listens music on a mobile phone, it is difficult to listened the sound. Sound quality on a mobile phone is worse than televisions and music players. These problems are attributed to major two factors. One is the lowness of the limit of rated input due to the smallness speakers, the output an electric signal to amplifiers can't be raised in order to turn up the volume. The other is low frequency reproduction capability compared with televisions and music players. The methodology with emphasis a sound of target to listened the loud had been proposed. However, this method is not effective in some cases of the application on mobile phones.

In this paper, the method of speech enhancement suitable for loudspeakers to listened a loud sound have proposed. In the proposed method, source signal is processed depend on each target applications to listened loud exploiting a characteristic of mobile phone speaker.

The process of designed digital filter verified from the computer simulation and the processed source signal's sound is the confirmed by subject rating. In addition, subject rating result shows the high effectiveness of proposed method for loudspeakers.

key words loudspeakers, digital filter, speech enhancement, auditory property

目次

第 1 章	序論	1
1.1	研究の目的とその背景	1
1.2	本論文の構成	2
第 2 章	小型スピーカの問題点	3
2.1	はじめに	3
2.2	白色信号による小型スピーカとモニタースピーカの比較	4
2.3	スイープ信号による小型スピーカとモニタースピーカの比較	5
2.4	まとめ	7
第 3 章	聴覚特性に基づく音声強調システム	8
3.1	はじめに	8
3.2	システムのアルゴリズム	8
3.3	システム構成	9
3.3.1	1/3 オクターブ分析	10
3.3.2	バンドパスフィルタ	11
3.3.3	聴覚特性 (A 特性)	13
3.3.4	音声強調フィルタ	14
3.4	システムの検証	17
	FFT → IFFT の検証方法	17
	FFT → IFFT の検証結果	18
3.4.1	FFT → 処理 → IFFT の検証方法	18
	FFT → 処理 → IFFT の検証結果	19
3.5	有効性の評価	20

3.5.1	計算機シミュレーションによる評価	20
3.5.2	音声による評価	23
	バンドパスフィルタによる音質の変化の検証	24
	バンドパスフィルタによる音質の変化の検証結果と考察	25
	アルゴリズムの検証	26
	アルゴリズムの検証結果と考察	26
	音声強調フィルタによる音声の検証	27
	音声強調フィルタによる検証結果と考察	27
3.6	まとめ	28
第 4 章	結論	30
4.1	本研究のまとめ	30
4.2	今後の課題	30
	謝辞	32
	参考文献	34
付録 A	中心周波数と上下限遮断周波数	35
付録 B	フーリエ変換	36
B.1	フーリエ変換の特徴	36
B.2	フーリエ変換の手順	36

図目次

2.1	白色信号による小型スピーカの周波数特性	4
2.2	白色信号によるモニタースピーカの周波数特性	5
2.3	スイープ信号による小型スピーカの周波数特性	6
2.4	スイープ信号によるモニタースピーカの周波数特性	6
3.1	アルゴリズム	9
3.2	提案するシステム構成図	10
3.3	最小通過周波数 ω_1 の決定方法	11
3.4	最小通過周波数 ω_2 の決定方法	12
3.5	バンドパスフィルタの通過域の決定方法	12
3.6	A 特性	14
3.7	音声強調フィルタの設計方法	15
3.8	A 特性を基に音声を強調するフィルタ	16
3.9	A 特性と耳の構造を基に音声を強調するフィルタ	16
3.10	A 特性を基に音声をフラットにするフィルタ	17
3.11	FFT → IFFT 前の信号	18
3.12	FFT → IFFT 後の信号	18
3.13	制御前の振幅 1 の正弦波 (440Hz)	19
3.14	制御後の振幅 2 の正弦波 (440Hz)	19
3.15	制御前の振幅 1 の正弦波 (440Hz)	19
3.16	制御後の振幅 2 の正弦波 (440Hz)	19
3.17	入力した白色信号	21
3.18	A 特性を基に強調した白色信号	22
3.19	A 特性と耳の構造を基に強調した白色信号	22

3.20 A 特性を基に減衰させた白色信号	23
---------------------------------	----

表目次

2.1	録音環境	3
3.1	設計したバンドパスフィルタの最小・最大通過周波数	13
3.2	FFT 環境	18
3.3	計算機シミュレーション環境	20
3.4	被験者実験の環境	24
3.5	バンドパスフィルタによる音質の変化の検証結果	25
3.6	アルゴリズムの検証結果	26
3.7	音声強調フィルタによる検証結果	28
A.1	中心周波数と上下限遮断周波数	35

第 1 章

序論

1.1 研究の目的とその背景

近年，携帯電話の普及率の増加により，携帯電話の用途は従来の音声通話以外にもミュージックプレイヤー，ワンセグ視聴，時計・アラームなど様々なエンターテインメントとしてのアプリケーションが要求され，変化している [1]．そのため携帯電話で使用されているスピーカには，音響特性の向上，低消費電力化，小型・薄型化などが求められ，従来使用していた周波数帯域の狭い圧電スピーカから周波数帯域の比較的広いダイナミックスピーカが現在多くの場合，使用されている [2][3]．しかしながら，同じ音声をテレビやプレイヤーなどのスピーカで再生したものと聞き比べると，携帯電話で再生した音声と音質が異なって聞こえる．また，携帯電話で音楽を再生すると音量が小さいため，外方向のスピーカに近寄らなければ聞こえにくかったり，外部の音にマスクされ聞こえないことがある．テレビやプレイヤーなどに内蔵されているスピーカの場合，スピーカのサイズが大きいため定格入力限界も高く，音が小さければ電気信号を増幅させるアンプの出力を上げることで，電力のエネルギーを増やし音を大きくすることができる [4]．しかし，携帯電話のスピーカの場合，スピーカを収納するスペースに限りがあるため，スピーカのサイズが $\phi 10\text{mm} \sim 20\text{mm}$ 程度と非常に小さい．そのため，定格入力限界が低く，大きい音を出せない．携帯電話で音量を上げようとアンプの出力を上げると定格入力の範囲を超え，音声波形が変形してしまい弾性歪みが生じてしまう．弾性歪みが生じた音声を再生すると音割れといった音質劣化が起こり，受聴者に不快感や違和感を与えてしまう．そのため，アンプの出力を上げ，電力のエネルギーを増やす方法で音を大きくすることはできない．そこで，目的の音声の周波数成分を抽出

し強調することで、強調した音声聞き取りやすくなり大きく聞こえる音声強調という処理方法が提案されている。音声通話の場合、周りの雑音があっても人の声だけを強調することで、聞き取りやすくなることが確認されている [5]。しかしこの場合、人の声だけを強調するため、ミュージックプレイヤーやワンセグ視聴のように楽器音や機械音など様々な音声が含まれている音声には有効ではない場合がある。なぜなら、受聴者や音声の種類によっては人の声より楽器音や機械音などの背景音を重視している場合があるからである。そこで本研究では小型スピーカに適した処理方法を提案し、携帯電話でよく使用されるアプリケーションを対象とした音声強調を行い、聴覚特性を基に音が大きく聞こえるシステムを提案する。

1.2 本論文の構成

本論文の構成を示す。

第2章では、携帯電話のスピーカの問題点をモニタースピーカの周波数特性と比較し説明する。

第3章では、提案する聴覚特性に基づく音声強調システムについて説明し、計算機シミュレーションの評価と被験者実験による客観的評価の結果について述べる。

そして最後に、これまでの実験結果から第4章で本論文のまとめと今後の課題について述べる。

第 2 章

小型スピーカの問題点

2.1 はじめに

携帯電話に内蔵されている小型スピーカは、テレビやプレイヤなどに内蔵されているスピーカに比べ、サイズが非常に小さいため、定格入力に限界が低く大きい音を出せないことがある。また、テレビやプレイヤなどで再生した音声と携帯電話で再生した音声を聞き比べると音質が異なって聞こえる。そこで本章では、本研究で扱う小型スピーカの周波数特性をモニタースピーカの周波数特性と比較し、小型スピーカの問題点を説明する。モニタースピーカとは、一般的なオーディオスピーカと違い、広範囲な周波数帯域の再生のバランスが良く、フラットな音声を再生することができるスピーカである [6]。測定に使用した録音環境を表 2.1 に示す。小型スピーカは、実際に携帯電話で使用されているものを使用している。入力信号には、白色信号とスイープ信号の 2 種類を用いて周波数特性を測定する、

表 2.1 録音環境

小型スピーカ	N901is(NEC 製) 内蔵スピーカ
モニタースピーカ	YAMAHA モニタースピーカ MSP5A
マイク	BEHRINGER ECM8000
録音ソフト	DSSF3(吉正電子製)
サンプリング周波数	48000Hz
量子ビット	16bit

2.2 白色信号による小型スピーカとモニタースピーカの比較

入力信号を白色信号とした場合の小型スピーカの周波数特性を測定した結果を図 2.1 に示し、モニタースピーカの周波数特性を測定した結果を図 2.2 に示す。

図 2.2 からモニタースピーカは、ほぼ全ての周波数帯域がバランスよく出力され、音声フラットな理想的なスピーカであるということが確認できる。図 2.1 から小型スピーカは、周波数帯域ごとの出力のバランスが悪く、約 1000Hz 以下の周波数帯域は鳴りにくいということが確認できる。そのため小型スピーカで再生した音声は、高い周波数帯域が低い周波数帯域をマスクしたり、鳴ってない周波数帯域があるため、テレビやプレイヤなどで再生された音声と音質が異なって聞こえることが確認できる。

