

聴覚特性に基づく高齢者に向け音声強調

Speech enhancement for Elderly People to based on auditory property

田母神 恒
Tamogami Wataru
法政大学情報科学部デジタルメディア学科
E-mail:wataru.tamogami.p2@cis.hosei.ac.jp

Abstract

Now Japan is old people society that is increased old people. An old person has problem daily life bring by decrease hearing skill with aging. An old person is different of sound volume and range than healthy people, that loss audible frequency range and increase threshold of audibility than normal threshold of audibility, especially high-tone range is prominence. Original sounds run noise rejection used by modulation spectrum, that more easy hearing of sound portion. However this operation ingenerate musical noise, bring a discomfort feeling by heard this sound. At the same time, combine of original sounds and after sounds, which create to stifle musical noise and emphatic sound portion. This signal adds volume control of equal-loudness curve, which made bell-clear voice. Experimental material was used TV sounds (news, variety, sports, Drama and CM) that difference sound volume and contents in noise. Appraisal method, many people compared original sound with after sound, which conduct subjective appraisal Finally, hearing checked there, loudness + modulation spectrum better than other.

1 はじめに

音声強調は難聴者支援や音声認識など、多くの技術分野において要求は極めて強い。食生活や医療技術の水準が上がり平均寿命は増加しているため、現在の日本は高齢者が増加し高齢化社会となっている。そして人間は歳を取ると共に聴覚能力が低下することが明らかになっており、これは一般的には耳が遠くなると言われ、加齢により聴力が低下する。20代前後をピークに聴力が低下し始め、高周波数帯から聞き取りづらくなっていく。高齢者は若年者と比べ聞こえに必要な音圧レベル(dB)が高く、ラウドネスの傾斜が高いので少し音を大きくすると、とても大きくなったように感じられ、嫌悪感やストレスを感じる(リクルートメント現象)。さらに、高齢者は残響下において大幅な文章了解度の低下を示したという結果が出ており音声を聞き取りづらい。このため、日常生活においても、騒音の多い場所でのアナウンスやBGM音量の大きいTV等の音が聞き取りづらくなり、支障をきたすことが容易に考えられる。加齢につれてこのような症状の発症率が高まる為、必要音素の強調と背景音抑圧の行える音声強調の重要性は高まると考えられる。しかし、聴力低下には個人差が大きく個人に対応した処理を施さなければ十分な効果は得られない。

本研究は高齢者に健常者と同様に音を聞こえるようにすることを目標とした。目標を達成する為に、高齢者と健常者の聞こえの違いを把握する必要がある。個人差に対応する必要がある。そのため、聴覚特性に着目し聴覚特性の差を用いることで聞こえの違いに対応し、聴覚特性を求める聴力テストを簡単化することで個人別に測定するコストの軽減を目指す。

2 加齢による聴覚の低下

人が聞こえる可聴音域は20Hzから20,000Hzであるといわれている。人間の声は、そのうちの400Hzから4000Hzの帯域内に重要な音情報が含まれている。つまり、人の声の明瞭性を確保するためには、この重要な帯域成分をしっかりと再生し、相手に伝える必要がある。聴力のレベルは、オーディオメーターを用いて測定するが、正常聴力の場合は、0dB近辺であり、難聴の程度が強くなるほどこの値が大きくなる。通常30dB以上が軽度難聴、50dB以上が中度難聴、70dB以上が高度難聴であるとされている。高齢者では、加齢とともに高周波の音が聞こえにくくなる。もちろん個人差があるが、70歳の高齢者では、2,000Hzで約30dB、4,000Hzで約40dB、8,000Hzでは約60dBもの聴力損失がみられるという報告もある。このような年齢に関連した聴力低下は老人性難聴とよばれている。聴覚の鋭敏性は聴神経の変化が原因となり、50歳位から少しずつ低下し始め、65歳以上では約30%が一定の聴力障害を起こしているといわれている。加齢による聴覚の低下は1. 内耳 2. 脳幹 3. 視覚野 と順番に低下していき、感覚系が先に低下し次に中枢系が低下していく。

2.1 外耳

- 外耳道の崩れや耳垢の増加による純音に対する感受性が低下
- 耳に入ってくる音声信号が外耳道の形の崩れや耳垢の増加で正常に聞こえにくくなり、余分な反響などを起こす。

