

フリーソフトを用いた 音声処理の実際

石井 直樹 著

コロナ社

まえがき

21世紀に入った頃から、音声処理を応用した技術がわれわれの生活のすみずみに浸透してきている。多くの国民がその恩恵にあずかりながら、気がついていない場合もある。いまでは、機械がその場面・状況に応じた音声を発するのは当たり前のことと捉えている。スマートフォン（スマホ）を含めて携帯電話の音声は、昔の固定電話の音声とはまったく異なる技術で伝えられているが、（気がついている方もいるが）違和感なく使用している。スマホに話しかけて、何か情報を得ているらしい光景を目にするのも増えてきた。

これらの技術は、機器内部のLSI（大規模集積回路）あるいは、あるいはアプリケーション（アプリ）としてメモリの中に隠蔽されており、その内容をうかがい知ることはできない。音響学、音声学、あるいは音声工学は、やや狭い専門分野の学問であり、大学あるいは専門学校ではそれを意図する特定の学生しか履修できない分野になっている。さらに、これら三つの分野はたがいに密接に関連しているが、これらを総合的に学ぶことすら難しいのが現状である（複数学部の講義を聴講し、自分で関連付けせねばならない）。

かたや、コンピュータ関連のハードウェアおよびソフトウェアの進展（変化）は目覚ましく、ひと時代前のものが使用できない、あるいは通用しないこともしばしば経験する。このコンピュータの力を最大限に利用して、音声あるいは音の信号（以下、サウンドと総称しよう）を自由自在に扱う手助けをするのが、本書の狙いである。最新（でなくてもよい）のコンピュータの性能は、処理能力、記憶能力、表示能力のすべての点で、サウンドを扱う上で申し分ない。しかし、それは「しかるべきソフトウェア（アプリ）」が備わっていればの話である。残念なことに、最新のコンピュータを購入しても、それだけでサウンドを扱える範囲は限られている。やりたいサウンド処理に応じたソフトウェアを入手し、それをコンピュータに組み込み、かつそのソフトウェアの扱い方に習熟して初めて、目的とするサウンド処理が可能となる。

従来、サウンド処理のソフトウェアは、特に専門的で複雑な処理をするものは、かなり高額なものが多く、比較的低価格なソフトは実行できる機能や性能が限られていた。しかし、最近「フリーソフト」と称される、無償で使うことができるソフトウェアで、機能的にも性能的にも高度なものが現れてきた。残念ながらそのようなソフトウェアは欧米の研究者が開発したもので、メニューおよびヘルプは英語で記述されており、ややハードルが高かった。だが、国内の先駆的な大学研究者を中心にこのようなソフトウェアを使いこなして、教材に使う方が増え、日本語の解説ドキュメントも公表されるようになってきた。

筆者は、2002年というWindows MEの時代に『音声工房を用いた音声処理入門』¹⁾（前著と呼ぶこととする）を出版し、類書もないことから、大学等で教科書、参考書などに採用さ

れ、また独習用の教材として利用されて、現在まで増刷を重ねてきた。しかし、執筆から15年以上も経過すると、さすがに内容の古い箇所が目につく。一つはパソコンのオペレーティングシステム（OS）である Windows の変遷、二つ目はパソコンでサウンドを扱う環境の変化、三つ目はサウンドを扱うフリーソフトの勃興である。

このような状況を鑑みて、本書『フリーソフトを用いた音声処理の実際』を世に問うこととした。すなわち、前著は「音声工房 Pro」というかなり高価なソフトウェアを使用することを想定して、サウンド処理を始めようとする読者を主対象としていた。それに対して、本書は無償で入手できるフリーソフトを使って、サウンド処理の実際を伝授しようとした。フリーソフトの種類によっては、実行できる処理の種類、精度、あるいは処理結果の表示法が異なる場合がある。読者が実施したい処理に適するフリーソフトを選択する指針をも与えている。

サウンドを扱うために用意するパソコンは、最近の機種である必要はない。ただし、モニタ（ディスプレイ）の解像度は XGA（1024 × 768 ドット）以上であるほうが操作しやすい。また、音声データは大きな記憶領域を必要とするので、大きめのハードディスクを備えていることが望ましい。

パソコンの OS は、本書で紹介するフリーソフトを扱うためには Windows に限る。ただし、最新の Windows 10 でなくとも、Windows 8/7 で十分である。すでにサポートが終了した Windows XP/Vista が搭載された旧機種でも構わない。

