

デジタル補聴器のための臨界帯域圧縮アルゴリズム

安 啓一 菱谷 誠人 荒井 隆行 村原雄二

上智大学理工学部電気電子工学科荒井研究室 〒102-8554 東京都千代田区紀尾井町 7-1

E-mail: {k-yasu, arai}@sophia.ac.jp, {mhishi99, mura}@splab.ee.sophia.ac.jp

あらまし 聴覚障害者の聴覚フィルタは、健聴者のものと比べてその帯域が広がっていることが報告されている。この特性を考慮し、臨界帯域ごとに帯域分割した信号の包絡成分を、周波数軸に沿って各帯域の中心周波数方向に圧縮するという新しいアルゴリズムを提案した。このアルゴリズムを実現するため、フィルタバンクを用いたFB法とFFT法という2つの手法を用いた。2種類の聴取実験を4名の聴覚障害者に対して行った。実験1として、0%（圧縮なし）から90%圧縮まで10%刻みの処理音声について、明瞭度テストを行った。実験2として、原音に対する20%、40%、60%、80%の圧縮率の処理音声の音質評価をMOS (Mean Opinion Score) 評価によって行った。実験1, 2より、被験者によって最適な条件(FB法/FFT法, 圧縮率)は異なるものの、提案法による音質・明瞭度に改善がみられた。

キーワード デジタル補聴器, 臨界帯域, 周波数圧縮, 聴覚障害

Critical-band compression algorithms for digital hearing aids

Keiichi YASU Masato HISHITANI Takayuki ARAI and Yuji MURAHARA

Dept. of Electrical & Electronics Eng., Sophia University 7-1, Kioi-cho, Chiyoda-ku, Tokyo, 102-8554 Japan

E-mail: {k-yasu, arai}@sophia.ac.jp, {mhishi99, mura}@splab.ee.sophia.ac.jp

Abstract In this paper, a novel algorithm for digital hearing aids was proposed, which considering the feature of the wider auditory filter of hearing impaired people. We have focused on this wider auditory filter shape and developed a method in which the critical band is compressed along the frequency axis. This algorithm is implemented by two approaches: the Filterbank (FB)-based approach and the FFT-based approach. To evaluate this algorithm, two experiments were conducted: articulation test for Experiment 1, and MOS evaluation for Experiment 2. As the results of the experiments, these results show the feasibility of frequency compression algorithm for the hearing impaired people.

Keyword digital hearing aids, critical-band, frequency compression, hearing impaired

1. はじめに

聴覚障害者のための補聴器の開発は近年目覚ましい成長を遂げている。補聴器のためのアルゴリズムの研究も多岐に進められ、実用化されている手法も多い。障害者自身が様々な種類の機能を搭載した補聴器を選ぶ時代になったといっても過言ではない。一般的に補聴器は雑音がない環境下ではとても高い了解度と音質を示す。しかし日常生活を営む程度の雑音環境下や、ホールなどの残響下での使用を前提とした場合、途端に性能が劣化する。雑音除去のアルゴリズムとして適応フィルタによる方法[1]やノイズサブトラクション法[2]が用いられている。音声の定常部抑圧の手法では残響下で聴覚障害者の音声明瞭度に改善が見られてい

る[3]。また、聴覚障害者の聞き取る音圧の感度が急激に変化する補充現象の低減のためのAGC(Automatic Gain Control)法[4]、話者位置特定のためのマルチマイクロフォンによるビームフォーミング[5]などの技術が実際に補聴器に応用されている。聴覚フィルタ論に基づいた研究としては、入力音声を帯域通過フィルタにより18帯域に分け、偶数番目の帯域信号のみをすべて足し合わせて右耳に、奇数番目の帯域信号の和を左耳に提示する手法が報告されている[6][7]。実験の結果、健聴者、聴覚障害者のどちらにも明瞭度に改善が見られた。しかしこの手法では右耳と左耳の聴覚特性が同じでなければならないなどという弱点がある。同じく周波数軸上で信号を操作する研究としては、入力信号