2.2 中耳

- 鼓室輪・中耳小骨の石灰化により信号振幅の減衰・絶対的感受性の低下
- 鼓室輪は鼓膜の周りを覆っている部分でこれが硬質化してしまい鼓膜が信号振幅を受け取りづらくなっている。
- 中耳小骨はつち骨・きぬた骨・あぶみ骨の3つの骨で構成され、その鼓膜から音を増幅し内耳へと伝える部分(通常では音圧を約22倍へと増幅する)であるがこの部分も硬質化が進み信号の増幅が小さくなってしまふ。そのため、音を伝える際に小さな音を聞き取れなくなる。

2.3 内耳

- 単一の神経細胞の振幅の閾値が増加
音として感知できる最低音が大きくなる。特に規定部分で顕著に現れ、高周波数成分が高齢者は聞きづらい原因となっている。
- 外毛細胞の減少による蝸牛マイクロホンの低下
外毛細胞が音波を受け取り蝸牛内で信号が強調されるが信号を受け取る外毛細胞自体が減少してしまうことで強調が行えず周波数の理解が出来なくなる。

一般的に外耳・中耳が障害された場合を【伝音性難聴】といい、内耳以降が障害された場合を【感音性難聴】という。伝音性

難聴は絶対的な音量が聞き取りづらくなるだけなので医学的な治療や補聴器を使用することで改善を図れるが、感音性難聴は医学的な治療で改善しにくく補聴器を使用して聞き取りの訓練をしていく必要がある。

3 高齢者が聴覚において自覚する問題

1. 背景に雑音がある場合の聞き取り
従来研究で多目的ホール、駅構内、空港などの広い空間で音声電気が拡声される際、室内のスピーカから放射された音声には同時に長い残響が掛かる。このような残響環境下ではしばしば音声の聞きとりが低下するが、特に聴覚障害者・高齢者・非母語話者ではその影響は大きいということが分かっている。また、周波数分解能が低下し多くの音から必要な音を取り出す能力が低下しているためにこのようなことが起こる。
2. ひずんだ音声の聞き取り
高齢者は通常で聞き取りを行う際に必要な音圧レベルが高い。そのうえ、歪んだ音声では各周波数ごとの聞き取りやすい周波数帯にも影響が出てしま音が響いてしま音が聞こえないわけでないが何を言っているかの理解が出来ない。
3. 通常の音声の聞き取り
高齢者は特定の周波数帯で音圧レベルが低いときはとても低く感じ、あるレベル(50dB程度)で急激に大きく感じるようになり、更に高いレベル(80dB程度以上)になると若年者と同程度の大きさに回復するというリクルートメント現象が現れている。また、高齢者は両耳の間での注意の切り替えに時間がかかることも聞き取り障害の原因と考えられる。
4. 高い音の聞き取り
年齢別の各周波数ごとの聴力障害発生率が以下の図1で、どの年代も高音域になるにつれて障害が発生しやすくなっている。

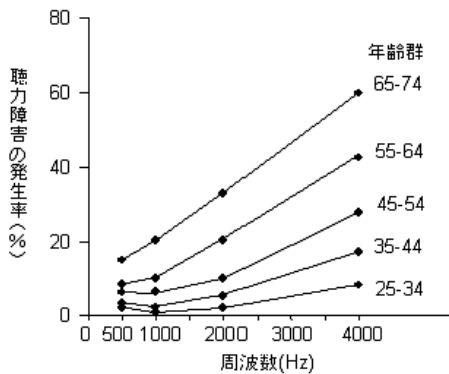


図1. 障害発生率

また、図2より性別によっては男性の方が女性より聴力消失が大きいですが、加齢による比率のほうが高く性別による差は誤差と考えられ、全ての人に加齢による聴覚障害が発生する可能性がある。

5. 時間分解能
短時間で多くの音声が入ってきたとしても高齢者は区切り(二つの音の休止)を若年者より長く取らないと知覚することが出来ず、前後の単語が融合してしまう。よく高齢者と会話をしている通常通り話していても速くて何を言っているかわからないと返されるのはこの為である。