パソコンにフリーソフトを導入するためにインターネット（以下ネット）にアクセスする場合、二つの注意点がある。一つは、コンピュータウイルスが仕組まれたソフトに気をつけることである。そのためには、本書に掲載した、信用できるサイト（窓の杜、Vector、SourceForge、など）からダウンロードすることである（SourceForge には最近悪い報道もあるが）。また、ダウンロードした圧縮ファイルは解凍する前に、ウイルス検査ソフトで確認すること（ただし、ウイルス検査ソフトは、最も高い割合でウイルスが仕組まれたソフトであることを知っておこう）。もう一つは、所望のフリーソフトの案内画面のそばに（より目立つ形で）配置されている有料ソフトのボタン（じつは、広告である）に注意すること。また、所望のフリーソフトを正しく選択した場合でも、（おまけの不要な）別のソフトを合わせてインストールさせようとするものもある。案内文（英語）をしっかりと読み、「自己責任」で対処していただきたい。

2018年8月

著 者

本書で使用している会社名、製品名は、一般に各社の商標または登録商標です。本書では®と™は明記していません。

肩付き数字は、巻末の文献番号を表す。

目 次

1. パソコンのサウンド機能

1.1 サウンドデバイス	1
1.1.1 サウンドデバイスとは	1
1.1.2 サウンドデバイスの中身は	1
1.1.3 サウンドデバイスの組み込み状況を調べる —「サウンド」—	2
1.2 ノートパソコン	3
1.3 タブレット型パソコン	4
1.3.1 Windows タブレットのサウンド機能	5
1.3.2 サウンド回路の特性	5
1.4 デスクトップ型パソコン	5
1.4.1 デスクトップパソコンのサウンド機能	5
1.4.2 内蔵のサウンド回路の性能	7

2. パソコン用音響機器

2.1 アナログ音響機器	8
2.2 パソコン用デジタル音響機器 — USB オーディオ機器	9
2.2.1 USB マイク	9
2.2.2 USB DAC	9
2.2.3 USB スピーカ (USB DAC 内蔵スピーカ)	10
2.2.4 USB ヘッドセット	10
2.2.5 USB オーディオインタフェース	10
2.2.6 USB について	11
2.2.7 USB 規格	12
2.2.8 USB 端子	12
2.2.9 USB ケーブル	13
2.3 ボイスレコーダ, デジタル録音機	13
2.3.1 ボイスレコーダ	14
2.3.2 音声符号化方式	14
2.3.3 メモリ容量, 録音時間	14
2.3.4 マイクロホン	14
2.3.5 パソコンとのデータの受け渡し — USB, μ SDHC —	15
2.3.6 デジタル録音機	15
2.4 Bluetooth 音響機器の利用	15

3. Windows におけるサウンドの扱い

3.1 Windows 10 の場合	16
3.2 Windows 8/8.1 の場合	18
3.2.1 サウンドレコーダ — Windows ストアアプリ —	18
3.2.2 サウンドレコーダ — デスクトップアプリ —	19
3.3 Windows 7 の場合	21

4. サウンド用フリーソフト

4.1 フリーソフトについて	23
4.1.1 フリーソフトの種類	23
4.1.2 サウンド用フリーソフト	25
4.2 SoundEngine Free 音声編集フリーソフト	31
4.2.1 概 要	31
4.2.2 SoundEngine Free の基本操作	31
4.2.3 SoundEngine Free の分析機能	32
4.2.4 SoundEngine Free のエフェクト機能	33
4.2.5 評 価	33
4.3 Audacity (A free audio editor and recorder)	33
4.3.1 概 要	33
4.3.2 Audacity の基本操作	34
4.3.3 Audacity の分析機能	34
4.3.4 Audacity のエフェクト機能	35
4.3.5 評 価	35
4.4 WavePad 音声編集ソフト	35
4.4.1 概 要	35
4.4.2 WavePad の基本操作	36
4.4.3 WavePad の分析機能	37
4.4.4 WavePad のエフェクト機能	37
4.4.5 WavePad のサンプル音声	37
4.4.6 評 価	38
4.5 SFS/WASP, SFSWin	38
4.5.1 SFS について	38
4.5.2 SFS/WASP, SFSWin の概要	38
4.5.3 SFS/WASP の基本操作	38
4.5.4 SFS/WASP の分析機能	40
4.5.5 SFSWin の基本操作	40

