

双方向同時通話型ハンズフリーインタホン

Full-Duplex Hands-Free Intercom

福島 実* ・ 吉田 恵一*
Minoru Fukushima Keiichi Yoshida

通話機筐体ごとに異なる音響結合特性に依存せずにエコー抑圧量の周波数特性が平坦となるエコーキャンセラと、刻々と変化する音響結合利得およびエコー抑圧量に追従して通話路に挿入する損失量の総和（総損失量）を逐次更新することを特徴とする総損失量適応型ボイススイッチを用いて、高音圧の周囲騒音環境においても短時間で総挿入損失量を低減して双方向同時通話状態を実現するハンズフリー通話アルゴリズムを考案した。

本技術の適用により、業界初の双方向同時通話型ハンズフリーインタホンを実現した。

A new algorithm for hands-free speech communication has been developed for enabling communication in a high noise environment by combining two echo cancellers to obtain a flat frequency characteristic of cancellation, independent from varying coupling characteristics of each terminal, and an adaptive voice switch, which continuously updates the total insertion loss into the communication path according to the varying coupling gain and echo cancellation quantity.

An algorithm for reducing the total insertion loss in a short time in a noisy environment has made the industry's first full duplex hands-free intercom possible.

1. ま え が き

近年、利便性、デザイン性、および省スペース化などの理由から、住宅で用いられるインタホンのハンズフリー化が進んでいる。ハンズフリー通話を可能とするには、機器から離れた位置においても受話音声聞き取り可能な音量を確保する必要がある。しかし、通話回路系の利得を単純に増加すると、スピーカからマイクロホンへの音の回り込み（音響結合）により形成される閉ループの一巡利得が増加するため、ハウリングや不快なエコーを生じる。これらを抑圧するための手段として、送受話音声信号のレベルに応じた配分で通話路に損失量を挿入するボイススイッチ¹⁾（以下、VSと記す）が一般的に用いられてきた。しかし、この方式は、比較的安価なハードウェアで構成できる利点があるが、原理的に片方向通話方式であるため周囲騒音の音圧が高くなると受話音声途切れるという問題がある。一方、この問題を解決するために、通話路に損失を挿入せずにマイクロホンの出力信号に含まれるエコー成分のみを相殺する手段であるエコーキャンセラ²⁾（以下、ECと記す）による双方向通話の実現をねらった製品例もみられる。しかし、ECを搭載しただけでは、通話開始直後の未収束状

態においてハウリングや不快なエコーを抑圧することができない。また、エコー抑圧量は刻々と変化し、かつ周波数特性をもつため、すべての周波数帯域において閉ループ一巡利得を所定値以下に抑える必要がある。

これらの課題を解決するため、筆者らは、独自のエコーキャンセラ二つと、総挿入損失量を通話系の一巡伝達利得に応じて更新する総損失量適応型ボイススイッチ（以下、AVSと記す）を核として構成し、高音圧の周囲騒音環境においても二つのECを偏りなく収束させることにより短時間での総挿入損失量の低減を可能とする通話アルゴリズムを構築した。

2. インタホンの技術的課題

類似の技術は、ハンズフリー電話機、会議システムなどにもみられる³⁾が、インタホンでは次のような特有の技術的課題がある。

(1) 多様な設置環境への対応

宅内端末はリビング、寝室、浴室などの各部屋、また、宅外端末は外壁面やサインポスト、集合住宅における共用玄関部などと設置環境はさまざまなため、周囲の残響

* 情報機器事業本部 情報機器 R & D センター Research & Development Center, Information Equipment & Wiring Products Manufacturing Business Unit

特性がそれぞれ異なる。したがって、いかなる設置環境においてもハウリングなく安定した通話性能を確保する必要がある。

(2) 高レベルの周囲騒音への対応

図1に示すように、宅内の周囲騒音としてはTV，警報機，換気扇の回転音などが存在し，また宅外では，道路騒音，工事騒音，風切り音などが存在する。これらの騒音の音圧は設置環境や時間帯などによりさまざまであるが，最大で80 dB以上となる場合もある。このような騒音下においても受話音声の途切れがない，安定した通話性能を確保する必要がある。

(3) 相手側通話機の音響特性への対応

相手側通話機（子器）としては，図2に示すように，インタホン子器，「ロビーインタホン」，宅内端末（内線端末）などがあり，それぞれ，音響特性（マイクロホン，スピーカの感度など）が異なる。したがって，相手側通話機の特性も考慮したシステム設計が必要となる。