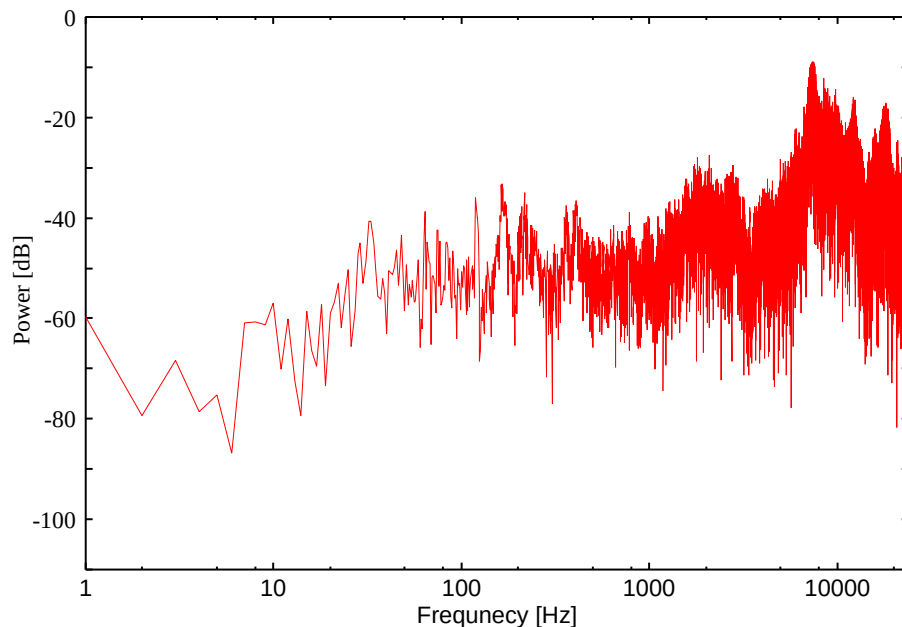


図 2.1 白色信号による小型スピーカの周波数特性

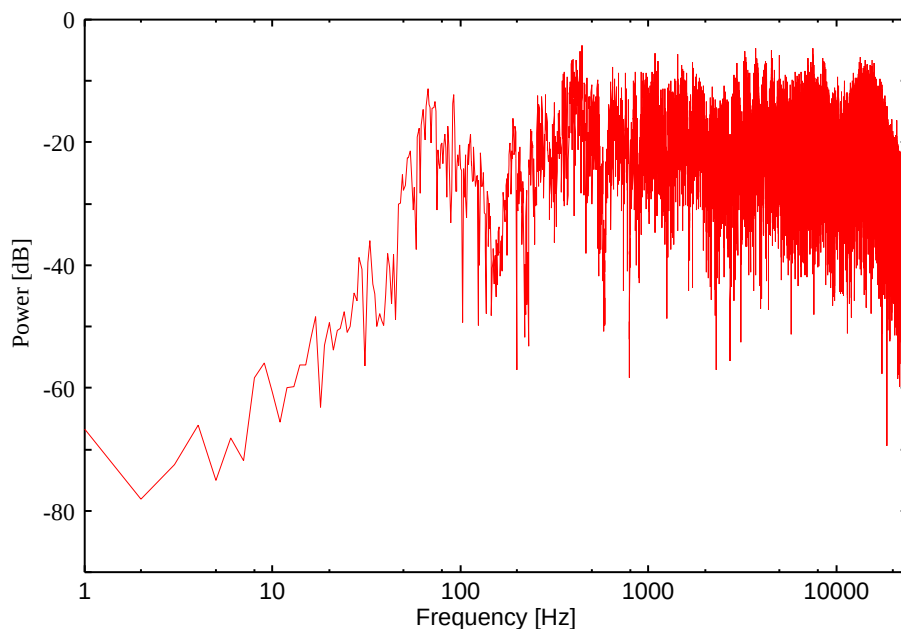


図 2.2 白色信号によるモニタースピーカの周波数特性

2.3 スイープ信号による小型スピーカとモニタースピーカの比較

次に白色信号を使用して、周波数特性を測定する場合、スピーカから出ている音か外部の騒音なのかを録音時に判断することができないため、目と耳で確認できるスイープ信号を使用する。スイープ信号とは、徐々に周波数が変化する信号で、輝度・クロマ信号の周波数帯域の測定やフィルタ特性、エイリアシングノイズのチェックを行う場合に使用する信号である。スイープ信号を使用することで、どの周波数から音として鳴り出すのか目と耳で判断できる。スイープ信号を用いて小型スピーカの周波数特性を測定したものを図 2.3 に示し、モニタースピーカの周波数特性を測定したものを図 2.4 に示す。

図 2.3 から小型スピーカは、約 550Hz から音が聞こえだし、他の周波数帯域の再生に強弱があることが確認できた。図 2.4 からモニタースピーカは、約 50Hz から音が聞こえだし、他の周波数帯域の再生に強弱がほとんどないことが確認できた。

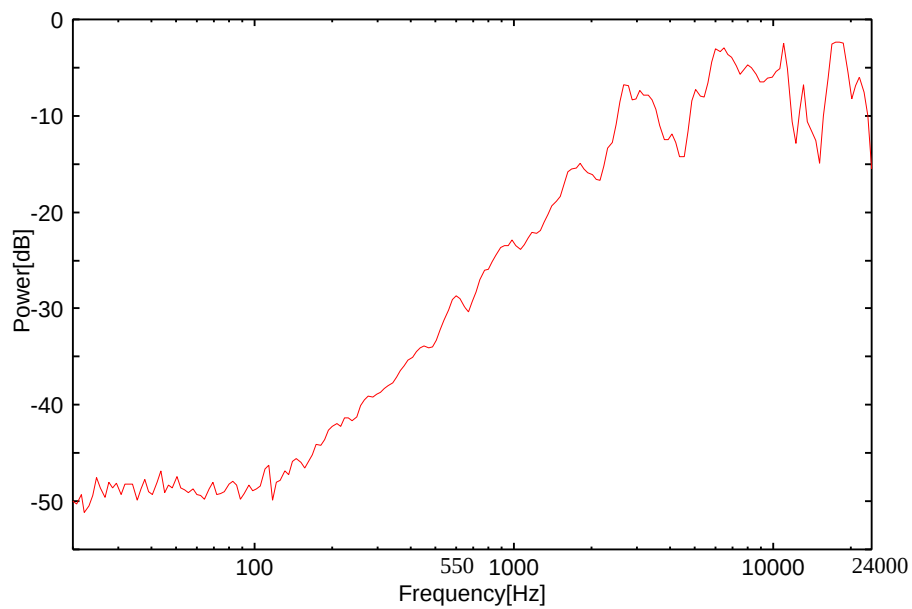


図 2.3 スイープ信号による小型スピーカの周波数特性

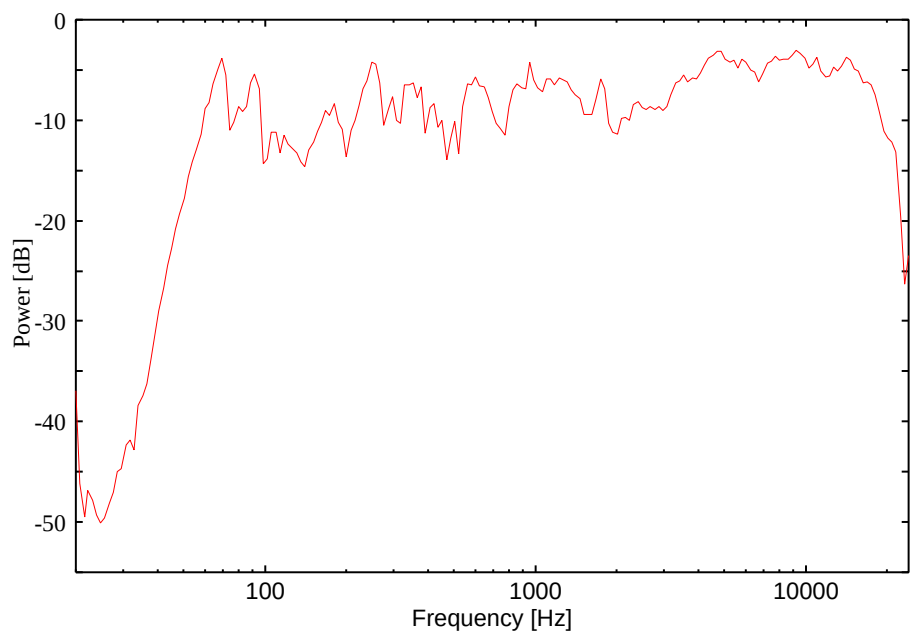


図 2.4 スイープ信号によるモニタースピーカの周波数特性

2.4 まとめ

本章では，白色信号とスイープ信号の 2 種類の信号を用いて小型スピーカとモニタースピーカの周波数特性の比較を行った．その結果，小型スピーカの問題点として低い周波数帯域に比べ，高い周波数帯域は鳴りやすいということが検証できた．また，周波数帯域によって音声を再生しにくい周波数帯域があることを聴覚で確認できた．そのため小型スピーカで音声を再生した場合，テレビやプレイヤで再生した音声と音質が異なることを検証できた．

次章では，本章で検証した小型スピーカの問題点に考慮し，提案する小型スピーカに適したシステムについて説明する．

第 3 章

聴覚特性に基づく音声強調システム

3.1 はじめに

本章では，提案する小型スピーカに適した聴覚特性に基づく音声強調システムについて説明する．まず，提案する小型スピーカに適したシステムのアルゴリズムとシステム構成について説明する．次にシステムのアルゴリズムに基づく処理方法について説明し，処理に用いる高速フーリエ変換と高速フーリエ逆変換の実験，処理方法について説明する．最後に，計算機シミュレーションによる評価と被験者実験による客観的評価を行い，提案するシステムの有効性を評価する．計算機シミュレーションによる評価では，入力信号の対象を白色信号とし，被験者実験では，入力信号を対象とするアプリケーションごとの音声として有効性の評価を行う．また本研究では，対象とするアプリケーションの音声をミュージックプレイヤー，音声通話（ハンズフリー），ワンセグ視聴の 3 種類とする．

3.2 システムのアルゴリズム

提案する小型スピーカに適した音を大きくするアルゴリズムについて説明する．携帯電話の小型スピーカの周波数特性からも分かるように携帯電話で音声を再生すると鳴りにくい，または鳴っていても音量が小さすぎて聞き取れない無駄な周波数帯域が低い周波数帯域と高い周波数帯域にある．その無駄な周波数帯域を削ることで削った周波数帯域の電力値を聞き取れる周波数帯域に使用でき，聞き取れる周波数帯域の電力値を増やせるため，電力のエネルギー量を一定のまま音を大きくすることができる．鳴りにくい，または鳴っていても聞き