4.5.6	SFSWin の編集機能	41
4.5.7	SFSWin の分析機能	41
4.5.8	SFSWin のラベル付与機能	43
4.5.9	SFSWin の信号音／合成音生成機能	43
4.5.10	その他 Windows 用のソフトウェア	44
4.5.11	評 価	44
4.6	Praat	45
4.6.1	概 要	45
4.6.2	Praat の基本操作	45
4.6.3	Praat の音声編集機能	47
4.6.4	Praat の音声分析機能	47
4.6.5	文字表記の付加 (Annotation)	49
4.6.6	Praat における音声変換機能	49
4.6.7	Praat における変声機能	50
4.6.8	Praat の音声合成機能	50
4.6.9	評 価	51
4.7	WaveSurfer	52
4.7.1	概 要	52
4.7.2	WaveSurfer の基本操作	52
4.7.3	WaveSurfer の音声編集／変換機能	53
4.7.4	WaveSurfer の音声分析機能	53
4.7.5	文字表記の付加 (Transcription)	54
4.7.6	評 価	54
4.8	Speech Analyzer	55
4.8.1	概 要	55
4.8.2	Speech Analyzer の基本操作	55
4.8.3	Speech Analyzer の音声編集機能	56
4.8.4	Speech Analyzer の音声分析機能	56
4.8.5	Speech Analyzer のラベリング機能	57
4.8.6	Speech Analyzer の特異な機能	57
4.8.7	評 価	58
4.9	SASLab Lite	58
4.9.1	概 要	58
4.9.2	SASLab Lite の基本操作	58
4.9.3	SASLab Lite の音声編集／変換機能	59
4.9.4	SASLab Lite の音声分析機能	60
4.9.5	SASLab Lite の信号音作成機能	61
4.9.6	SASLab Lite の特異な機能	62
4.9.7	評 価	62

4.10 Raven Lite	62
4.10.1 概 要	62
4.10.2 Raven Lite の基本操作	63
4.10.3 Raven Lite の音声編集機能	64
4.10.4 Raven Lite の音声分析機能	65
4.10.5 Raven Lite の特異な機能	65
4.10.6 評 価	66
4.11 その他のフリーソフト	66
4.11.1 音声工房および音声録聞見	66
4.11.2 Sound Analysis Pro 2011	66
4.11.3 Wavosaur (Audio Editor with VST Support)	67
4.11.4 XMedia Recode (メディアファイル変換ソフト)	69
4.11.5 Moo0 ボイス録音機 (ストリーミング音声録音ソフト)	70
4.11.6 恋 声	70
4.11.7 SoX (Sound eXchange)	71

5. 音 と 音 声

5.1 音の基本知識	72
5.1.1 音	72
5.1.2 音の波形	74
5.1.3 サウンドソフトにおける波形の表示	75
5.2 音声の特徴	78
5.3 音のデジタル化	79
5.3.1 デジタル化とは	79
5.3.2 標本化 (サンプリング)	79
5.3.3 量子化	80
5.3.4 デジタル化に際しての注意	81
5.3.5 アナログ信号に復元する際の注意	83
5.4 音声デジタル化と音質	83
5.4.1 標本化周波数と音質	84
5.4.2 量子化ビット数と音質	84
5.4.3 過負荷雑音	87
5.5 音声伝送容量の削減	88
5.5.1 モノラル化, 狭帯域化	88
5.5.2 片方向通信用の音響符号化	88
5.5.3 高能率波形符号化法	89
5.5.4 Windows のオーディオ CODEC	89

5.5.5	高度の音声符号化方式	90
5.5.6	音声音響符号化技術	90
5.5.7	ロスレス圧縮（可逆圧縮）	90

6. サウンド波形の編集

6.1	サウンド波形の表示・観測	91
6.1.1	Audacity による波形表示	92
6.1.2	SoundEngine による波形表示	93
6.1.3	Wavosaur による波形表示	94
6.1.4	WavePad による波形表示	95
6.1.5	SFS/WASP による波形表示	95
6.1.6	SFSWin による波形表示	96
6.1.7	Praat による波形表示	97
6.1.8	WaveSurfer による波形表示	99
6.1.9	Speech Analyzer による波形表示	100
6.2	サウンド波形の操作・編集	101
6.2.1	振幅を変える	101
6.2.2	音のレベルを合わせる	103
6.2.3	音の分割／切貼り／切出し	104
6.2.4	音のミキシング	105
6.2.5	ステレオ（2チャンネル）信号の操作	106
6.2.6	反響（エコー）と残響（リバーブ）	107
6.3	フィルタ	107
6.4	雑音除去	108
6.4.1	雑音区間の除去	108
6.4.2	ノッチフィルタ	110
6.4.3	残響・反響の除去・軽減	111
6.4.4	人声の除去（Vocal Remover）	111
6.5	発声速度，声の高さ，継続時間の変更	112
6.5.1	音声波形の一部を変更する — 継続時間と発声速度を変える —	112
6.5.2	音声データの全体に発声速度と高さを変更 — リサンプリング —	113
6.5.3	声の高さのみを変える — ピッチシフト —	114
6.6	変声，声質変換	115
6.6.1	男声⇔女声変換	115
6.6.2	ボイスチェンジャ，変声機	118
6.6.3	ヘリウムボイス	118
6.7	信号音の作成	118