の広帯域スペクトルを低周波域に圧縮するという手法がある[8]。この手法は、高音漸傾型や高音急墜形などの高域の聴力が低い聴覚障害者が、残存する低域の聴力を活用することを目的としている。この実験の結果では、改善を示した被験者と明瞭度の減少を示した被験者がいたが、平均としては改善が見られなかった。

本論文では、聴覚障害者の聴覚フィルタの形状は健聴者のものと比べて広がっている[9]という特徴に着目し、近い周波数同士の選択性(周波数選択性)を改善する新しい手法を提案する。本手法において入力音声は臨界帯域を考慮した帯域数に分割される[10]。分割された各帯域において、スペクトル包絡を臨界帯域ごとの中心周波数方向に圧縮されるという手法である。本論文ではこの手法を臨界帯域圧縮と呼ぶ[11][12]。Fig. 1 に本手法の概略を示す。また、Fig. 2 に入力された音声と処理音声のスペクトログラムを示す。

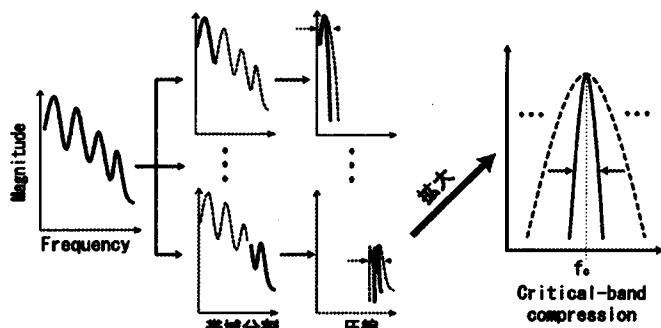


Fig. 1 臨界帯域圧縮の概念図。入力音声は臨界帯域に沿って帯域分割された後、各臨界帯域の中心周波数方向に情報が圧縮される。

2. 臨界帯域圧縮のためのアルゴリズム

臨界帯域圧縮のアルゴリズムを実現するために、二つの手法を提案する。

2.1. FB (Filter Bank) 法

まず、入力音声を臨界帯域の中心周波数に従い帯域分割を行う[10](標本化周波数 16000Hz のときは 21 帯域, 11025Hz のときには 19 帯域)。このフィルタバンクはカイザー窓による FIR フィルタである。信号を $x_{org}[k, n]$ ($k = 1, 2, \dots, 21$) と置き、この信号からヒルベルト包絡 $E_{org}[k, n]$ を求める。一方同じ入力信号に対し帯域幅を圧縮率に応じて狭めた信号を $x_{cmp}[k, n]$ 、その狭帯域信号から求めたヒルベルト包絡

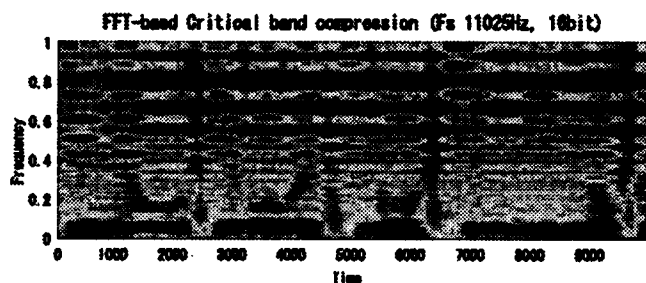
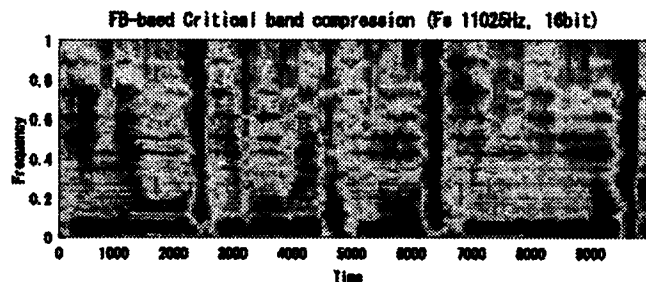
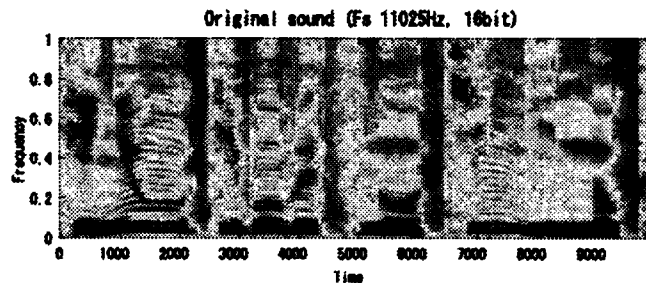