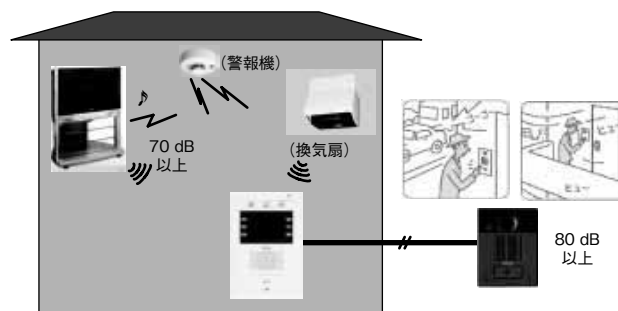


図1 住環境におけるさまざまな周囲騒音



図2 相手側通話端末の種類

2.1 構築したアルゴリズム

構築したアルゴリズムをハンズフリーインタホン親機に搭載したときの子器との間の通話システム構成図を図3に示す。図3においては，親機－子器間の音声通信はアナログ2線伝送方式（ペア線を介して双方向に信号を伝送する方式）によるものとしており，Hは2線－4線変換回路である。構築したアルゴリズムは親機の通話回路内に搭載されるものであり，二つのEC（EC1，EC2），およびAVSを核として構成される。また，T-ATT，R-ATTはそれぞれ送話路および受話路へ挿入される減衰器を表し，その利得 G_T ， G_R は挿入利得制御部によって刻々と更新される。ここ

で，AVSから親機のスピーカ，マイクロホン側をみたとき，R-ATTからの出力点（Rout）からスピーカに至り，音響結合路を経てマイクロホンへ到達し，さらに送話路を経てT-ATTへの入力点（Tin）へ帰還する回り込み経路 α が存在する。また，AVSから子器側をみたときには，T-ATTからの出力点（Tout）からHでの回り込み（ $\beta 1$ ），および子器での音響結合による回り込み（ $\beta 2$ ）を経てR-