6.7.1	作成できる信号音の種類	118
6.7.2	信号音の成形	119
6.7.3	合図音の作成	120
6.7.4	複合正弦音の作成	120

7. 言語音声の特徴と音声分析

7.1	言語音声の特徴	122
7.2	音声分析とは	123
7.2.1	スペクトル分析	124
7.2.2	音声生成器官に関する物理量の分析	124
7.2.3	その他の分析法	124
7.3	音声パワーとその時間変化	125
7.4	基本周波数とその時間変化	126
7.5	パワースペクトル	128
7.6	スペクトル包絡 (LPC スペクトル)	133
7.7	パワースペクトルの時間的变化を表示する方法	134
7.8	スペクトログラム	134
7.9	フォルマントとその時間変化	137
7.9.1	フォルマント軌跡の分析	138
7.9.2	フォルマントの表現法	140
7.10	その他の分析法	142
7.10.1	調波性 (Harmonicity)	142
7.10.2	ケプストラム分析	143
7.10.3	声門パルス分析	143
7.10.4	点過程分析	144
7.11	音声分析の応用	145
7.11.1	音声符号化	145
7.11.2	音声合成	146
7.11.3	音声認識	146

8. 言語音声の波形と特徴量の観測

8.1	言語音声の波形の観測	148
8.2	母音の波形と特徴量の観測	150
8.3	子音の波形と特徴量の観測	153
8.3.1	無声閉鎖音	153

8.3.2	有 声 閉 鎖 音	155
8.3.3	摩 擦 音	156
8.3.4	鼻 音	157
8.3.5	半 母 音	159
8.3.6	発声様式の変化	160
8.4	長 音	163
8.5	連 母 音	164
8.6	韻 律 的 特 徴	164
8.6.1	発 話 速 度	165
8.6.2	ア ク セ ン ト	165
8.6.3	イ ン ト ネ ー シ ョ ン	166
8.6.4	リ ズ ム	167
8.6.5	感 情	168
8.6.6	個 人 性	168

9. 特殊な発声音声の分析

9.1	歌 声 の 分 析	169
9.1.1	唱 歌	169
9.1.2	May J. の 声	171
9.1.3	ホ ー ミ ー	172
9.2	い ろ い ろ な 発 声	174
9.2.1	ひ そ ひ そ 声	174
9.2.2	だ み 声	176
9.2.3	し わ が れ 声 (嗚 声)	178
9.2.4	裏 声, フ ェ ル セ ッ ト な ど	180
9.2.5	腹 話 術	181
9.3	動 物 音 声 の 分 析	183
9.3.1	哺 乳 類	183
9.3.2	鳥 類	185
9.3.3	鳥 類 (キ ュ ウ カ ン チ ョ ウ)	189
9.3.4	蛙 (カ ジ カ ガ エ ル)	191
9.3.5	そ の 他	191

引用・参考文献	193
---------	-----

索 引	194
-----	-----

1. パソコンのサウンド機能

パソコンを立ち上げると、「ジャララーン」という音が出る（OSのバージョン、あるいは設定状況によって異なるかもしれないが）。このように、パソコンはサウンドを再生する機能が組み込まれている。さらに、マイクロホンを接続すると音声をパソコンに取り込むこともできる。パソコンのサウンド機能は、パソコンの形態（ノート／タブレット／デスクトップ／その他）により異なるので、自分が使用するパソコンに備えられている機能を把握・理解しておく必要がある。

1.1 サウンドデバイス

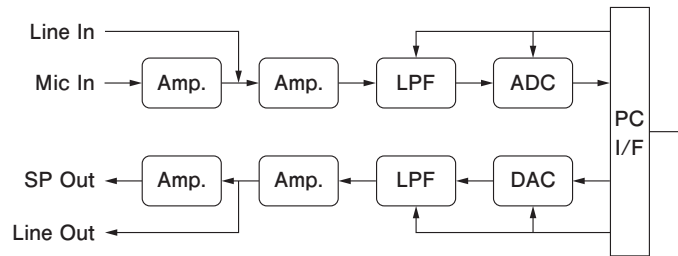
1.1.1 サウンドデバイスとは

パソコンでサウンド（音）機能を扱うハードウェアを**サウンドデバイス**と呼ぶ（デバイスは通常、装置と訳されるが、コンピュータの世界では、ある機能を果たす部品といった意味で用いられる）。従来は、パソコン本体の印刷配線基板（**プリント基板**と呼ぶことが多い）とは別の基板の形態をしていたので、サウンドカード、あるいはサウンドボードなどと呼ばれていた。最近のパソコンでは、サウンド機能を果たす電子部品は、パソコン本体の**基板**（**マザーボード**と呼ばれる）上に直接取り付けられていることが多い。ノートパソコンの場合はほとんどがそうであり、デスクトップパソコンもこのタイプが増えてきた。サウンド信号を本格的に扱う人は、パソコン本体に組み込まれたサウンドデバイスの機能性能では満足できず、拡張スロットに装着する特別のサウンドボードや、パソコン外部に設置してパソコンとデジタル接続する装置を使用している。このように、サウンドデバイスはさまざまな形態をしている。