取れない無駄な周波数帯域を削るため、音質に変化はない。提案する小型スピーカに適した処理方法のアルゴリズムを図 3.1 に示す。

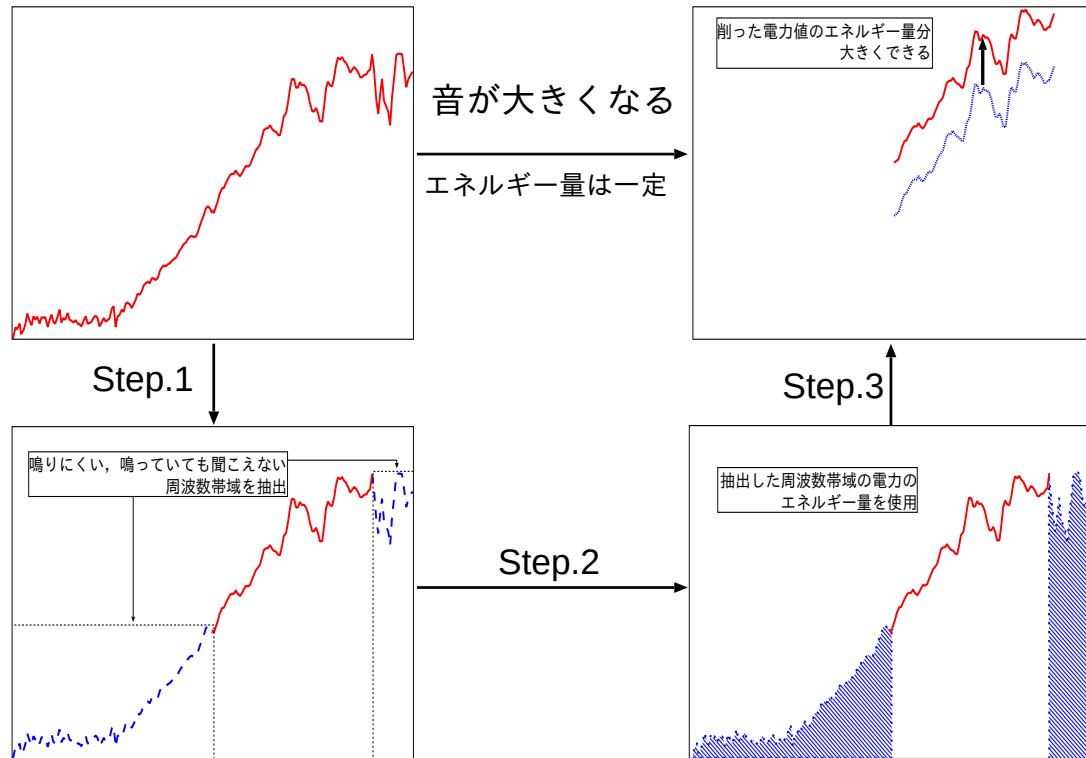


図 3.1 アルゴリズム

このアルゴリズムを基に小型スピーカに適したシステムを構成し、音を大きくする。

3.3 システム構成

提案するシステムは図 3.2 に示すように、デジタルフィルタの設計方法と高速フーリエ変換 (FFT)、高速フーリエ逆変換 (IFFT)、スピーカ、マイクで構成されている。まず、本システムでは音声を入力する前に、提案するシステムのアルゴリズムに基づく小型スピーカに適したデジタルフィルタを設計するため、点線の四角で囲まれた動作を行い、対象とするアプリケーションごとの音声強調フィルタを設計する。次に、入力信号を FFT で周波数領域に変換したものと設計した音声強調フィルタを畳み込み、周波数帯域ごとの音圧の処理を行う。最後に処理を行った信号を IFFT で時間領域信号に変換し、評価する。

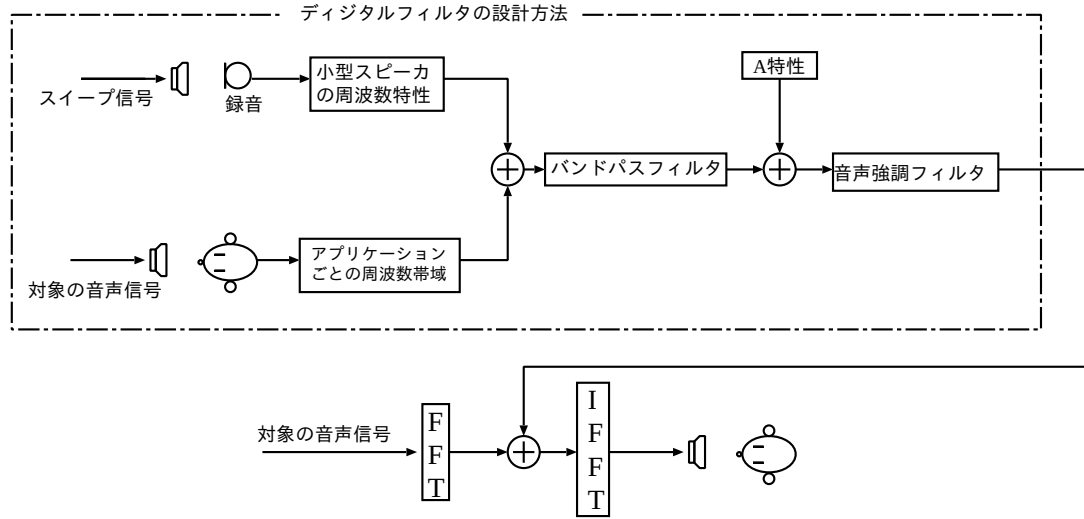


図 3.2 提案するシステム構成図

3.3.1 1/3 オクターブ分析

システムのアルゴリズムで説明した無駄な周波数帯域を削るバンドパスフィルタを 1/3 オクターブ分析で設計する。1/3 オクターブ分析とは、ISO の規格にある 1000Hz を基準に 1/3 オクターブごとに周波数の分析を行う、また、それぞれのバンドの中心周波数から上下限遮断周波数を求め処理を行う分析方法である。本研究ではバンド数を 1Hz～24000Hz までの 30 バンドとする。中心周波数 f_m の下限遮断周波数 f_1 を式 (3.1) から求め、上限遮断周波数 f_2 を式 (3.2) をから求める [7]。

$$f_1 = f_m \div \sqrt[6]{2} \quad (3.1)$$

$$f_2 = \sqrt[6]{2} \times f_m \quad (3.2)$$

また、上限遮断周波数と次のバンドの下限遮断周波数の間に周波数が存在してしまうため、バンドごとの処理を行うことができない。そこで、上限遮断周波数と次のバンドの下限遮断周波数を同じ周波数にすることで、上限遮断周波数と下限遮断周波数の間に周波数が溢れないようにする。また 0Hz は直流成分としているため、0Hz のパワースペクトルを 0 とする [8]。中心周波数から求めた上下限遮断周波数を付録 A の表 A.1 に示す。次に小型スピーカのアプリケーションごとのバンドパスフィルタの設計方法について説明する。

3.3.2 バンドパスフィルタ

まず、提案するシステムのアルゴリズムで説明した無駄な周波数帯域を削るバンドパスフィルタの通過域を小型スピーカの周波数特性と対象の音声ごとの周波数帯域から決定する．小型スピーカの周波数特性から通過域の最小通過周波数 ω_1 を決定し、対象の音声ごとの周波数帯域から通過域の最大通過周波数 ω_2 を決定する．

ω_1 の決定方法だが、小型スピーカの周波数特性を基に今回は音として聞こえ出した周波数を通過域の ω_1 と決定する．そこで ω_1 のパワーである -30dB を最小可聴値 (聴覚しきい値) とし、パワーが -30dB 以下の周波数は音として聞き取れないため削る．しかしミュージックプレイヤー用のバンドパスフィルタは、主に強調する音声を歌声と設定したため、基音の高い歌声では ω_1 を変更することでより効率的に無駄な周波数帯域を削ることができる．ハイパスフィルタでミュージックプレイヤー用の音声を処理し、音質に影響がない中心周波数をミュージックプレイヤー用の ω_1 と決定する． ω_1 の決定方法を図 3.3 に示す．

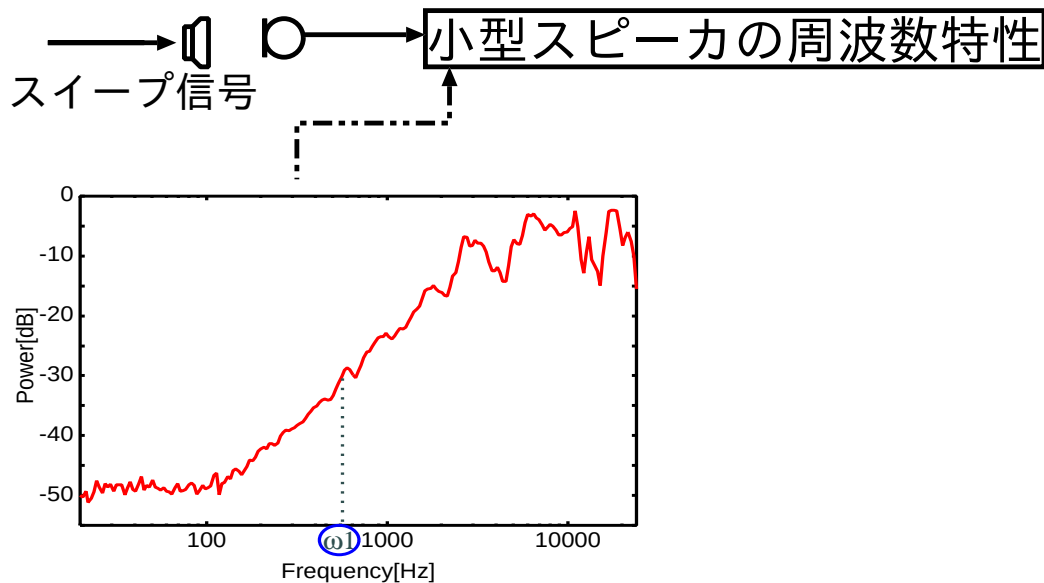


図 3.3 最小通過周波数 ω_1 の決定方法

ω_2 の決定方法だが、扱う音声ごとにより効率的に無駄な周波数帯域を削るため、扱う音声信号をローパスフィルタに通し小型スピーカで再生した結果、音質に影響がない中心周波

数を通過域の ω_2 と決定する．扱う音声ごとに ω_2 を変更するのは，ミュージックプレイヤーの場合と音声通話の場合だと含まれる最大周波数が異なるためである． ω_2 の決定方法を図 3.4 に示す．