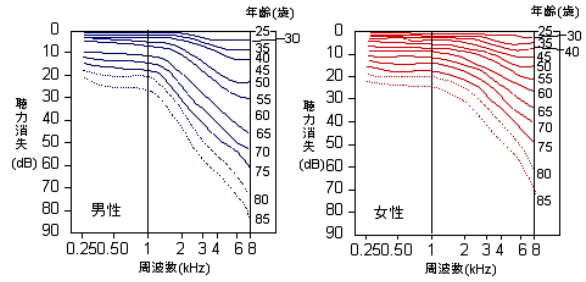


図2. 聴力減衰率

4 聴力特性に基づく音声強調

前章の1~5を考慮すると通常の音声のまま音声強調しても高齢者には聞き取りづらい部分が発生する恐れが高い。本研究では前章の1~4の問題を改善することを目的とする。聴覚特性を簡易推定を行う為にお夏レベルの基準音を作成・作成した音を用いて聴力テストにより最小可聴域を測定する。元音源から音声を抽出し、その抽出した信号を窓幅512のハンギング窓を用いてスペクトルにし、聴力テストによって計測した各個人のラウドネス曲線と健常者のラウドネス曲線の差分を補完する。

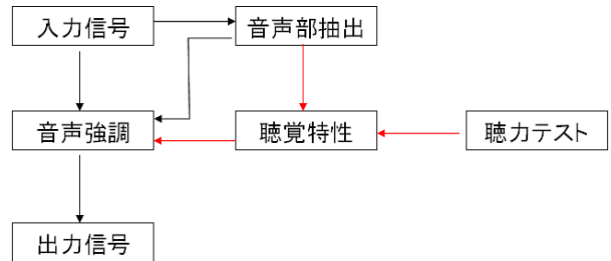


図3. 提案手法構造図

4.1 ラウドネス曲線

人間の聴覚で感じる、音の大きさの感覚をラウドネスと呼ぶ。1000Hzの音圧(dB)を基準とし、phon(ホン)という単位を用いる。基準の1000HzにおいてはdBとphonの値は等しくなる。音の周波数を変化させたときに等しいラウドネスになる音圧レベルを測定し、等高線として結んだものである。同ラウドネス曲線上の各周波数ごとの聞こえる音の大きさは等しく聞こえる。端になるにつれて聞き取る為の必要音圧が上昇している。これにより人は低域と高域に対しての感度が悪く、中域の感度が高いのがわかる。これは健常者に対するデータであり、高齢者のラウドネス曲線は健常者と異なる形になるのは容易に考えられる。

4.2 聴力検査

健常者のラウドネス曲線はあるが、高齢者のラウドネスは形状は異なることは容易に考えられるが定義などはされていない。これは、聞こえ方にも個人差があり決まった形をとらず、難聴の具合などで変動が激しい為である。上記のラウドネス曲線は健常者の平均によって生成されている。実際に高齢者の方はそのように聞こえているのかを調べる為に聴力テストを行った。

4.2.1 検査音測定・作成

音量を定義するものにdB(デシベル)があるがdBにも特性があり、人の聴覚に対応したA特性、比較的平坦な周波数特性を持っているC特性、C特性より更に広い周波数範囲にわたり平坦な特性を持っているZ特性(あるいはF特性)の3つがある。A特性は特に騒音計などの騒音計測に用いられ中域が大きく低域・高域は小さく重み付けされている。C特性は衝撃音(周波数帯で見ると幅が広がるもの)の測定に用いられている。Z特性は重み付けが一切なくFLATであり、音響センサとして用いられている。通常聴力測定で使われるオーディオメータの基準

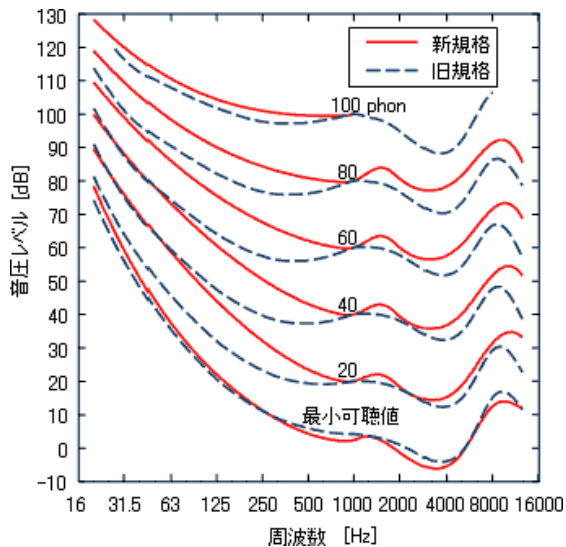


図 4. ラウドネス図 [11]