1.1.2 サウンドデバイスの中身は

サウンドデバイスの一般的な機能ブロック構成を図1.1に示す。パソコン本体に組み込まれた簡易なサウンドデバイスでは、一部の回路が省略されていることもある。あるいは、別の方式の符号復号器を組み込んでいるなど広い用途のサウンドデバイスもある。

Line In（ライン入力）あるいは**Mic In**（マイク入力）端子から入った信号は、**Amp.**（増幅器：amplifier）により適正な振幅にされたのち、**LPF**（低域フィルタ：low pass filter）に

図 1.1 サウンドデバイスの機能ブロック構成[†]

入れられ、切断周波数以上の成分が除去される。その出力が **ADC** (**A-D 変換器**：analog to digital converter) でデジタル信号に変換 (**符号化**と呼ばれる) され、インタフェース部 (I/F) を経由して、他のパソコン回路に送られる。一方、パソコン内部からのデジタル形式のサウンド信号は、インタフェース部を経由して、**DAC** (**D-A 変換器**：digital to analog converter) に入れられ、アナログ信号に変換 (**復号化**と呼ばれる) される。DAC の出力は LPF にて不要な成分を除去されたのち、アンプを通過して **SP Out** (スピーカ出力)、あるいは **Line Out** (ライン出力) 端子に出力される。なお、パソコン側で設定した標準化周波数や量子化ビット数 (後述) などの条件は、インタフェース部を経由して、ADC、DAC、および LPF に送られ、その条件で符号化・復号化が行われるよう制御される。なお、LPF の切断周波数は、標準化周波数の半分程度に設定される。

1.1.3 サウンドデバイスの組み込み状況を調べる —「サウンド」—

あなたのパソコンにどのようなサウンドデバイスが組み込まれている (あるいは接続されている) か、もし組み込まれているならどのような製品、あるいは特性のものであるか、を調べる方法を説明しよう。この方法は、パソコンのメーカー/機種、基本ソフトウェア (Windows など) のバージョン、および組み込まれているサウンドデバイスによって若干異なる。

Windows 10 の場合を例にして、サウンドデバイスの組み込み状況を調べよう。コントロールパネルから [サウンド] を指定すると、[サウンド] という名前のウィンドウが現れる (別図 1.1[†])。このウィンドウには、再生、録音、サウンド、通信という四つのタブがあり、[再生] のタブが選択されているはずである。タブの下の表示欄には、あなたのパソコンに組み込まれている、あるいは接続されている [再生デバイス] が何段かに分かれて表示されている。そのうち、アイコンの右下にチェックマークが付され、その右に [既定のデバイス] と書かれたデバイスが、現在サウンド再生用に設定されているものである。なお、右端の 10 本の横棒は (簡易) レベル計である。

[†] 本書に掲載した図のカラー版、および本書では掲載しないが、参考になる図面をコロナ社のホームページ <http://www.coronasha.co.jp/np/isbn/9784339009163> に掲載しているので、参考にしていただきたい。

なお、ある再生デバイスを選択し、右クリックして現れるメニュー(別図 1.2) から、[スピーカーの設定]、[テスト] その他の機能を実行することができる。また、デバイスのない白紙のところを右クリックして、[無効なデバイス]、[切断されているデバイス] を表示するかどうか、設定することができる。

[サウンド] のウィンドウで [録音] のタブを選択すると、同じように、パソコンに組み込まれている [録音デバイス] の一覧が表示され(別図 1.3)、接続・有効化されたデバイスがあれば、アイコン右下にチェックが入り、[既定のデバイス] と表示されている。[既定のデバイス] と設定されているデバイス ([マイク] などと表示されている) を選択し、右下の [プロパティ] ボタンを押すと、[マイクのプロパティ] と記されたウィンドウが現れる(別図 1.4)。そのウィンドウの [レベル] タブを押し、現れるスライダを左右に動かすことにより、マイクから入力された音量を調整することができる(別図 1.5)。

あなたのパソコンに新たにサウンドデバイスを追加した場合には、パソコンがそのデバイスを扱うためのソフトウェア(デバイスドライバ、あるいはデバイスマネージャ)を自動的に探索して組み込んでくれる。コントロールパネルで新しいデバイスを選択して有効化すれば、そのデバイスを使用することができる。