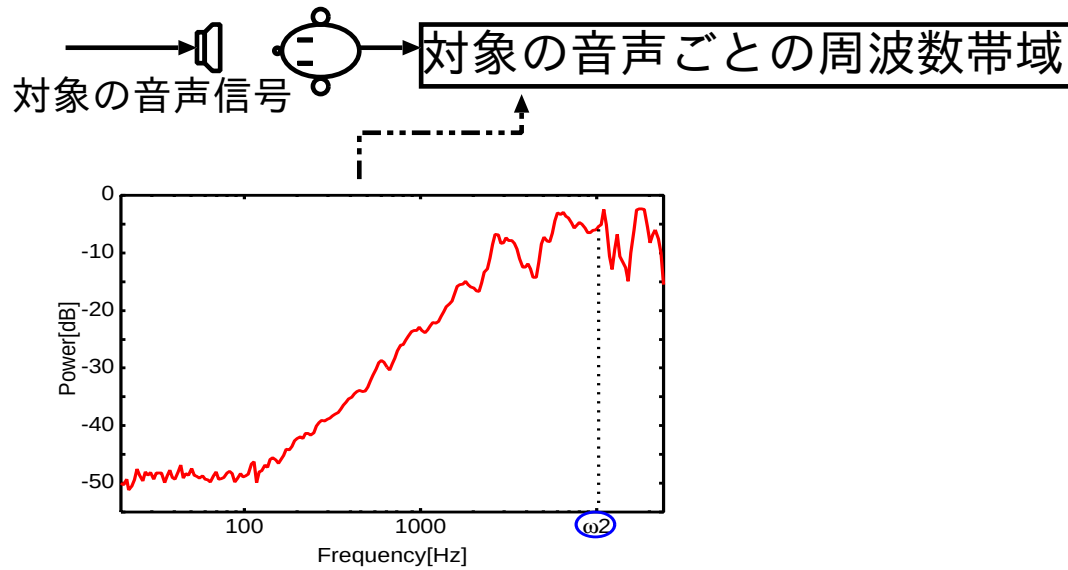


図 3.4 最小通過周波数 ω_2 の決定方法

これにより決定した ω_1 , ω_2 を基にバンドパスフィルタの通過域の決定方法を図 3.5 に示す．

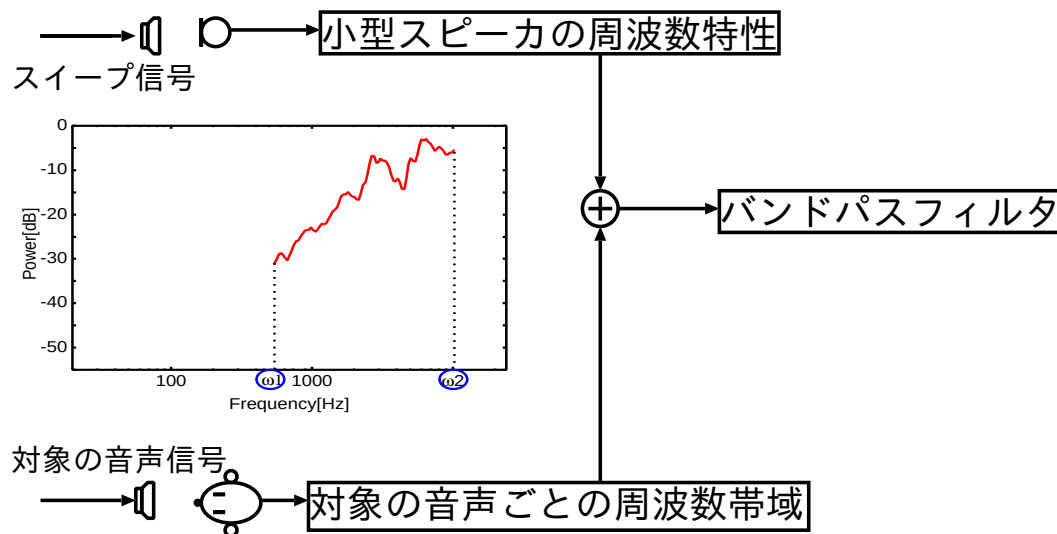


図 3.5 バンドパスフィルタの通過域の決定方法

設計した扱う音声ごとのバンドパスフィルタの最小・最大通過周波数を表 3.1 に示す.

表 3.1 設計したバンドパスフィルタの最小・最大通過周波数

対象の音声 \ 通過域の周波数	最小通過周波数	最大通過周波数
ミュージックプレイヤー用	中心周波数 1000Hz	中心周波数 16000Hz
音声通話用	中心周波数 630Hz	中心周波数 10000Hz
ワンセグ用	中心周波数 630Hz	中心周波数 16000Hz

次に, 扱う音声ごとに設計したバンドパスフィルタに重み付けを行う聴覚特性について説明する.

3.3.3 聴覚特性 (A 特性)

人間の聴覚には, 同じ音の大きさでも周波数が変われば, 音の大きさが変わって聞こえる音響心理学の等ラウドネス曲線で表される特性がある. この聴覚特性を用いた信号処理の例では, MP3 等の音声圧縮技術がある [9]. 等ラウドネス曲線とは, 1000Hz の音の大きさを基準とし, 周波数を変化させながら同じ音の大きさに聞こえる周波数の音圧レベルをプロットした曲線である [10]. JIS 規格にある等ラウドネス曲線を基に補正値をプロットした A 特性を図 3.6 に示す [11].

図 3.6 から人間の聴覚特性について以下の 2 点がわかる.

- 低音が極端に感度が悪くなる.
- 2000Hz~5000Hz は, 非常に感度が良い.

本研究では, 後者の聴覚特性に着目し, 設計したバンドパスフィルタに重み付けすることで, 音が大きく聞こえる音声強調フィルタを設計する. また, 人間の耳の構造上, 約 3000Hz の周波数が外耳骨で共振する周波数とされているため, この特性も考慮し音声強調フィルタを設計する [12].

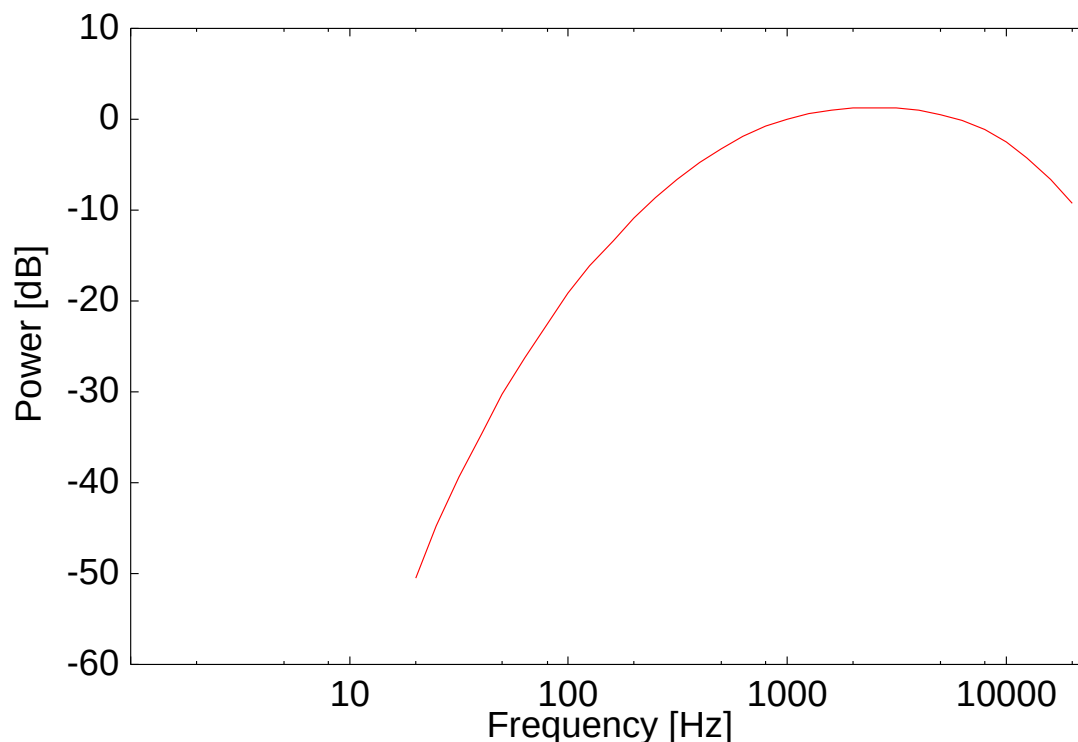


図 3.6 A 特性

3.3.4 音声強調フィルタ

次に、アプリケーションごとに設計したバンドパスフィルタに A 特性を基に重み付けを行い、特性の異なる音声強調フィルタを設計する．音声強調フィルタの設計方法を図 3.7 に示す．

A 特性の曲線だが、周波数帯域ごとの値を使用しているため実際は曲線ではなく、周波数帯域ごとに不連続な波形になる．そのため、重み付けを行い設計するフィルタも周波数帯域ごとに不連続である．周波数帯域ごとに不連続なフィルタで処理を行っているが、スプライン補間で周波数ごとに連続にしたフィルタで処理した音声と比較した結果、音質の変化がほぼなかったため、不連続なフィルタでも問題ないことを確認している．

設計した音声強調フィルタを図 3.8, 3.9, 3.10 に示す．図 3.8 は、A 特性を基に聴覚上感度の良いとされる 2000～5000Hz のパワーを +6dB 強調することで音が大きく聞こえるフィルタである．図 3.9 は、A 特性と耳の構造を基に図 3.8 で示されるフィルタの約 3000Hz

のパワーを +3dB 強調することで、音が大きく聞こえるフィルタである。図 3.10 は、2000～5000Hz の周波数帯域のパワーを -6dB 減衰させることで、2000～5000Hz 以外のパワーを強調できるため、音声フラットになり音が大きくなるフィルタである。この設計した音声強調フィルタを使用することで、処理を行った音声は入力範囲 (16bit) を超える場合があるが、適切な正規化を行うことでエネルギー量は一定のままである。入力範囲に満たない場合も同様である。