Fig.2 臨界帯域圧縮のスペクトログラム。入力音声(上段)に対してFB法により80%圧縮した音声(中段)、FFT法により80%圧縮した音声(下段)。臨界帯域の中心周波数に情報が圧縮されているのがわかる。

を $E_{cmp}[k, n]$ とすると、キャリア成分 $\cos\theta_{cmp}[k, n]$ は
$$\cos\theta_{cmp}[k, n] = x_{cmp}[k, n] / E_{cmp}[k, n] \dots \textcircled{1}$$

と表すことが出来る。音声の重要な情報の大部分は包絡成分に含まれているので、キャリア成分を圧縮しても音質にはあまり影響は無い。原音のヒルベルト包絡 $E_{org}[k, n]$ と狭帯域のキャリア成分 $\cos\theta_{cmp}[k, n]$ を掛け合わせることで k 番目の帯域における信号 $y[k, n]$ を求めると

$$y[k, n] = E_{org}[k, n] \cos\theta_{cmp}[k, n] \dots \textcircled{2}$$

となる。最後に全帯域にわたって帯域ごとの信号 $y[k, n]$ を足し合わせることで最終的な出力信号を得る。周波数圧縮率 $R\%$ で圧縮された音声は、各臨界帯域で原音の帯域幅に対し、 $(100 - R)\%$ の帯域幅に周波数圧縮される。Fig. 3 は FB 法のブロックダイアグラムである。

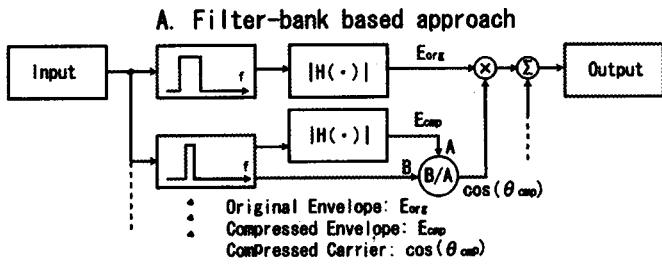


Fig. 3 フィルタバンク法のブロックダイアグラム (21帯域に分割された中の一つの帯域を示す)

2.2. FFT 法

次に二つ目の手法、FFT 法について説明する。入力信号は窓長 32ms、フレームシフト 8ms のハミング窓を通過後、FFT 処理され周波数領域に移される。そして位相成分と振幅成分に分けられ、振幅成分は臨界帯域の各中心周波数方向に圧縮され、圧縮後に成分のなくなった部分には 0 が代入される。圧縮された振幅成分と元の位相成分を掛け合わせ、IFFT により時間領域に戻される。フレームシフトされた信号は OLA(Overlap add)を施され出力される。周波数圧縮率 $R\%$ で圧縮された音声は、各臨界帯域で原音の帯域幅に対し、 $(100 - R)\%$ の帯域幅に周波数圧縮される。Fig. 4 に FFT 法のブロックダイアグラムを示す。