ここでは、Windows 10 が搭載されたデスクトップ型パソコンを例に、サウンドデバイスの搭載状況を調べる方法を紹介したが、ほかのバージョンの Windows、あるいは型の異なるパソコンでもほぼ同様である。つぎの節では、いろいろのタイプ(ノート/タブレット/デスクトップ)のパソコンに備わっているオーディオ端子と、その音響特性について調べてみる。

1.2 ノートパソコン

ノートブック型パソコン(以下、ノートパソコンと呼ぶ)には、通常マイクロホン端子とヘッドホン端子が備わっている(図 1.2 参照)。ヘッドホンへのサウンド出力の音量調整ボタンが付いている場合もある。また、パソコン本体に小形のラウドスピーカが組み込まれている場合も多い。

ノートパソコンの機種によっては、マイクロホン端子とヘッドホン端子とが合体し、ヘッドセット端子(4極のミニプラグでヘッドホンマイクを接続する)になっているものもある。



図 1.2 ノートパソコンのマイクロホン端子とヘッドホン端子(右側は USB 端子)

4 1. パソコンのサウンド機能

る。パソコンのマイクロホン端子は、通常プラグインパワー（plug-in power）方式に対応しており、電源を必要とするプラグインパワー方式のマイクロホンを接続することができる。また、マイクロホン端子はモノラルであり、ステレオのマイクロホンを接続しても、モノラル信号しか録音されない。なお、この端子にプラグインパワー方式でないダイナミックマイクロホン（カラオケ用など）を接続しても、不具合になることはない。

ヘッドホン端子はステレオになっており、ヘッドホンあるいはアンプ内蔵スピーカなどを接続する。

サウンド回路の特性

ノートパソコンの内部に収められているプリント基板には、大小さまざまな電子部品がぎっしりと搭載されており、その一部にサウンド機能を果たす部品が装着されている。マイクロホン端子からの配線が、このサウンド用の電子回路に接続されている。このような状況であるから、いくら回路実装上の工夫をしても、サウンド用電子回路には電氣的雑音の混入は避けられない。

あるノートパソコン（性能的にも、価格的にもハイエンドの機種である Let's Note CF-B10）のマイク入力系の SN 比を実測すると、約 43 dB であり、別の小形のノートパソコン（Dynabook SS M10）の場合は、約 56 dB であった。このように、サウンド系の性能は、必ずしもパソコン性能に追従するものではなく、むしろ製造メーカーの設計能力によっているようである。

ノートパソコンを用いて高い SN 比で音声を録音するには、後述の USB オーディオ機器を利用するのがよい。ほとんどすべてのノートパソコンには、外部デジタル機器と接続するために、USB 端子が装備されている。USB オーディオ機器側のマイクロホンから音を取り込み、デジタル出力端子から、USB ケーブルによりノートパソコン側にデジタル信号で送るのである。このほうが、マイクロホンが取り込んだ微弱な電気信号をデジタル変換する場所が雑音源から遠くなり、結果的によいサウンドを取り込むことができる。上述のハイエンドのノートパソコンの場合、USB オーディオ機器を経由してマイク録音する方法では、SN 比は約 62 dB まで改善した。

このような事情は、タブレット型パソコンや、デスクトップ型パソコンでも同様である。

1.3 タブレット型パソコン

アップル社が iPad という名称の携帯情報端末を発売して以降、タブレットと呼ばれる装置が世の中に受け入れられようになった。iPad に続いて、アンドロイドタブレットと

Windows タブレットが出現した。iPad のユーザインタフェースは Microsoft 社にも影響を与え、Windows 8 の誕生に結び付いたのであろう。

Windows タブレットは、小形・薄型のノートパソコンとみなすことができ、たいていのサウンドソフトも正常に動作する。タブレット端末は軽くて携帯に便利で、バッテリーも長時間もつので、サウンドソフトを搭載したタブレットにはそれに適した使い方も考えられる。

1.3.1 Windows タブレットのサウンド機能

Windows タブレットには、ヘッドホン端子が装備されているが、マイク端子はない（機種により異なるかもしれないが）。その代わりに、小形のマイクロホンが埋め込まれている。また小形のスピーカも装備されている。マイクロホンは背面または側面に配置されており、モニタ画面を見ながらの発声では、十分な音量では録音できない。2.2.1 項で述べるように、USB 端子に接続する USB マイクロホン（USB マイク）などを利用したほうがよいであろう（携帯性はやや下がるが）。なお、Windows タブレットの USB 端子は、C 型と呼ばれる小形のものが採用されていることが多い。このような場合は、A 型-C 型の USB ケーブル、あるいは変換アダプタを使用すればよい。