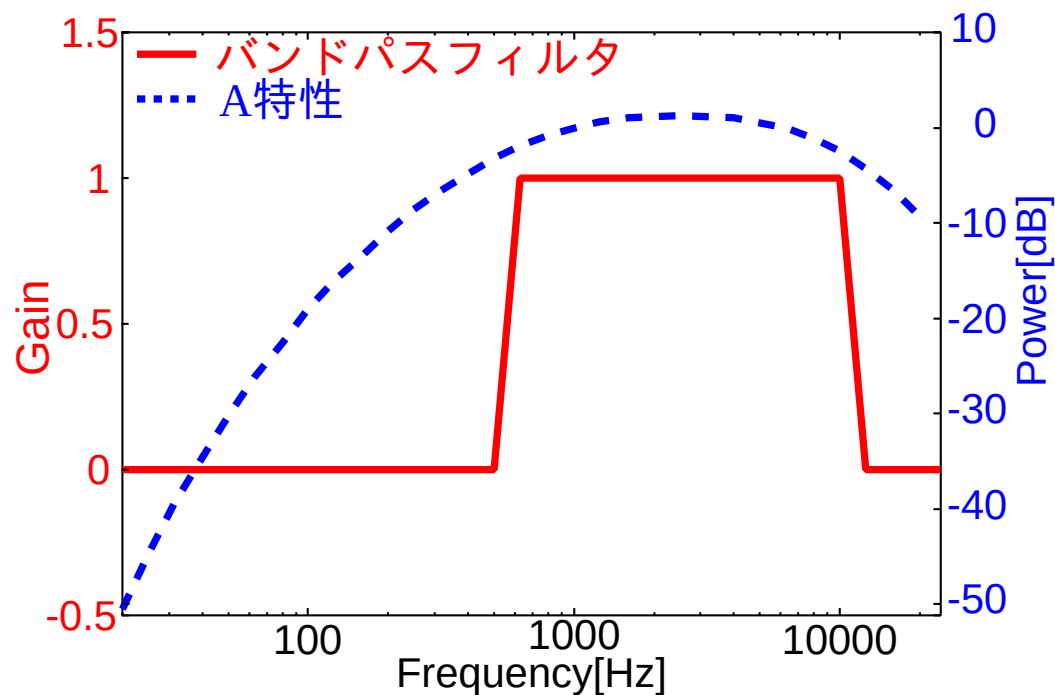


図 3.7 音声強調フィルタの設計方法

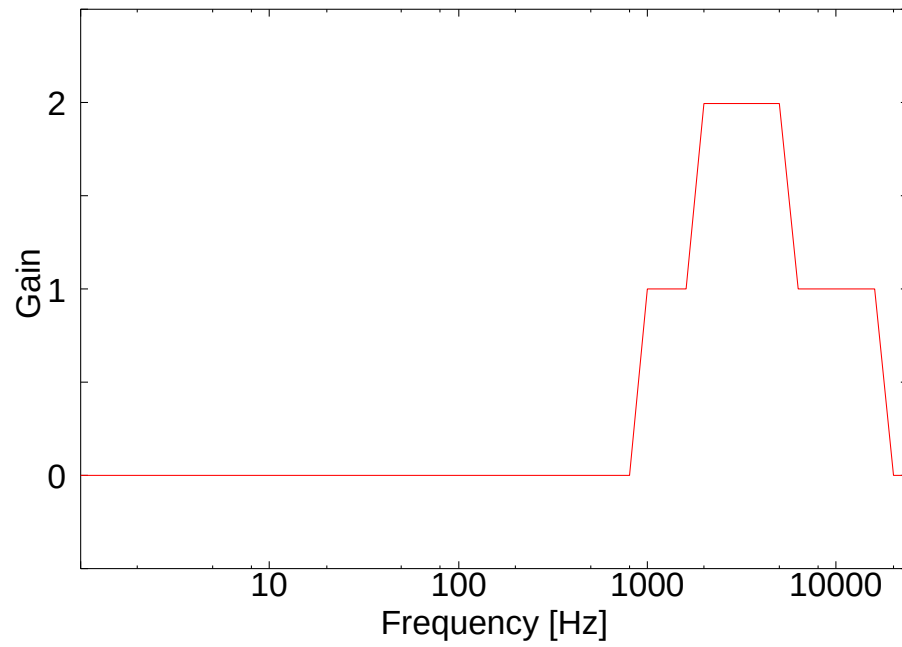


図 3.8 A 特性を基に音声を強調するフィルタ

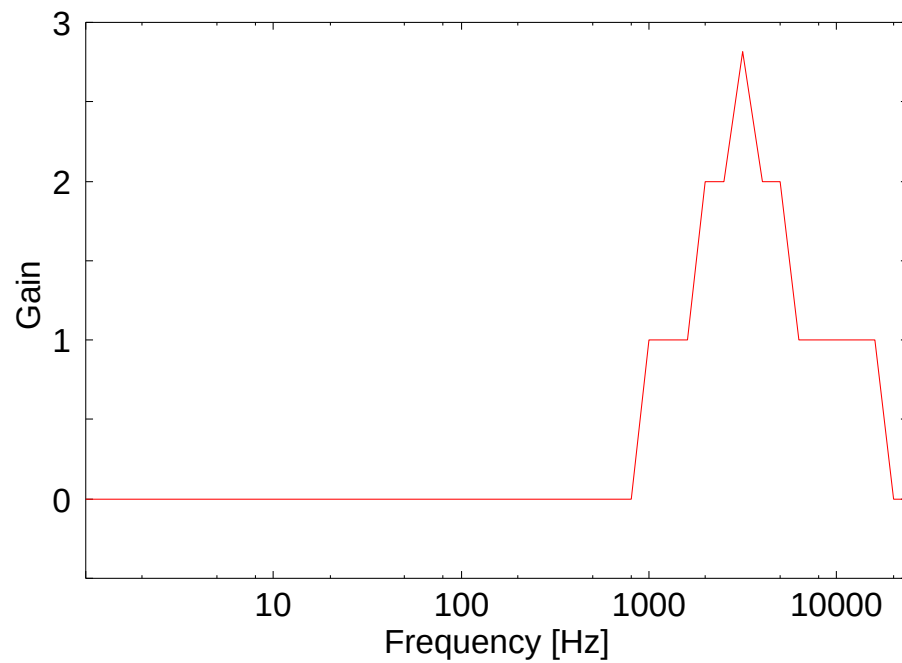


図 3.9 A 特性と耳の構造を基に音声を強調するフィルタ

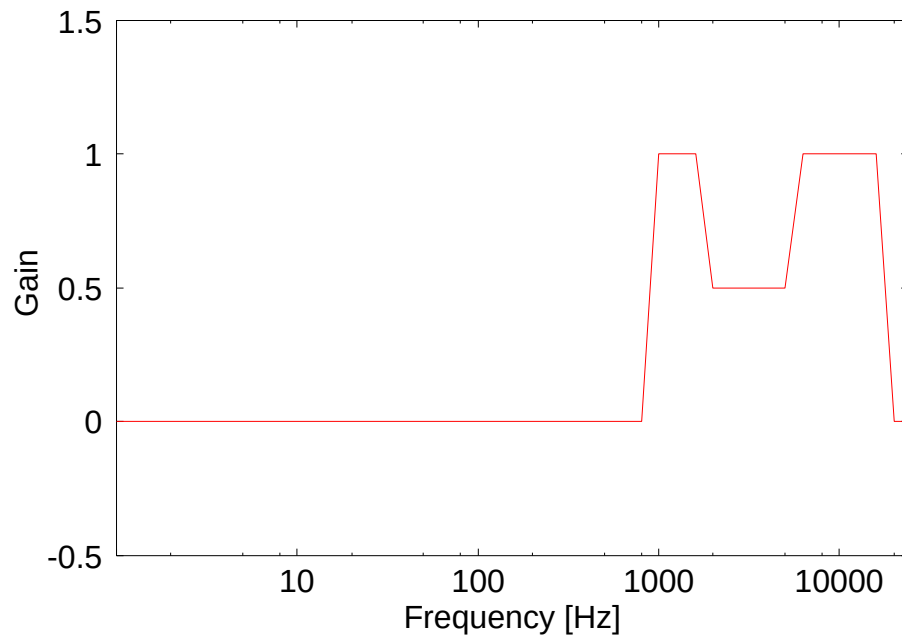


図 3.10 A 特性を基に音声をフラットにするフィルタ

3.4 システムの検証

1/3 オクターブ分析で決定したバンドごとの周波数帯域を処理するために，入力信号を FFT で周波数領域に変換し，バンドごとに音声強調フィルタを畳み込み，音圧を処理し IFFT で時間領域信号に変換する．そのため，FFT → IFFT を行った場合に信号に歪みが発生してはならない．そこで，FFT 後の信号を周波数領域に変換し，IFFT によって時間領域信号に変換した時，処理前と処理後の信号に振幅以外の変化がないことを検証する必要がある．検証には計算機シミュレーションを用いて，入力信号と FFT → 処理 → IFFT を行った信号を比較し，変化がないことを検証する．

FFT → IFFT の検証方法

まず，FFT → IFFT について動作の検証を行う．検証を行う FFT 環境を表 3.2 に示す．

表 3.2 の条件で FFT → IFFT を行った信号と行う前の元の信号に変化がないことを検証

表 3.2 FFT 環境

FFT プログラム	FFTW
サンプリング周波数	48000Hz
窓関数	ハミング窓

する．また，FFT に使用した窓関数 $\omega(x)$ は，式 (3.3) で求められる．

$$\omega(x) = 0.54 - 0.46 \cos 2\pi x, if(0 \leq x \leq 1) \quad (3.3)$$

FFT → IFFT の検証結果

FFT → IFFT の検証には白色信号を使用する．入力信号を図 3.11 に示し，FFT → IFFT を行った信号を図 3.12 に示す．

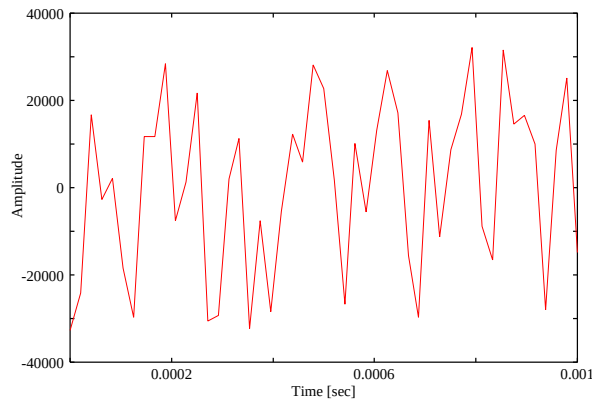


図 3.11 FFT → IFFT 前の信号

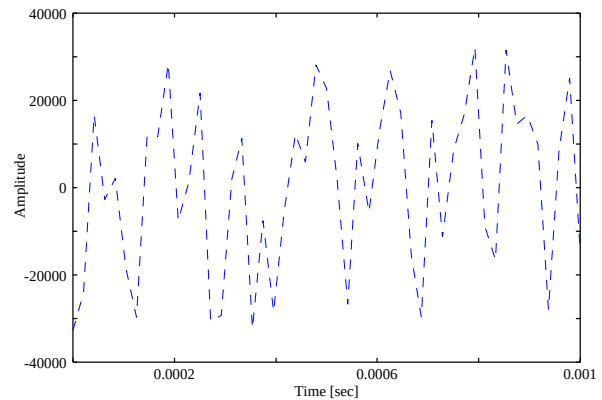


図 3.12 FFT → IFFT 後の信号

図 3.11 と図 3.12 の波形を確認して，入力信号と FFT → IFFT を行った後の信号に変化がないことが確認できる．

3.4.1 FFT →処理→ IFFT の検証方法

次に，FFT 後に処理を行った場合の検証方法について説明する．検証には，入力信号として振幅 1 の 440Hz の正弦波を使用し，440Hz の周波数の音圧を 2 倍にするフィルタを畳み

込み，時間領域と周波数領域の結果を検証する．実験環境は先で用いた表 3.2 と同様である．

FFT →処理→ IFFT の検証結果

検証結果を時間領域信号と周波数領域で確認した．時間領域信号の結果を図 3.13, 3.14 に示し，周波数領域の結果を図 3.15, 3.16 に示す．

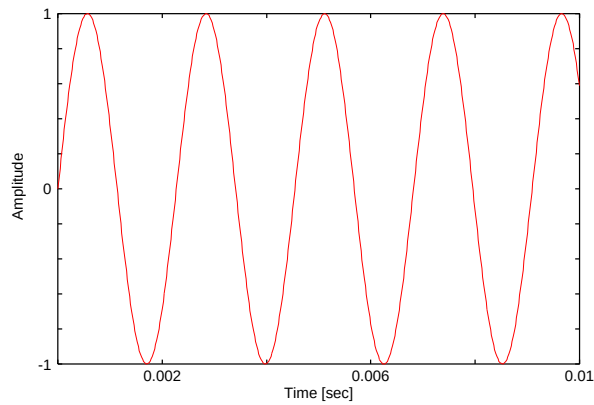


図 3.13 制御前の振幅 1 の正弦波 (440Hz)

