

©阪内 澄宇 牧野 昭二 (NTT ヒューマンインタフェース研究所)

## 1 はじめに

拡声通信システムにおいて、通話品質の劣化要因である音響エコーを防ぐためには、適応フィルタや非線形処理等のエコーキャンセラ装置が不可欠である。これらの装置を設計する際に、エコーをどの程度抑圧すべきかを明らかにすることは、大変重要である。従来評価されてきた所要エコー抑圧量(以下、所要量)のほとんどは、フルバンド方式のエコーキャンセラを念頭においた検討[1]であった。そのため、信号を複数の帯域に分割し、それぞれの帯域毎にエコー抑圧処理を行なうサブバンド方式のエコーキャンセラ(以下、SBEC)[2]に、そのまま適用することはできない。

本報告では、SBEC 設計に必要な周波数帯域毎の所要量を明らかにするため、模擬実験システムを用いた主観評価を行なった。また、得られた所要量に基づき、SBEC における適応フィルタのタップ数の割り当てを検討した。

## 2 実験システム

模擬実験システムは、7kHz 帯域を有する 4 線回線構成の対向拡声通信システムを想定した。図 1 にシステムの概略図を示す。

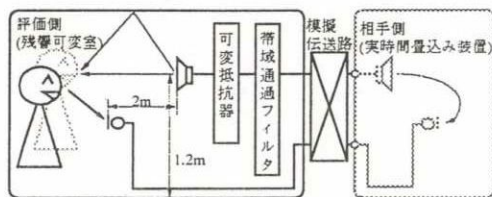


図 1: 模擬実験システムの構成

この評価実験は、評価側で発声した音声は、相手側の室内伝達特性を模擬した実時間量込み装置を通り、再び戻ってくるエコーに対して所要量を決定するものである。ここで、所要量の評価に用いられるエコー信号は、SBEC で使用されるものと同様の帯域通過フィルタを通り、拡声の際にはすでに各帯域に制限されている。この各帯域毎のエコーに対し、可変抵抗器により損失を挿入し、模擬的にエコー抑圧を行ない所要量を決定する。なお、送受話感度は ITU-T 勧告 P.34 に従い設定しており、挿入した損失量が 0dB の状態を所要量 0dB とした。

所要量を決定する評価カテゴリーは、実装置の設計にあたり最も妥当な所要量を決定することを念頭に

\*Consideration on frequency domain echo return loss required for audio teleconference systems  
By Sumitaka Sakauchi and Shoji Makino (NTT Human Interface Laboratories)

置き、「エコーを聞き分けようと注意深く聴くと、幾分残留エコー感がある。」程度とした。このカテゴリーは従来の評価における、検知限と許容限の中間範囲にあたる。

評価パラメータは、設計する SBEC に対応した 32 分割の各帯域(帯域幅 250Hz)、伝送遅延時間および受聴状態の相違(発話・傍聴)である。ここで受聴状態については、デスクトップ会議システム(発話中心)および大規模 TV 会議システム(傍聴を含む)を想定し、それぞれの相違を明らかにすることを目的としている。なお、実時間量込み装置内の音響結合は -2dB とした。

## 3 評価実験方法と結果

### 3.1 実験 1(受聴状態および遅延時間の影響)

はじめに各帯域の所要量が、受聴状態の相違および伝送遅延の大小により、どのように影響を受けるかを評定者 1 人について大まかな傾向を調べた。

評価パラメータの 1 つである受聴状態の相違については、評定者自身が短文を発声し、同時に所要量を評価する場合を「発話状態」とし、疑似口より短文を発声し、評定者は発声せずに所要量を評価する場合を「傍聴状態」とした。

もう一方のパラメータである遅延時間としては、伝送遅延なしと伝送遅延あり(約 200ms)を模擬した。ここで伝送遅延なしの場合においても、SBEC の分割合成フィルタの一巡処理遅延(約 28ms)は加味した。なお、評価側および相手側のエコー経路インパルス応答継続時間は、200ms とした。こうして求めた周波数帯域別の所要量を図 2 に示す。

図の横軸は周波数(Hz)、縦軸は所要量(dB)を表す。図中の実線は伝送遅延が小さい場合、破線は大きい場合の結果を、●および\*は、それぞれ発話、傍聴の結果を示している。

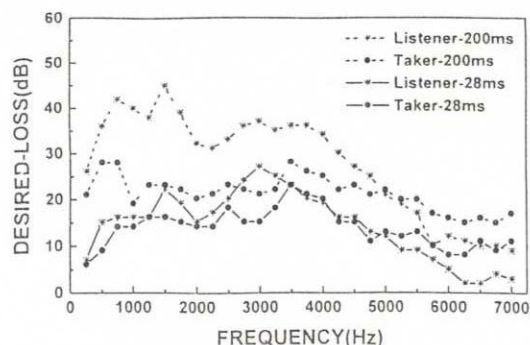


図 2: 受聴状態と遅延時間の所要量に対する影響

これらの実験結果より次の 3 点に分かる。(1) 遅延が小さい場合の所要量は、受聴状態の相違に影響さ

れない。(2) 遅延が大きい場合の所要量は、発話に比べ傍聴の方が大きい。(3) 受聴状態に関わらず、遅延が大きいと低域部の所要量が増加する。

### 3.2 実験2(必要十分な所要量の決定)

3.1の結果から、受聴状態が傍聴で遅延時間大のとき、より大きい所要量が必要であることが確認された。これより実験2においては、拡声通信システムの通信条件を以下のように想定し、さらに評定者数を増やし実際の装置設計に必要な十分な所要量を求めた。