1.3.2 サウンド回路の特性

ノートパソコンの場合と同じように、あるいはそれ以上に、あまり良好な特性は期待できないであろう。Windows タブレット（Epson TB01S）に USB オーディオ機器を接続した場合、マイク録音系の SN 比は約 33 dB という低い値であった。

1.4 デスクトップ型パソコン

従来、デスクトップ型パソコンといえば、黒または白っぽい色の四角い箱で、モニタやキーボードを外付けしたものであるというイメージであった。それが近年では、モニタの後部に本体が一体化されたものや、小さな弁当箱タイプのものが現れ、ついにはスティック型（USB 端子を有する IC レコーダ状—実際には、HDMI 端子である）のものまで現れた。

したがって、これらパソコンのサウンド入出力端子もさまざまであり、装備していないものもある。ここでは旧来のやや大きめの筐体のデスクトップパソコンを想定して、そのサウンド機能を調べることにする。

1.4.1 デスクトップパソコンのサウンド機能

大きな筐体のデスクトップパソコンでは、サウンドカードの実装法に 2 種類がある。一つ

索引

【あ行】

アクセント核	165
アクセント句	165
アンチエイリアシングフィルタ	82
イコライザ	108
イントネーション	166
韻律の特徴	164
裏声	180
エイリアシングひずみ	82
音	72
—の大きさ	73
—の三要素	73
—の種類	123
—の高さ	73, 122
—の強さ	123
オフセット	82
オーバフロー	81
オープンソースソフトウェア	24
折返しひずみ	82
音圧レベル	77
音場	78
音色	74
音声	78
音声工房	66
音声特徴量	123
音声分析ソフト	25
音声録聞見	66
音節	167
音紋	134

【か行】

概周期的波形	149
可逆圧縮法	90
カクテルパーティー現象	78
隠れマルコフモデル	147
過負荷雑音	87
期間限定版	24
基底膜	136
機能限定版	24
基板	1
球面波	72
強勢	165
空間図	140
グラフィックイコライザ	108
クリエイティブ・コモンズ	24
ケプストラム	143
言語音声	122
検知限	107
恋声	70

高域強調フィルタ	108
高域通過フィルタ	108
高速フーリエ変換	130
高能率波形符号化法	89
高能率符号化	89
コントロールパネル	2

【さ行】

サウンドデバイス	1
サウンド用フリーソフト	23
サウンドレコーダ	18
嗄声	178
散布図	140
シェアウェア	23
時間的变化	123
指向性	73
ジッタ	144
シマ	144
自由空間	73
周波数スペクトル	128
純音	74
瞬時音圧	74
試用版	24
シラブル	167
しわがれ声	178
信号音	118
スイープ信号	119
スペクトル微分量	66
スペクトル分析	124
スペクトル包絡	131, 133
スペクトル密度	128
スペクトログラム	134
スムージングフィルタ	83
生物音響学	183
声紋	134
声門パルス	143
セクション	37
接近音	159
ゼロクロス	33, 104, 150
線形予測	90
線形予測法	133, 145
線形量子化	80
線スペクトル対	147
増幅器	1
ソナグラム	134
ソノグラム	124, 134
ソフトウェア	16
疎密波	72

【た行】

帯域阻止フィルタ	108
----------	-----

帯域通過フィルタ	108
ダイナミックイコライザ	108
縦波	74
だみ声	176
単語編集	105, 146
単発的な波形	149
チャープ	119
長音	163
超音波	74
長時間平均スペクトル	54
超低周波音	74
調波性	142
調波成分対雑音比	142
低域強調フィルタ	108
低域通過フィルタ	108
低域フィルタ	1
低域ろ波器	82, 83
デジタルフィルタ	82, 107
適応差分PCM	89
テキスト合成	146
デシベル	73
デスクトップアプリ	18
デバイスドライバ	3
デバイスマネージャ	3
点音源	72
点過程	144
頭声	180
動物音声分析ソフト	25
トーン	119

【な行】

ナイキスト周波数	82
二重母音	164
音色	74
ノイズ	119
ノッチフィルタ	108

【は行】

ハイパスフィルタ	108
ハイブリッド符号化	146
ハイレゾ音源	86
ハイレゾリユーション音源	80
波形	72, 74
波形接続型	146
波形符号化	146
波形符号化法	89, 145
パブリックドメイン	
ソフトウェア	24
パルス符号変調	14, 80
パルス密度変調	79
パワースペクトル	128