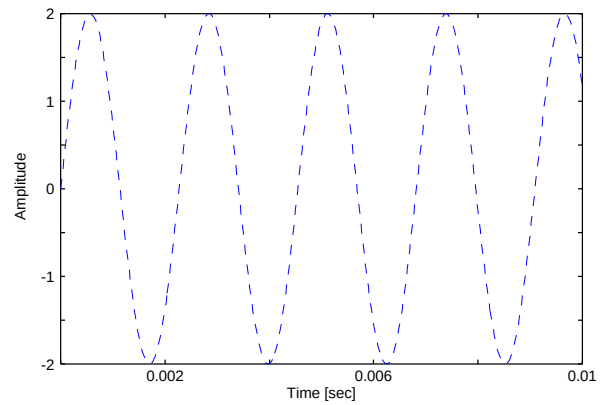


図 3.14 制御後の振幅 2 の正弦波 (440Hz)

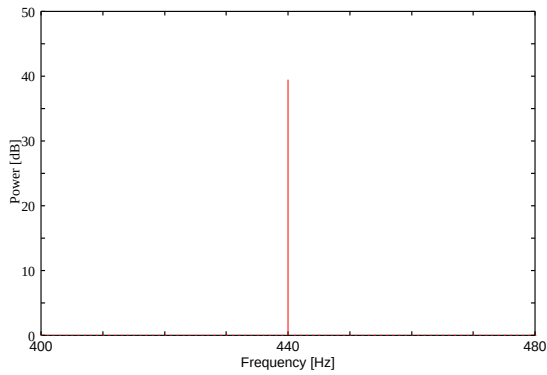


図 3.15 制御前の振幅 1 の正弦波 (440Hz)

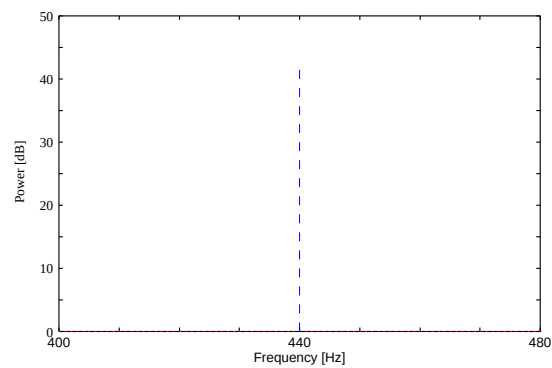


図 3.16 制御後の振幅 2 の正弦波 (440Hz)

FFT 後の実部を x_{re} とし，虚部を x_{im} としたパワースペクトル (Power) を式 (3.4) から求めた．

$$Power = (x_{re})^2 + (x_{im})^2 \quad (3.4)$$

図 3.13 の正弦波の振幅を 2 倍にしたものが，図 3.14 である．図 3.13, 3.14 から FFT 後

に 440Hz の正弦波の振幅を 2 倍に処理し、振幅以外に変化がないことが確認できる。また、図 3.15, 3.16 の結果から周波数領域のパワースペクトル以外の変化がないことが確認できる。これらの結果から、入力信号と FFT →処理→ IFFT を行った信号に振幅以外の変化がないことを検証した。

次に設計した音声強調フィルタの検証を行い、システムの有効性を評価する。

3.5 有効性の評価

提案するシステムの有効性の評価方法として、計算機シミュレーションによる評価と処理した音声による評価の 2 点で行う。計算機シミュレーションによる評価方法として、入力信号を設計した音声強調フィルタで処理し、バンドパスフィルタで設定した通過域以外のカットと音声強調フィルタで設定した周波数帯域ごとの音圧の処理を検証し、評価する。音声による評価方法として、システムのアルゴリズムの検証と設計したバンドパスフィルタ、音声強調フィルタで処理した結果を被験者実験で評価する。

3.5.1 計算機シミュレーションによる評価

設計した音声強調フィルタの計算機シミュレーションによる評価として、バンドパスフィルタで設定した通過域以外のカットと聴覚特性を基に設定した周波数帯域の音圧の処理について周波数領域で確認し、正常に処理していることを検証する。検証には図 3.8, 3.9, 3.10 で示した異なる特性の音声強調フィルタを使用する。計算機シミュレーションの環境を表 3.3 に示す。入力に使用した白色信号を周波数領域に変換した結果を図 3.17 に示す。

表 3.3 計算機シミュレーション環境

サンプリング周波数	48000Hz
入力信号	白色信号
FFT プログラム	FFTW

図 3.8 で示したフィルタで処理した音声を周波数領域に変換した結果を図 3.18 に示す．図 3.18 から設定した 2000～5000Hz の周波数帯域の音圧を +6dB に処理できていることが確認できる．図 3.19 で示したフィルタで処理し音声を周波数領域に変換した結果を図 3.19 に示す．図 3.19 から，図 3.9 を基に中心周波数 3150Hz の音圧だけを +3dB に処理できていることが確認できる．図 3.10 で示したフィルタで処理した音声周波数領域に変換した結果を図 3.20 に示す．図 3.20 から設定した中心周波数 2000～5000Hz の音圧を -6dB に処理できていることが確認できる．図 3.18，3.19，3.20 から異なる特性を持つ音声強調フィルタを使用してもバンドパスフィルタで設定した通過域以外のカットと設定した周波数帯域の音圧を制御できていることが確認でき，設計した音声強調フィルタで正常に処理できることを検証した．