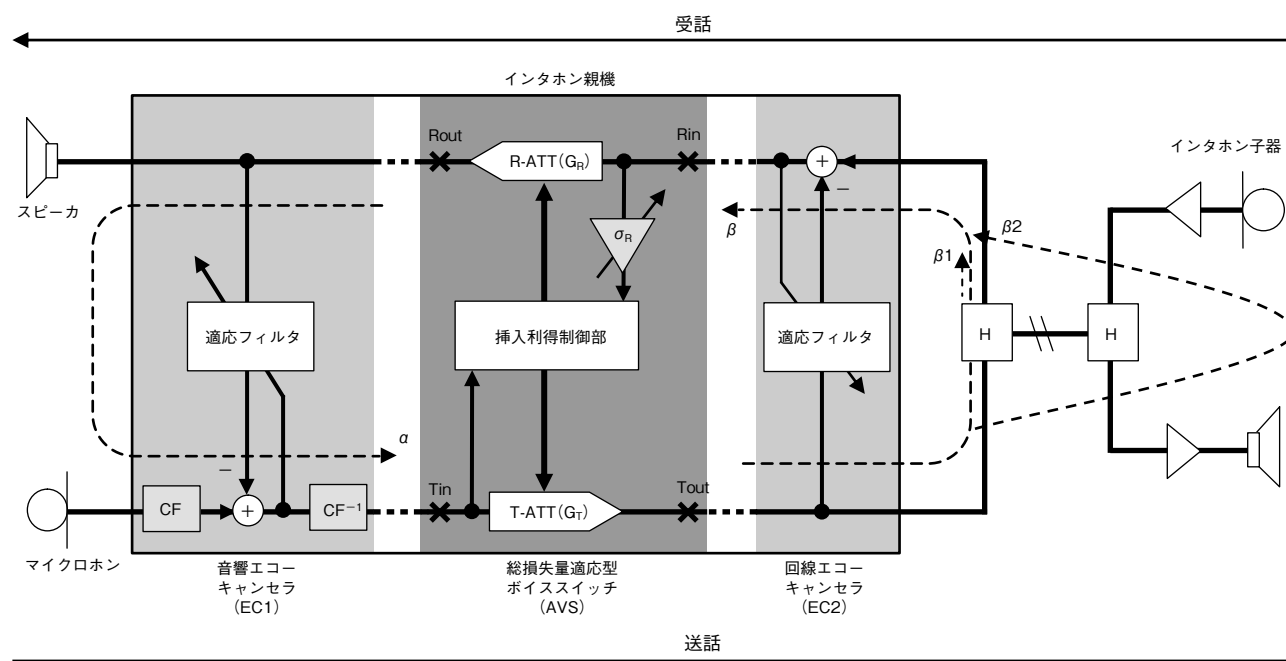


図3 構築したアルゴリズム

ATTへの入力点 (Rin) へ帰還する回り込み経路 β が存在する。二つのECは、送受話信号に対して損失を与えずにこれらの回り込み α 、 β の利得のみを低減する。これにより、閉ループ ($\alpha \rightarrow T-ATT \rightarrow \beta \rightarrow R-ATT$) の一巡利得が減少するためハウリングに対する余裕度 (以下、ハウリングマージンと記す) が増加する。その結果、二つのECが収束するのに伴ってAVSにおける総損失量 ($G_T \cdot G_R$)⁻¹ が低減するため、双方向同時通話性能を実現することができる。

2.2 VSにおける総損失量制御

一般に、VSは通話状態の推定結果に基づき二つの減衰器の利得 G_T 、 G_R を調整する。すなわち、親機側話者が発声状態と推定した際には G_T を最大値 (通常は 0 dB)、 G_R を最小値 ($-L_t$ dB) とし、子器側話者が発声状態と推定した際には G_T を最小値、 G_R を最大値とする。推定された通話状態に対する G_T 、 G_R の遷移を図4に示す。

従来のVSでは総損失量 L_t は時刻 n に依存せず一定で

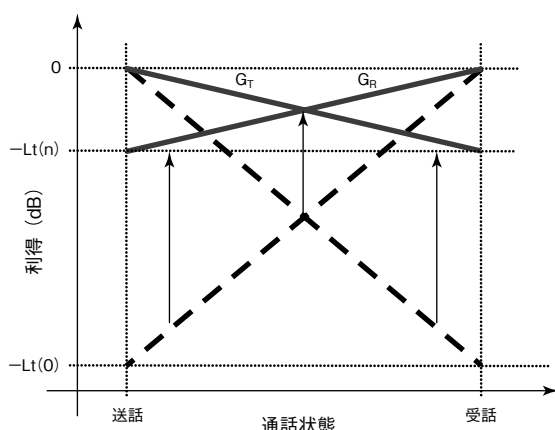


図4 挿入利得 G_T 、 G_R の遷移

あり、 $L_t(n) = L_t(0)$ としていた。ここで、 $L_t(0)$ は L_t の初期値であり、ECが未収束の通話開始直後においてもハウリングを生じないために必要な損失量を示しており、子器も含めた通話機筐体の音響特性、設置環境における残響特性などをもとに設計される値である。

一方、AVSにおいては、刻々と変化する二つの回り込み α 、 β の利得を推定し、それらの値に基づいて L_t を更新することに特徴がある。すなわち、通話開始後、しばらく会話をを行うと二つのECが収束し、回り込み α 、 β が低減する。その結果、通話系に形成されている閉ループの一巡利得 G_{LP} が低減する。ここで、 G_{LP} は式 (1) から求められるものであるが、 α 、 β が周波数特性をもつことから G_{LP} も周波数特性を有することになり、 G_{LP} が 1 (= 0 dB) を超える周波数が存在するとその周波数でハウリングが生じる。

$$G_{LP}(f) = |\alpha(f)| \cdot G_T \cdot |\beta(f)| \cdot G_R = L_t^{-1} |\alpha(f)| \cdot |\beta(f)| \quad (1)$$

したがって、VSの L_t は次式を満足する必要がある。

$$L_t > \text{MAX}[|\alpha(f)| \cdot |\beta(f)|] \quad (2)$$

ただし、 f は周波数を表す。

ここで、話者が通話機に近づくなどのエコー経路の変動や高音圧の周囲騒音などの外乱によってECの抑圧量が低下した場合には、式 (2) の右辺は大きく変動する。したがって、 L_t の算出の際には、ある程度のハウリングマージンを考慮して大きめの値にする必要がある。一方、ECが正常に収束した場合には、速やかに L_t を低減し、双方向同時通話状態へ移行させなければならない。これらを両立させるためには、 $\text{MAX}[|\alpha(f)| \cdot |\beta(f)|]$ を精度良く推定する必要がある。筆者らが提案するAVSにおいては、次節で述べるアルゴリズムによって $\text{MAX}[|\alpha(f)| \cdot |\beta(f)|]$ を求める。