パワーボタン 125
 バンドストップフィルタ 108
 バンドパスフィルタ 108
 半母音 160
 鼻音 157
 鼻音化 162
 ひずみ 90
 非線形量子化 80
 ひそひそ声 174
 ピッチ 74
 ピッチシフト 114
 ピッチ同期波形重畳法 112
 ピッチパターン 126
 標本化 79
 標本化周波数 79
 ピンクノイズ 119
 ファルセット 180
 フィルタ 107
 フォルマント 137
 フォルマント軌跡 137
 フォン 73
 不規則な波形 149
 複号化 2
 複合正弦音 120
 復標本化周波数 83
 腹話術 181
 符号化 2
 符号復号器 89
 プラグインパワー 4
 プリエンファシス 131
 フリーソフト 23
 フリーソフトウェア 24
 プリント基板 1
 ブルートゥース方式 8
 プレイビューモード 22
 分解能 80
 分析合成符号化 146
 分析合成符号化法 145
 文節 146
 ペアリング 15
 閉鎖休止区間 153
 平面波 72
 ヘッドボイス 180
 ヘリウムボイス 118
 変声 50
 変声機 118
 ボイスチェンジャ 118
 ボイスレコーダ 14, 16
 母音三角形 141
 ホーミー 172
 ホワイトノイズ 119

【ま行】

摩擦音 156
 マザーボード 1
 窓関数 130
 ミキシング 105
 ミックスボイス 180
 ミュージック 20
 無音 119
 無声化 161
 無声閉鎖音 153
 メル尺度 139
 モータセオリー 78
 モダン UI アプリ 18
 モーラ 165

【ら行】

ラベリング 49, 54, 57
 ラベルの挿入 39
 離散フーリエ変換 132
 リサンプリング 113
 リズム 167
 量子化 79
 量子化雑音 86
 量子化精度 80
 量子化ビット数 72
 連母音 164
 録音/編集ソフト 25
 ロスレス 90
 ロスレス符号化法 90
 ローパスフィルタ 82, 108

【わ行】

わたり 135

【アルファベット】

AAC 18, 19, 90
 ADC 2
 ADPCM 89
 A-D 変換 79
 A-D 変換器 2
 Amp 1
 Annotation 49
 ANR 123
 Audacity 33
 Bluetooth 15
 CC ライセンス 24
 CODEC 89
 DAC 2
 D-A 変換 79
 D-A 変換器 2
 dB 102
 DFT 132

DPCM 89
 DSD 79
 DTFM 信号 120
 EVS 90
 FFT 130
 FLAC 90
 HMM 147
 HNR 142
 Line In 1
 Line Out 2
 LPC 133
 LPC スペクトル 133
 LPF 1
 LSP 147
 m4a 形式 17
 MBSOLA 43
 Mic In 1
 MP3 88
 PCM 方式 14
 PCM 80
 PDM 79
 Phase Vocoder 71
 Praat 45
 PSOLA 70, 112
 Raven Lite 2.0 62
 RMS 125
 SASLab Lite 58
 SFS 38
 SFS/WASP 38
 SFSWin 38
 Sound Analysis Pro 2011 66
 SoundEngine Free 31
 SoX 71
 spectral derivatives 66
 Speech Analyzer 55
 SP Out 2
 Transcription 54
 USB 9
 USB DAC 9
 USB オーディオインタフェース 10
 USB オーディオ機器 9
 USB ヘッドセット 9
 USB マイク 9
 VoLTE 90
 WavePad 35
 WaveSurfer 52
 Wavosaur 67
 Windows Media Player 22
 Windows ストアアプリ 18
 WMA 22
 wma ファイル 20

— 著者略歴 —

1965年 京都大学工学部電子工学科卒業
1965年 日本電信電話公社（現、NTT）電気通信研究所勤務
1982年 日本電信電話公社電気通信研究所音声入出力方式研究室室長
1987年 NTTアドバンステクノロジー株式会社勤務
2005年 横浜国立大学産学連携推進本部勤務
2013年 退職

フリーソフトを用いた音声処理の実際

Details of Speech Processing Techniques Using Free Softwares

© Naoki Ishii 2018

2018年12月28日 初版第1刷発行

検印省略

著者 石井直樹
発行者 株式会社 コロナ社
代表者 牛来真也
印刷所 新日本印刷株式会社
製本所 有限会社 愛千製本所

112-0011 東京都文京区千石 4-46-10
発行所 株式会社 コロナ社
CORONA PUBLISHING CO., LTD.
Tokyo Japan

振替00140-8-14844・電話(03)3941-3131(代)
ホームページ <http://www.coronasha.co.jp>

ISBN 978-4-339-00916-3 C3055 Printed in Japan

(森岡)



<出版者著作権管理機構 委託出版物>

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつど事前に、出版者著作権管理機構（電話 03-3513-6969, FAX 03-3513-6979, e-mail: info@jcopy.or.jp）の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。