音が dB SPL (音圧レベル) での定義をされている為、音圧の測定には騒音計測器を用いて周波数に重み付けをしない Z 特性での録音を行った。

- 測定環境
サンプリング周波数は 48kHz・量子化ビット数は 16 ビット、計測音の時間は 3 秒間。聴力検査に用いた音は各 125・250・500・1000・2000・4000・8000Hz の sin 波を使用し、音量は dB SPL で統一。dB 定義は HL を使用しオーディオメータと同様にした。
- 使用機器
logicool A-370P
小野測器 LA-5560
MathWorks MATLAB

検査に使用した音源はすべて MATLAB により生成を行い、音漏れ等ないように音響カブラを追加で作成し取り付け付けた。人工耳は型紙を一般的な外耳道のサイズの直径及び長さにし内側にフェルトシールを貼ってある。音響カブラも同様にヘッドホンの形に合わせ型紙で作成している。

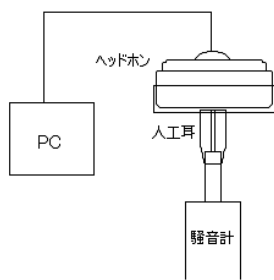


図 5. 定義音測定装置

測定の際、通常状態で約 40 ~ 50dB の音を検出してしまふ為、音量の大きいものしか測定は出来なかった。dB は音の強さを図す単位ではなく基準の音に対する比較単位で

$$10 \log(p_b/p_a)$$

の式によって算出される。はじめに計測した dB SPL は音圧レベルで音の強さの二乗になる為、算出式は

$$10 \log(p_b/p_a)^2 = 20 \log(p_b/p_a)$$

となる。計測した値から逆算し計測の出来なかった dB を作成した。測定できた dB のパワーの比較を行ったところ dB のパワーの差がきわめて値が大きく比較しづらいため、底を 10 とし

た対数をとった。その結果 dB の対数パワーは図 6 の通り直線状になっていた。

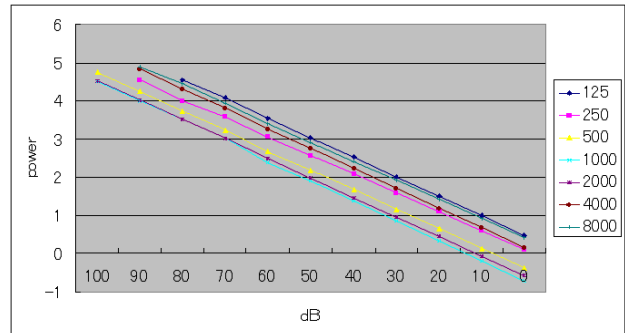


図 6. デシベルごとの対数パワー

作成した音を用いて PC とヘッドホンで聴力テストを行った。今回計測したのは最小可聴域で作成した 125 ~ 8000Hz の 7 つを 5dB 刻みで 0 ~ 100dB の音量で小さい音量から聞いてもらう。3 度流し聞こえない場合は次の大きさに移行する。聞こえた場合は次の周波数に移行する。

4.2.2 比較

作成した音を用いて聴力テストを行った。計測した高齢者と健常者のオーディオグラムから推測した最小可聴閾域は図 7 のようになった。

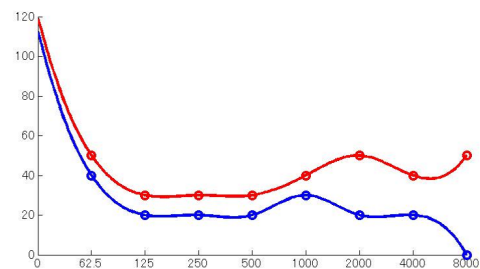


図 7. 赤：20 代男性 青：60 代男性

その結果、予想通り健常者と比較して高音域の可聴閾値が上がっていた。特に高音域では差が大きく生まれている。逆に低域では目立った差は見受けられず、健常者と同等か多少悪い程度となった。