エコー経路インパルス応答継続時間(≒残響時間)は、勧告 G.167 では400msとされているが、実際のTV会議室を想定し300msとした。伝送遅延時間は勧告では250ms以上とされており、また所要量が300ms前後で飽和する[1]ことから300msとした。周囲騒音は、マスクを疑似口からの発話音声に絞り現象を単純化するため、30dBA以下にした。なお、受聴状態はより検知限の厳しい傍聴とした。

評定者は音声研究に携わる25歳から48歳の男女13名である。各帯域に分割したエコーをランダムに提示し、評定者自身に損失を挿入してもらった。得られたデータは信頼区間0.01でt検定を行ない統計処理をした。その結果を図3に示す。

ここで求められた所要量は、実験1の傍聴の結果と周波数軸に対する傾向が等しいことがわかる。ここで、全体的な所要量が増加している理由は、遅延時間が実験1に比べ大きくなっていることによる。

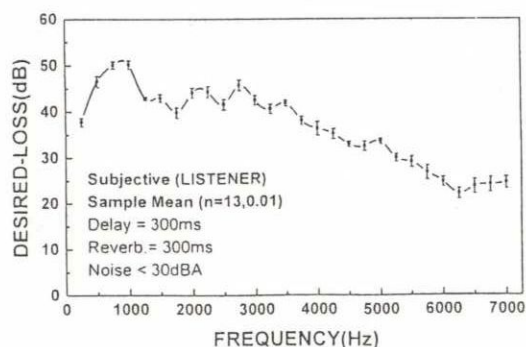


図3: 必要十分な周波数帯域別所要エコー抑圧量

### 3.3 エコーマスクングの遅延による影響

3.1でまとめた(1)(2)は、従来のフルバンドにおける結果[3]と、良く合致している。さらにSBECの設計に最も重要な周波数軸に関連した(3)の結果が、本実験によってはじめて明らかになった。遅延が大きくなるに従い低域部の所要量が増加する現象は、定性的には次のように理解できる。

遅延が小さい時には発話音声とそのエコーがほぼ同時に聞こえるため、エコーは発話音声に周波数軸上で均等にマスクされる。このため可聴感度の良い中域部(2~4kHz)で所要量が極大値を取る。これに対し遅延が大きくなると、発話音声とそのエコーに

時間的なズレが生じる。そのため、エコーのマスクされる確率が小さくなり、音声パワーの周波数分布に従うため、所要量は低域部でより大きくなる。

### 4 適応フィルタのタップ数の検討

3.2で得られた所要量から、SBECの適応フィルタのタップ数を帯域別に求める。タップ数 $L$ は室内のインパルス応答継続時間 $T_{Rf}(s)$ との関係より、

$$L = \frac{1}{60} \cdot \frac{T_{Rf}}{T_S \cdot M} \cdot (\text{Desired Loss}) \quad (1)$$

として求めることができる[4]。ここで、Desired Lossは所要量(dB)、 $T_S$ はサンプリング間隔(s)、 $M$ は間引き数である。

式(1)と図3の結果より求めた各帯域の必要タップ数(最大値により規格化)の一例を図4に示す。ここで式(1)中の各帯域のインパルス応答継続時間には、室容積87m<sup>3</sup>、残響時間300msの実験室での実測値を用いた。

図4よりタップ数は、音声音圧の大きな低域部により多く割り当てる必要があり、中域部で幾分増加するものの、高域部になるに従い割り当てを少なくして良いことが分かった。

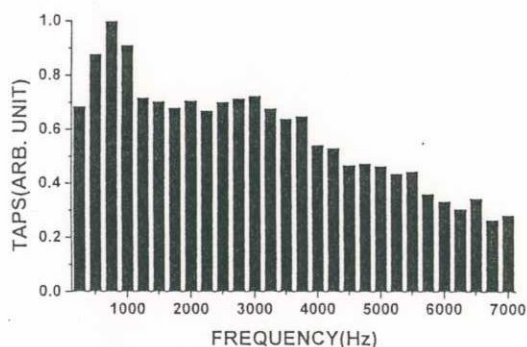


図4: 聴感に基づいたタップ割り当ての一例

### 5 まとめ

SBECの装置設計のため、周波数帯域毎の所要エコー抑圧量を主観評価実験により検討した。まず、受聴状態の相違および遅延時間の大小によって、所要量がどのように影響を受けるのかを確かめた。これをもとに、傍聴かつ伝送遅延が大きい状態での所要量を主観評価により求めた。その結果SBECにおいて、低域部で所要量が増加するように適応フィルタのタップ数の割り当てを行なう必要があることを明らかにした。

謝辞 日頃御指導頂く、当所北脇音声情報研究部部長、小島GL、羽田研究主任に深謝する。

### 参考文献

- [1] N. Kishimoto, et al., *ICC88*, 8.4, June (1988).
- [2] S. Makino, et al., *Proc. ICASSP96*, pp.945-948 (1996).
- [3] H. Yasukawa, et al., *Proc. ICASSP86*, pp.1309-1312 (1986).
- [4] 古川 他, 信学技報 *CS85-101*, (1985).