次に，音声による評価を行う．

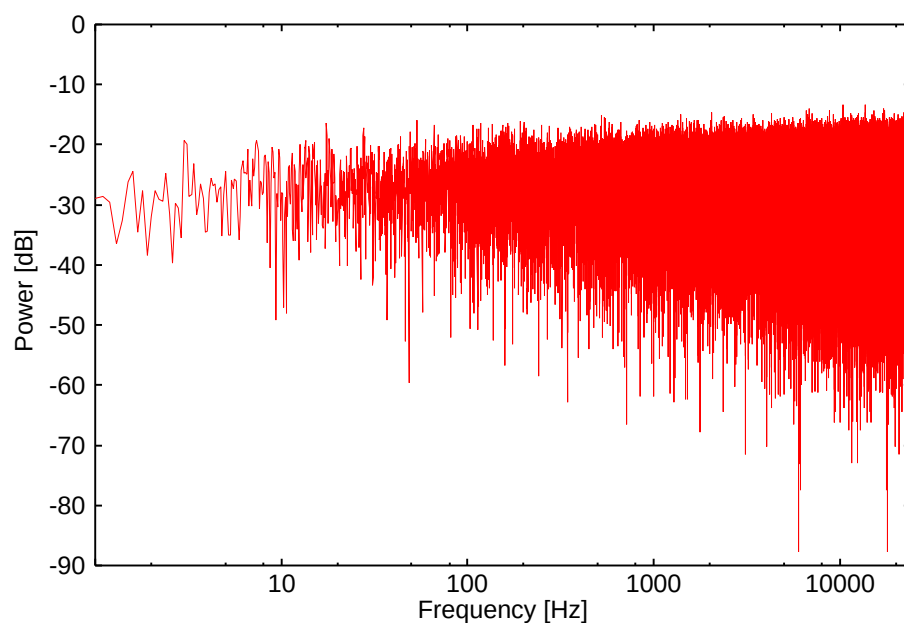


図 3.17 入力した白色信号

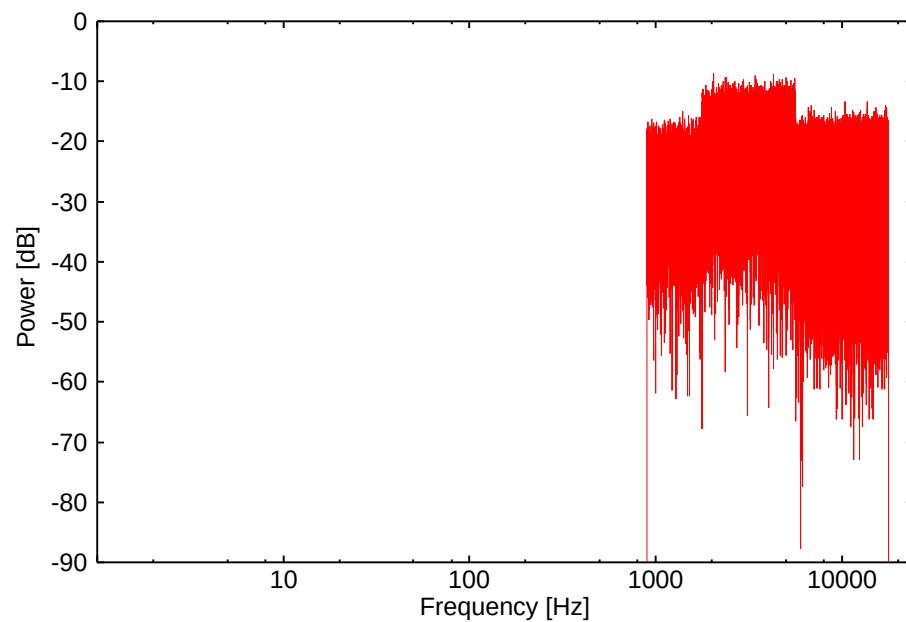


図 3.18 A 特性を基に強調した白色信号

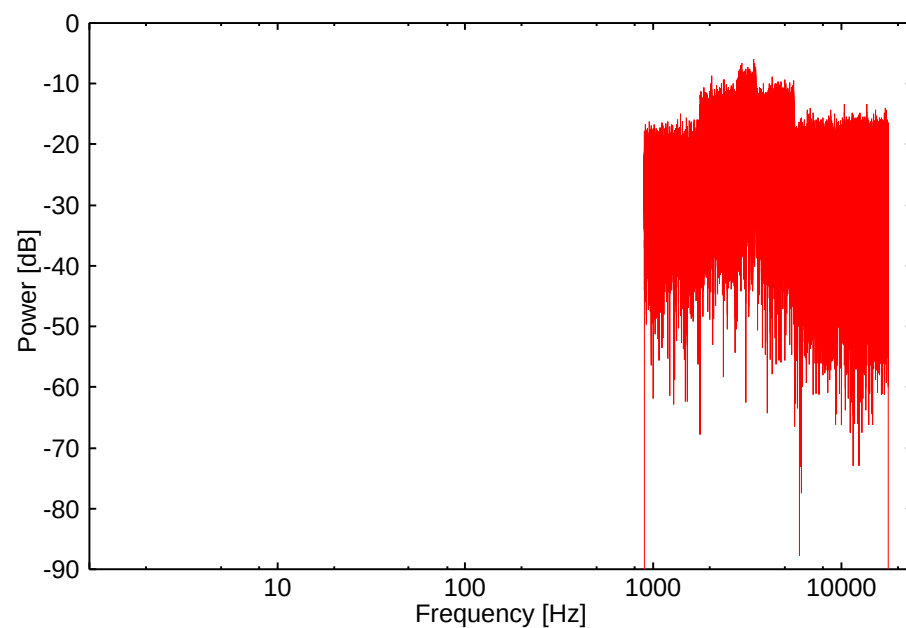


図 3.19 A 特性と耳の構造を基に強調した白色信号

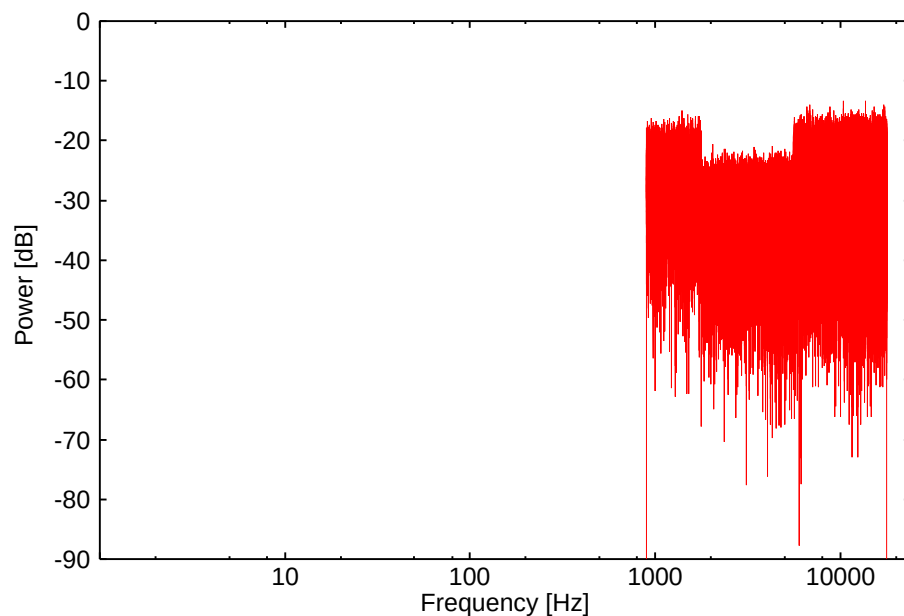


図 3.20 A 特性を基に減衰させた白色信号

3.5.2 音声による評価

音声による評価項目を示す.

- バンドパスフィルタによる音質の変化検証
- アルゴリズムの検証
- 音声強調フィルタによる音声の検証

まず、音声による評価として設計したバンドパスフィルタによる音質の変化がないか、原音を再生したものと聞き比べ評価する。次に、原音をそのまま再生するより提案する手法で処理した音声の方が音が大きくなるかシステムのアルゴリズムについて検証を行う。最後に、対象の音声信号に作成した音声強調フィルタで処理した音声が大きく聞こえるか評価する。入力に使用する音声を 1～7 の音声とする。1, 2 の音声は、音声通話の状況を想定し人の声を録音した音声である。3～6 の音声は、ミュージックプレイヤーを想定した音声である。7 の音声は、ワンセグの音声を想定した音声である。また、小型スピーカから再生される音声を聞く距離は、実際の使用時を想定しているため決めていない。人によっては、近付いて

聞く人もいたし、1m ぐらい離れて聞く人もいたため、その影響がどのようなになるかも検証する。被験者実験による環境を表 3.4 に示す。

表 3.4 被験者実験の環境

人数	5 人
性別	男性
年齢	20 歳代

- 対象の音声信号

1. 男性の声 (音声通話)
2. 女性の声 (音声通話)
3. J-pop : 男性歌手 (ミュージックプレイヤ)
4. J-pop : 女性歌手 (ミュージックプレイヤ)
5. 演歌 : 男性歌手 (ミュージックプレイヤ)
6. 演歌 ; 女性歌手 (ミュージックプレイヤ)
7. 会話とテレビの背景音 (ワンセグ視聴)

バンドパスフィルタによる音質の変化の検証

設計したバンドパスフィルタによる音質の変化がないか原音と聞き比べを行い検証する。検証方法として、1～7 の入力信号をそのまま再生したものとバンドパスフィルタを通し無駄な周波数帯域をカットした音声の音質の変化について検証を行う。

- 評価項目

- A. 音質は変わらない
- B. 少し変わったが聞いていても違和感なく聞ける
- C. かなり変わり原音とは全然違う音声になった

バンドパスフィルタによる音質の変化の検証結果と考察

評価結果を表 3.5 に示す。

表 3.5 バンドパスフィルタによる音質の変化の検証結果

評価項目 入力音声	A	B	C
1 の場合	3 人	2 人	0 人
2 の場合	4 人	1 人	0 人
3 の場合	4 人	1 人	0 人
4 の場合	4 人	1 人	0 人
5 の場合	4 人	1 人	0 人
6 の場合	4 人	1 人	0 人
7 の場合	5 人	0 人	0 人
計	28 票	7 票	0 票

この結果からどの音声もバンドパスフィルタによる音質の変化がほとんどないことが確認できる。特に 7 の場合は、他のバンドパスフィルタよりも通過域が広いワンセグ用のバンドパスフィルタを使用しているため、他のバンドパスフィルタに比べ音質に変化がないということがこの結果から確認できる。音質が変わったと答えた人にどのように音質が変わったか聞くと「少し楽器の音が無くなった」という意見や「全体的に高音になったような気がする」という意見があった。しかし、そう思う人もいれば音質に変化がないように聞こえたという人もいたため、聞く人が音を聞いている時に歌声に集中して聞くのか、楽器の音に集中して聞いているのかといった違いによる評価結果だと考えられる。6 の音声の場合、原音が音割れしている音声だったため、バンドパスフィルタを通すことで音割れが聞こえなくなり、「音質が良くなった」という意見もあった。また、「スピーカの近くで音声を聞くと音質が変わったように聞こえるが 50cm 程度離れると変わらない」と意見もあった。

アルゴリズムの検証

次に、システムのアルゴリズムが携帯電話に適しているか検証するため、バンドパスフィルタを通しパワーを相対的に大きくした音声と原音をそのまま再生した音声の音の大きさについて検証を行う。

● 評価項目

- A. 音の大きさは変わらない
- B. 少し大きくなったような気がする
- C. 大きくなったとすぐにわかる

アルゴリズムの検証結果と考察

評価結果を表 3.6 に示す。

表 3.6 アルゴリズムの検証結果

評価項目 入力音声	A	B	C
1 の場合	0 人	2 人	3 人
2 の場合	2 人	2 人	1 人
3 の場合	1 人	0 人	4 人
4 の場合	4 人	1 人	0 人
5 の場合	0 人	0 人	5 人
6 の場合	0 人	2 人	3 人
7 の場合	3 人	1 人	1 人
計	10 票	8 票	17 票

この評価結果から無駄な周波数帯域を削り、削った周波数帯域のエネルギー量を聞き取れ

る周波数帯域に使用するというシステムのアルゴリズムが小型スピーカに適していること確認できる。特に、5 の場合だと全員が大きくなったと評価した。これは、5 の音声は演歌なので楽器音や機械音があまり含まれてなく、主に歌声だけであることからバンドパスフィルタで削った結果歌声だけを強調でき音が大きく聞こえたのだと考えられる。逆に 4 の場合は、原音に楽器音や機械音、歌声と含まれていて原音自体が他の音声に比べると音が大きく聞こえるように加工された音声だったため、同じ処理を行っても 5 の場合のようにはならなかったと考えられる。

音声強調フィルタによる音声の検証

評価方法として、それぞれの音声を A～D に示す扱う音声ごとの音声強調フィルタで処理し、音声ごとに最も音が大きく聞こえた処理を選択する。

- 処理した音声

- A. 設計したバンドパスフィルタを通し相対的に大きくした音声
- B. 図 3.8 で示したフィルタを使用し処理した音声
- C. 図 3.9 で示したフィルタを使用し処理した音声
- D. 図 3.10 で示したフィルタを使用し処理した音声

音声強調フィルタによる検証結果と考察

評価結果を表 3.7 に示す。

この結果から音声によって、バンドパスフィルタを通し相対的にパワーを増やした音声より本システムで設計した音声強調フィルタで処理した方が音が大きく聞こえることがわかった。また、C の処理を行った 3, 4, 5, 6 の場合だと音声音が音割れを起きてしまうことが確認でき、C の処理をミュージックプレイヤーに使用することはできないということが確認できた。また、B の処理を行った音声が多いため 2000～5000Hz を強調することで音が

表 3.7 音声強調フィルタによる検証結果

入力音声 \ 処理した音声	A	B	C	D
1 の場合	1 人	0 人	0 人	4 人
2 の場合	0 人	4 人	1 人	0 人
3 の場合	2 人	1 人	音割れ	2 人
4 の場合	1 人	4 人	音割れ	0 人
5 の場合	4 人	0 人	音割れ	1 人
6 の場合	3 人	2 人	音割れ	0 人
7 の場合	0 人	2 人	3 人	0 人
計	11 票	13 票	4 票	7 票

大きくなることが確認できた。しかし、B の処理を行った音声は周波数帯域のパワーを変え
るため、音質が少し変わってしまう。2 番目に音が大きく聞こえた A の処理を行った音声な
ら音質もほとんど変わらず、音を大きくできるということが確認できた。また 3 番目に音が
大きく聞こえた D の処理を行った音声だが、1, 3, 5 の場合に評価を得たことから主に男性
の声が含まれている音声に対して有効であることが確認できた。また、バンドパスフィルタ
による音質の変化の検証結果で述べたのと同様に、人によって大きく聞こえる処理方法が異
なることから、ミュージックプレイヤーの場合だと歌声が大きくなったから音が大きく聞こえ
たと思ったり、楽器の音や背景音のような BGM が大きく聞こえたから音が大きく聞こえた
というように人によって捉え方が異なるのだと考えられる。

3.6 まとめ

小型スピーカで音声を再生する場合、エネルギー量を増やして音を大きくできないという
問題からエネルギー量を一定のまま音を携帯電話に適した音を大きくするシステムのアルゴ