B. FFT-based approach

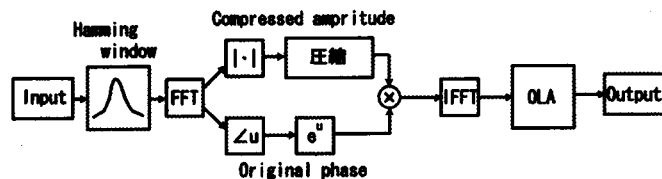


Fig. 4 FFT 法のブロックダイアグラム。FFT 後の処理の上段が振幅、下段がキャリアである。振幅の圧縮は周波数軸方向に沿って臨界帯域ごとに行われる。

3. 実験と結果

上記のアルゴリズムを検証するために聴覚障害者を対象に聴取実験を行った。実験には 4 人の聴覚障害者が参加し、いずれも WHO が定めるところの重度難聴(90dB 以上)であった。実験 1 として明瞭度テスト、実験 2 として音質の MOS 評価を行った。実験は防音室内で行われ、被験者は補聴器をつけた状態で実験に参加した。刺激音提示にはスピーカ(BOSE Speaker 402)およびアンプ(SONY V777ES)が用いられ、提示する刺激音は実験前に各被験者に聞きやすい適当な音量に合わせられた。

3.1. 実験 1 ～単語明瞭度テスト～

刺激音として、日本人男性による「私の名前は ___ ___ です」という文章を提示した。___ ___ の部分には VCV の無意味音節が入り、V は /a/ で統一したのに対し、C の部分は 14 の子音のうちから 1 つずつを割り当てた。PC 画面上に各子音に対応した 14 個のボタンが表示され、2 度繰り返し聞いた後に強制選択させた。圧縮率は 0% から 90% (原音を 0% とする) までの 10 種類、計 140 種類の刺激音がランダムに提示された。実験は FB 法、FFT 法の 2 通りについて行った。