4.3 聴覚特性補完

等ラウドネス曲線は 1000Hz の音を基準とし、それと同等の音量に聞こえる。高齢者のリクルートメント現象に対応する為に、各周波数をラウドネス曲線の差分で補正を行い健常者と同等に聞けるようにする。聴力の衰えには共通点もあるが個人差やばらつきが大きい為、各個人に対応したラウドネスを準備するのが最も望ましい。今回は健常者 3 人の聴覚特性の平均と高齢者の聴覚特性の差分 (図 8) を使用している。

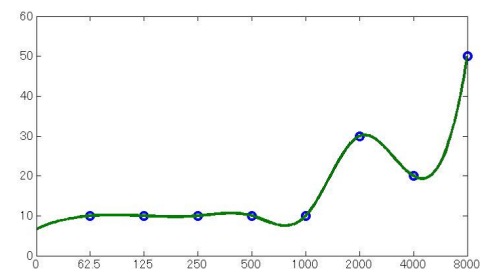


図 8. 聴覚特性の差分

計測した最小可聴域を元に等ラウドネス曲線を作成する。

$$S = a(p^{2\alpha} - p_t^{(2\alpha)})$$

S はラウドネス, p は純音の音圧, p_t は最小可聴域 (音圧出現), α はラウドネスのベキ指数で, a は定数である。この式は Zwislocki and Hellman 及び Lochner and Burger によって提案されたラウドネス関数である。これにより得られたラウドネスを 256 点で生成し, ランニングスペクトルに掛け合わせていき聴力特性に対応した元音源が作成される。あまりに低・高音域の音は高齢者だけでなく健常者も聞き取ることが出来ない為, ラウドネス曲線に重みを掛けて 500~2000Hz を 1 とし端に行くにつれて 2 次曲線的に重みを減らしている。

4.4 変調スペクトル

本研究では音性抽出に変調スペクトルを用いた。変調スペクトルとは全ての音声信号を短時間の窓関数を用いて切り出し, 音声信号全体を短時間信号にする。その短時間信号をスペクトル分析を行いスペクトルにし, 時系列に並べたランニングスペクトルを再度スペクトル分析を行うこと変調スペクトルが得られ, これはランニングスペクトルの (時間・周波数の 2 次元データ) であったことと異なり, (周波数・変調周波数) の 2 次元データとなっている。

変調スペクトル上で雑音なし音声は約 10Hz の変調周波数帯に成分が収束するのに対し, 雑音付き音声は 1Hz 以下に成分が集中する。これは音声成分と比べ雑音は変化が小さく定常である為, 低域に集中している。これを用いて音声に集中した成分を取り出せる。これに聴覚特性補完を行った音源と合成させることでもとの音源の減衰させたい周波数を減衰させることができる。変調スペクトルは BGM・ざわめき・野外騒音に効果がよく, 音声を歪まない程度に抽出を行うので今回の手法に使用した。

4.5 実験

実験に使用した音源は TV 音源を使用し, 内容としてはニュース・バラエティ・スポーツ・CM を使用した。

テレビ音源の録音は iodata 製 TV キャプチャボード「GV-MVP/GX2」を使用し, サンプリング周波数 48kHz, 量子化ビット数 16bit にて録画し映像と分離させた。

人の声は高くても 6kHz 以上にはならない為, 44.1kHz のサンプリング周波数から 16kHz にダウンサンプリングし処理を加えていた。

音声信号を変調スペクトルを用いて抽出を行い, 抽出した音声に聴力テストにより計測した聴覚特性の差分型に音声を強調する。処理後の音声を元音源に加え音声を強調する。これにより BGM や雑音を抑え音声を強調できる。