リズムの提案を行い，アルゴリズムに基づくバンドパスフィルタを設計し，バンドパスフィルタに聴覚特性を重み付けし設計した音声強調フィルタで音声の処理を行った．システムの有効性を評価するため，計算機シミュレーションによる評価と音声による評価を行いシステムの有効性を評価した．計算機シミュレーションによる結果は正常に動作できていることが確認できた．しかし，音声による評価を行った結果，音声や受聴者によってシステムの有効性に差があることが確認できた．そのため，A 特性以外の重み付けの方法を考える必要がある．

第 4 章

結論

4.1 本研究のまとめ

携帯電話で音楽を再生すると音量が小さく、外方向のスピーカに近付かなければ聞こえにくいという問題がある。その原因である小型スピーカの問題点を検証し、小型スピーカに適したシステムのアルゴリズムとアルゴリズムを基にシステムの処理方法を提案した。また、提案したシステムの処理方法で携帯電話に適したデジタルフィルタを設計し、音が大きく聞こえるようにした。設計したデジタルフィルタの計算機シミュレーションによる評価と音声による評価を行い、システムの有効性を評価した。しかし、受聴者や音声によってシステムの有効性に差があるといった問題や実際にシステムを実装した場合のことを想定できていないという問題がある。今後、これらの問題を解決する必要がある。

4.2 今後の課題

今後の課題として 3 点挙げられる。

まず 1 点目として、バンドパスフィルタの通過域の最大通過周波数の決定方法が聴覚による主観的な決定方法だったため、より適切な決定方法にする必要がある。

次に 2 点目として、今回は対象の音声ごとにバンドパスフィルタを手動で選択しなければならないため、携帯電話のマイクを使用し、周波数解析を行い自動でバンドパスフィルタを選択できる方法を検討する。また、実装にあたり処理はフィルタを通すだけなので処理時間はほとんどかからないと考えられるが、自動でバンドパスフィルタを選択する際、どのくら

い時間が必要なのかを検証する必要がある．ハンズフリーの場合などでは，リアルタイムで処理できなければならないため，バンドパスフィルタを選択することに時間がかけられない．

最後に 3 点目として，今回は設計したバンドパスフィルタに A 特性を重み付けしたが，周波数帯域の音圧を変化させるため，音質が多少変化してしまうという問題点がある．そこで，音声通話の場合だとフォルマント強調を使用すれば音質にほぼ変化を与えず音を大きくできると考えられる．しかし，音楽のように複数の音が含まれているミュージックプレイヤーやワンセグなどの場合の処理方法を考えなければならない．A 特性以外の重み付けを行うことで，本研究で生じたシステムの有効性の差を軽減できると考えられる．

謝辞

本研究を行うにあたり、御指導ならびに御助言を頂いた福本昌弘教授に深く感謝致します。また本研究の副査をして頂きました浜村昌則准教授、繁桝博昭講師に感謝致します。

福本先生には、2年次からアドバイザーとして、単位の取得状況や新年度の予定など様々なことを心配して頂き有難うございました。また就職活動の時は、履歴書の書き方や積極的に活動しない私に厳しい言葉を頂き、その結果5月に内定を取ることができたことも感謝致します。3年次には家庭があるにも関わらず休みの日や夜遅くまで、での悪い我々に付き添って頂き感謝致します。福本先生の「社会に出たら…」、「新聞を読め」、「姿勢が悪い」という口癖からもわかるように、常に我々のことを気にかけて頂き、この福本研究室で学んだことを糧に来春から、カッコいい(外見的な意味ではなく)社会人になれるよう頑張ります。

また、研究のことや趣味のことで優しく相談にのって頂いた福本研究室の兄貴的な存在の佐伯幸郎助教にも深く感謝致します。正月などの休みの日にも研究を手伝って頂き有難うございました。佐伯先生には忙しい中、論文の添削やプレゼンの作り方、発表練習など御迷惑をおかけしたにも関わらず、意見を出して頂いたり、一緒に考えて頂き有難うございました。また、晩ご飯やコーヒーなど様々な差入れ、研究に集中できる環境作りなど色々気を使って頂き有難うございました。おかげで効率よく研究を進めることができました。

また、3年次に輪講でプログラミングを教えて頂いた福富英次助手にも深く感謝致します。福富先生に3年次に教わったプログラミングをいざ研究で使うとなったときに忘れてしまって、厳しい言葉を頂きくじけそうになったのですが、細かいことも分かりやすく教えて頂き、有難うございました。「僕はここに居ていいんだ」と思いました。福富先生のおかげで人並みにプログラミングができるようになり、研究を行えました。また人生で初めてのクリスマスケーキ有難うございました。とてもおいしかったです。

また、中間発表会やオープンキャンパスの時などにドーナッツやパンを差入れして下さった修士2年の徳久翔太氏にも深く感謝致します。での悪い後輩で心配や苦勞をかけた

と思いますが、研究のことや研究に使う機材の使い方などを教えて頂き有難うございました。また、私生活でも共通の趣味のお店でジュースの差入れ有難うございました。徳久さんのひきの強さには感服しました。

また、楽しい時間を提供してくれた同研究室の卒業生の皆さんにも深く感謝致します。被験者実験で貴重なコメント有難うございました。この場をお借りして、厳しい言い方をして皆さんを傷付けたかもしれないことを深くお詫びします。

また、3年生の泰地優太氏には先輩として、就職活動や卒業研究について何も教えてあげられなかったことをお詫び致します。来年度からもしかしたら1人になり、オープンキャンパスの準備や卒業研究など大変かもしれませんが、屈強な先生方からたくさんのことを学んでください。

最後にこの場をお借りして、私の大学生活を支えてくれた家族はもちろんのこと、関わりのある全ての皆様に御迷惑をおかけしましたが、温かく見守って頂いたり、時には厳しく御指導頂いたことをお詫びするとともに深く感謝致します。有難うございました。

参考文献

- [1] 財団法人 地域流通経済研究所, “携帯電話の用途は「メール、電話、カメラ、時計・アラーム」”, http://www.dik.or.jp/pdf/press_0802_keitai.pdf
- [2] 平成 18 年度 特許出願技術動向調査報告書, “最新スピーカ技術-小型スピーカを中心に- (要約版)”, 2007.
- [3] 文責 メカライフ編修委員, “知って納得！メカランド スピーカ編”, 日本機械学会誌 Vol.110 No.1066, P726, 2007.9.
- [4] 藤井信生, “アナログ電子回路の基礎”, 昭晃堂, 2004.
- [5] 富士通研究所, “音声信号処理技術 (携帯電話) 携帯電話「らくらくホン」の「はっきりボイス」”, http://jp.fujitsu.com/group/labs/techinfo/techguide/list/voice_p05.html
- [6] 山本武夫, “モニタスピーカの変遷”, 日本音響学会誌 49 巻 9 号, 1993.
- [7] 小野測器, “騒音計とは 11-1 オクターブ分析について”, <http://www.onosokki.co.jp/HP-WK/c.support/newreport/noise/souon-11.htm#mark11-1>
- [8] 三谷政昭, “やり直しのための信号数学”, CP 出版社, 2005
- [9] 守谷健弘, “音声符号化”, 電子情報通信学会, 1998.
- [10] ISO226(2003), “Acoustics-Normal equal-loudness-level contours”, 2003.
- [11] JIS C 1509, “電気音響-サウンドレベルメータ (騒音計)”, 2005.
- [12] 堺 久雄, 中山 剛, “聴覚と音響心理”, コロナ社, 2002.
- [13] H.P. スウ, 佐藤平八 訳, “フーリエ解析”, 森北出版, 1979.
- [14] 三谷政昭, “ディジタルフィルタデザイン”, 昭晃堂, 1997.

付録 A

中心周波数と上下限遮断周波数

第 3 章 3.3 で使用した 1/3 オクターブ分析における中心周波数と上下限遮断周波数の範囲を表 A.1 に示す.

表 A.1 中心周波数と上下限遮断周波数

バンド数	中心周波数	下限遮断周波数	上限遮断周波数
1	中心周波数以外 (低音)	1	35
2	40	35	44
3	50	44	56
4	63	56	70
5	80	70	89
6	100	89	111
7	125	111	140
8	160	140	178
9	200	178	224
10	250	224	280
11	315	280	353
12	400	353	445
13	500	445	561
14	630	561	707
15	800	707	898
16	1000	898	1122
17	1250	1122	1403
18	1600	1403	1781
19	2000	1781	2227
20	2500	2227	2806
21	3150	2806	3535
22	4000	3535	4454
23	5000	4454	5612
24	6300	5612	7071
25	8000	7071	8908
26	10000	8908	11136
27	12500	11136	14030
28	16000	14030	17817
29	20000	17817	22449
30	中心周波数以外 (高音)	22449	24000

付録 B

フーリエ変換

B.1 フーリエ変換の特徴

フーリエ変換は信号解析に使用される一般的な方法である．ある任意の時間信号を周波数領域で表したものである．フーリエ変換には，離散周期信号を扱う離散フーリエ変換 (Discrete Fourier Transform, DFT)，離散フーリエ変換によって周波数領域に変換した信号を時間領域信号に変換する離散フーリエ逆変換 (Inverse Discrete Fourier Transform, IDFT)，離散フーリエ変換の数値計算を効率よく行う高速フーリエ変換 (Fast Fourier Transform, FFT)，高速フーリエ変換によって周波数領域に変換した信号を時間領域信号に変換する高速フーリエ逆変換 (Inverse Fast Fourier Transform, IFFT) がある．

B.2 フーリエ変換の手順

フーリエ変換の計算手順について説明する．周期信号は三角多項式での近似が可能である．これを無限級数で表すと，

$$f(t) = \frac{1}{2}a_0 + a_1 \cos t + a_2 \cos 2t + \cdots + b_1 \sin t + b_2 \sin 2t + \cdots \quad (\text{B.1})$$

のような式になる．ただし， $a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_n$ は定数である．この式は次のように表すことができる．

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt) \quad (\text{B.2})$$

式 (B.2) を $f(t)$ のフーリエ級数と呼ぶ．

周期を T とした関数は、 $-\frac{T}{2} \leq t \leq \frac{T}{2}$ の範囲の信号の周期的な繰り返しになる．このような周期を持つ関数は $\omega_0 = \frac{2\pi}{T}$ とおくと

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos n\omega_0 + b_n \sin n\omega_0) \quad (\text{B.3})$$

と表すことができる．さらにこの式をオイラーの公式を用いて複素指数として表すと

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t} \quad (\text{B.4})$$

が得られる．ただし、 c_n は、

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) e^{-jn\omega_0 t} dt \quad (\text{B.5})$$

と表すことができる． C_n は、正弦波および余弦波のスペクトルを表す．フーリエ変換はこれを使用したもので、信号を周波数成分に分解する．信号 $f(t)$ に対してフーリエ変換を行うと、

$$F[f(t)] = \int_{-\infty}^{\infty} f(t) e^{-j\omega_0 t} dt \quad (\text{B.6})$$

となる．