Fig. 5 は FB 法の処理音を提示したときの圧縮率に対する正解率の変化を示している。Fig. 6 は FFT 法の場合である。

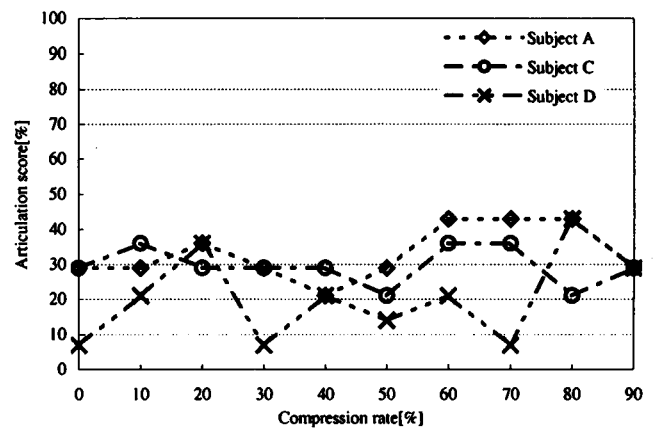


Fig. 5 各被験者の周波数圧縮率の変化に対する明瞭度の変化 (FB 法)。原音の圧縮率を 0% とした。

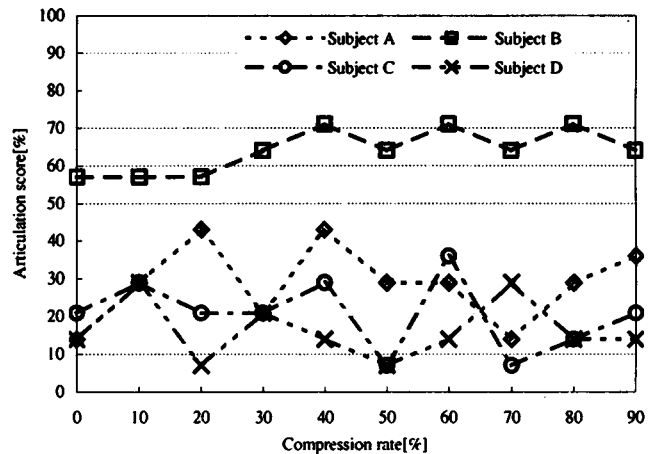


Fig. 6 各被験者の周波数圧縮率に対する明瞭度の変化 (FFT 法)。原音の圧縮率を 0% とした。

3.2. 実験2 ～MOS評価～

刺激音は NTT-AT「音素バランス 1000 文音声データ CD-ROM vol. 1」コーパスより 6 文(男性 3 文, 女性 3 文)をそれぞれ 20%, 40%, 60%, 80% の圧縮率で圧縮したものを提示した(原音の圧縮率は 0% とした). PC の画面上には, 提示される文と「原音」「処理音」という二つのボタンが表示され, 被験者は繰り返し刺激音を聞けるようにした. 「明瞭性」「音質」の二つの項目について「原音」の評価を「3」とした場合の「処理音」の評価を 1-5 の 5 段階 MOS 評価で行った. 刺激音は合計 48 種類(6 文*4 圧縮率*2 回)ランダムで提示された. 実験 1 と同じく FB 法, FFT 法それぞれについて実験を行った. FB 法で処理した音声について, 各被験者の圧縮率の変化に対する MOS 評価を Fig. 8 に示す. また, FFT 法で処理した音声について, 被験者ごとの圧縮率の変化に対する MOS 評価は Fig. 9 に示す.

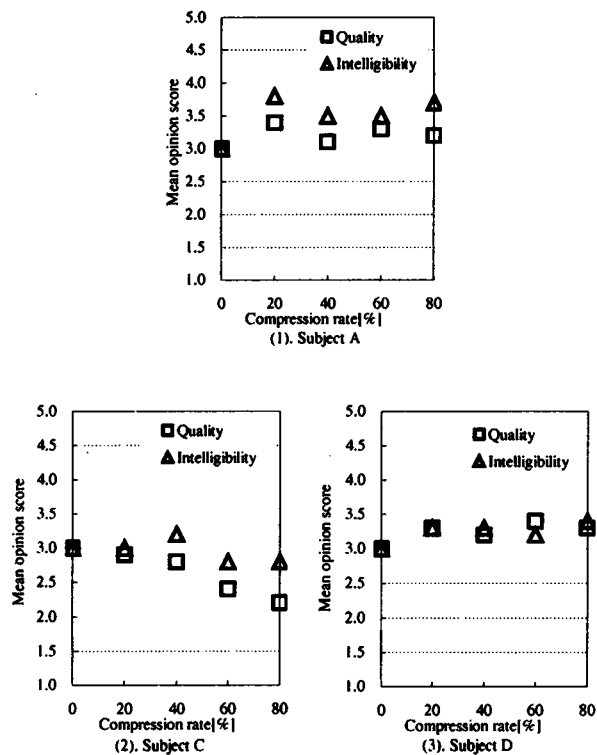


Fig. 8 FB 法で処理した音声の圧縮率の変化に対する各被験者ごとの MOS 評価

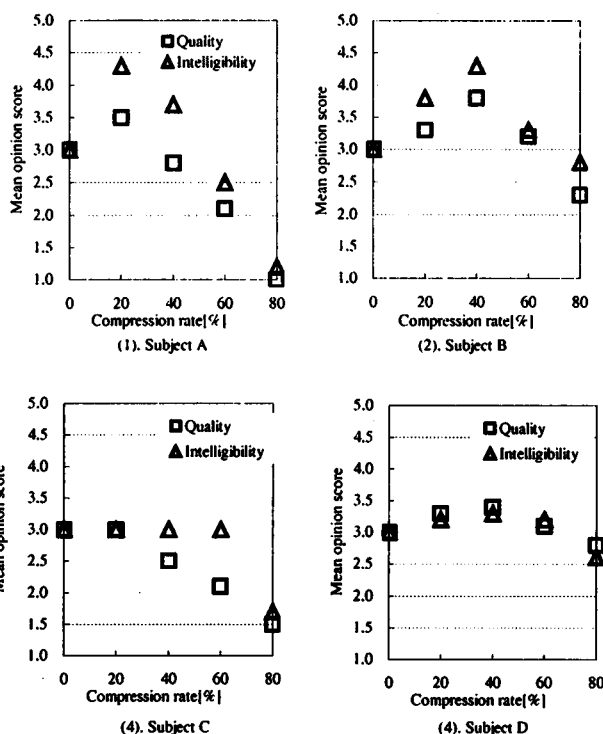


Fig. 9 FFT 法で処理した音声の圧縮率の変化に対する各被験者ごとの MOS 評価

4. 検討と考察

4.1. 単音節明瞭度テストについて

被験者内において FB 法と FFT 法の結果の間に相関はみられない. しかし全ての被験者において, 周波数圧縮後の音声の方が原音に比べて改善が見られる圧縮率が存在する. つまり, 臨界帯域圧縮の有効性が確認された.