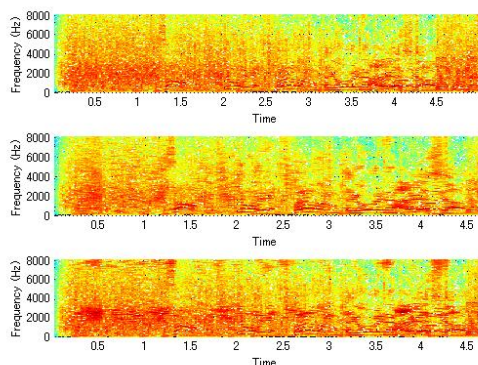


図 9. スペクトル比較 元音源はバラエティ番組。時間は約 15 秒

5 評価

作成した信号を評価する為に

a : 元音源

b : 従来の音性強調 (変調スペクトルを用いた音性強調)

c : 提案手法

の主観評価を行った。評価の仕方は 3 つに対し, 1 : 音声の聞きやすさ 2 : 音の聞きにくさで順位を付ける。このとき聞き逃しや再度繰り返し聞くことも可能とした。評価者は健常者 3 人と高齢者 3 人の計 6 人である。元音源を 3 とし, 5 段階で評価。(向上していれば 5) 補完に使用した差分は高齢者一人と健常者 3 人の平均の差分を使用した。

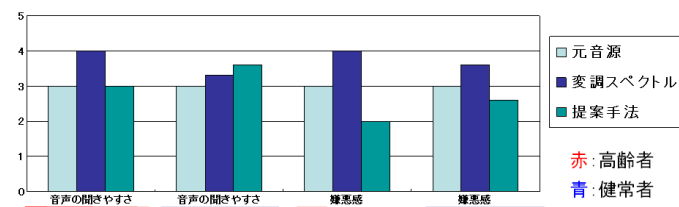


図 10. 主観評価結果

6 おわりに

音性聞き取りの結果としては変調スペクトルを用いた音声強調が最も聞きとり評価が高く, 提案手法は 2 番目となった。協力してもらった高齢者からは提案手法に対し「音はクリアになっているが音声が遠く感じ聞き取りにくい」、変調スペクトルを用いた音声強調には「音が若干歪んでいるが音声が近く感じ聞き取りやすい」といった意見があった。聞きにくさの結果については提案手法が最も評価が高くなった。元音源と比較するとどちらも向上し, 変調スペクトルを用いた音声強調と異なる点から聞き取りに対して改善は行われていると判断できる。

今後の課題としては聴覚を今回は手動にて計測を行ったが高齢者が一人で検査でき, そのまま差分計算までを GUI で簡単にできるようにする必要があり, 実際に想定している環境はテレビやスピーカーなどの為, 距離による伝達特性の変化を考える必要がある。

参考文献

- [1] 和田 他, “重み付け変調スペクトルを用いた雑音認識” IEICE Technical Report SIS2005-50 (2005-11)
- [2] 石田 他, “ランニングスペクトルを用いた雑音にロバストな音声認識” Tech. Rep. of IEICE DSP2003-36(2003-06)
- [3] ビジュアル生理学 ”http://bunseiri.hp.infoseek.co.jp/”
- [4] 扇長 他 “残響音場における高齢者及び聴覚障害者の文章理解度” TECHNICAL REPORT OF IEICE EA94-75(1994-12)
- [5] 近藤 勉 “よくわかる高齢者の心理” ISBN4-88848-642-5 (2001.4)
- [6] デニス・C. パーク, ノバート シュワルツ, Norbert Schwarz, Denise C. Park, 口ノ町 康夫, 坂田 陽子, 川口 潤 “認知のエイジング 入門編” ISBN978-4-7628-2361-9(4-7628-2361-9) C3011
- [7] 立木 孝・村井 和夫 著 “よくわかるオーディオグラム” 金原出版 (2005)
- [8] 倉片 憲治 “聴覚特性の加齢変化を考慮した音環境設計” 産総研 人間福祉医工学研究部門フォーラム '01/12/04 資料
- [9] J. J. Zwislocki and Hellman and J. F. Burger “On the psychophysical law,” J. Acoust. Soc. Am. 32, 924 (1960).
- [10] J. P. A. Lochner, and J.F. Burger, “Form of the loudness function in the presence of masking noise,” J. Acoust. Soc. Am. 33, 1705-1707 (1961).
- [11] 産総研” http://www.aist.go.jp/aist_j/press_release/pr2003/pr20031022/pr20031022.html ”聴覚の等感曲線の国際規格 ISO226 が全面的に改正に 2003/10/22.