4.2. MOS 評価について

Fig. 8, 9 より, 被験者 C 以外については原音よりも処理後の音声の方が明瞭性, 音質について, 改善が見られる点がある. FB 法と FFT 法のスコアを比べてみると FFT 法の処理の結果はスコアの山がはっきりと出ている.

また, 先行研究では圧縮率を FB 法では 30%, FFT 法では 40% に固定していたために被験者の MOS 評価には大きな差が出てしまった [11] が, 今回, 様々な圧縮率の処理音を作ることによって各被験者に適当な圧縮率を求めることが出来た. 具体的には被験者 A については, 圧縮率 20% の時, 被験者 B では FFT 法により 40% 圧縮した時, また, 被験者 D では FB 法を用いたときに 80% の圧縮率時に改善が見られた.

5. 結論

本研究では聴覚障害者の聴覚特性を考慮した周波数圧縮のアルゴリズムを提案し、実際の聴覚障害者に対して聴取実験を行ったところ、被験者にとって最適な条件(FB法/FFT法, 圧縮率)が異なるものの、提案法による改善が見られた。

これからの課題としては DSP による実時間処理、音声信号のスペクトル形状に合わせて分割する帯域をダイナミックに変化させる手法や、広帯域での周波数圧縮[8]と本論文の手法を組み合わせる手法に取り組みたい。

文 献

- [1] N. Magotra, P. Kasthuri, Y. Yang, R. Whitman, F. Livingston, Multichannel adaptive noise reduction in digital hearing aids, Proc. of IEEE Int. Symp. Circuits Syst., Vol. 1998, No. 6, IV, pp.582-585, 1988.
- [2] H. Levitt, N. A. Neuman, R. Mills, and T. Schwander, A digital master hearing aid, J. Rehabil. Res. Dev. 23, pp.79-87, 1986.
- [3] A. Kusumoto, T. Arai, T. Kitamura, M. Takahashi, and Y. Murahara, Modulation filtering of speech as a preprocessing against reverberation for the hearing-impaired, Proc. of ICASSP, II, pp.853-856, 2000.
- [4] B. C. J. Moore, and B. R. Glasberg, A compression of four methods implementing automatic gain control (AGC) in hearing aids, Brit. J. Audiol. 22, pp.93-104, 1988.
- [5] B. Window, A microphone array for hearing aids, IEEE Circuit and Systems Magazine, Vol.1, pp.26-32, 2001.
- [6] D.S. Chaudhari, and P.C. Pandey, Critical band splitting of speech signal for reducing the effect of spectral masking in bilateral sensorineural hearing impairment, Proc. of ISSPA, pp.22-25, 1999.
- [7] D.S. Chaudhari, and P.C. Pandey, Dichotic presentation of speech signal with critical band filtering for improving speech perception, Proc. of IEEE ICASSP, Vol.6, pp.3601-3604, 1998.
- [8] 坂本真一, 館野誠, 大友晋一, “難聴者における周波数圧縮音声の了解度特性,” 日本音響学会, vol.55, no.6, pp.424-434, 1999.
- [9] B. R. Glasberg, and B. C. J. Moore, Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments, J. Acoust. Soc. Am., 79, 1020-1033, 1986.
- [10] E. Zwicker, G. Flottorp, and S. S. Stevens, Critical band width in loudness summation, J. Acoust. Soc. Am., 3, 548-557, 1957.
- [11] K. Yasu, K. Kobayashi, K. Shinohara, M. Hishitani, T. Arai, and Y. Murahara, Critical-band compression method for digital hearing aids, Proc. of Forum Acusticum, Sevilla, 2002.
- [12] K. Yasu, K. Kobayashi, K. Shinohara, M. Hishitani, T. Arai, and Y. Murahara, Frequency Compression of Critical band for Digital Hearing Aids, Proc. of JCA2002, Nanjin, China, 2002.