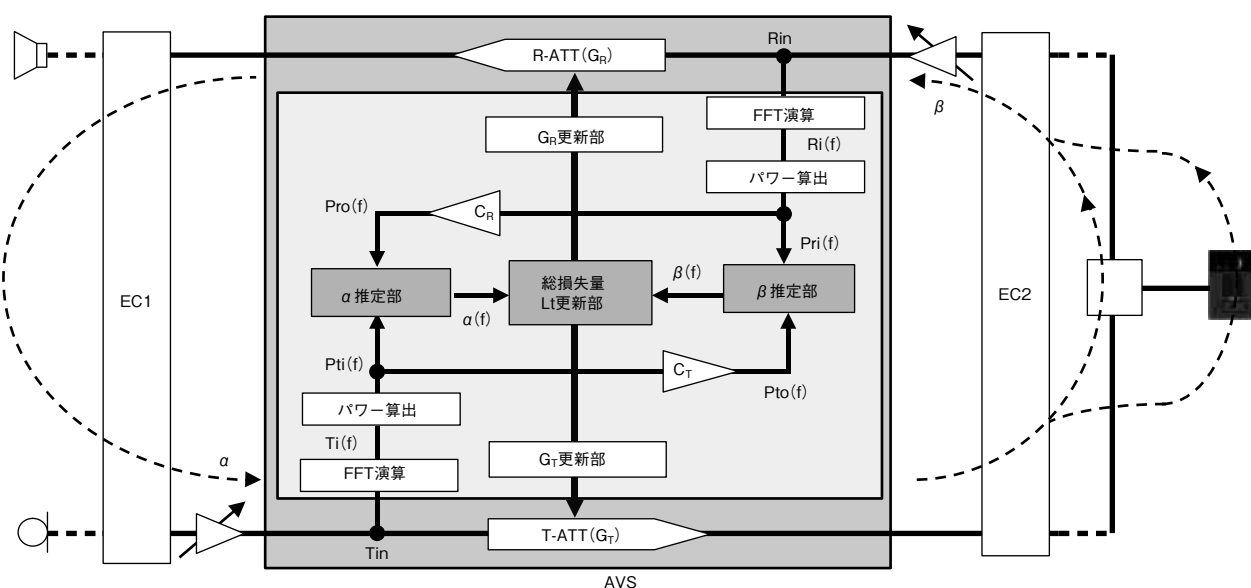


図5 総損失量 L_t の更新処理

2.3 回り込み利得 α 、 β の推定

音響側回り込み利得 α 、回線側回り込み利得 β の推定処理部、およびこれらの推定値に基づく総損失量 L_t の更新処理部に着目したブロック図を図5に示す。本アルゴリズムにおいては、まず、送話路および受話路に挿入される減衰器への入力信号 T_{in} 、 R_{in} に対してFFT演算を施し、それぞれのスペクトル $T_i(f)$ 、 $R_i(f)$ を求める。次に、これらのスペクトルに対して瞬時パワー演算を行い、 $P_{ti}(f)$ および $P_{ri}(f)$ を求める。さらに、これらの瞬時パワーに対して C_T および C_R を乗算し、 $P_{to}(f)$ 、 $P_{ro}(f)$ を算出する。ただし、 C_T 、 C_R は通話路への挿入利得 G_T 、 G_R とともに遷移する乗算係数を表し、 $C_T \cdot C_R = -L_t(0)$ の関係をつねに保つ。すなわち、 C_T 、 C_R は図4における破線に沿って通話状態に応じて遷移する。 α 推定部および β 推定部においては、このようにして得られた $P_{ti}(f)$ 、 $P_{ri}(f)$ 、 $P_{to}(f)$ 、 $P_{ro}(f)$ を用いて、次式のように $\alpha(f)$ 、 $\beta(f)$ を求める。

$$\alpha(f) = \frac{P_{ti}(f)}{P_{ro}(f)} \quad (3)$$

$$\beta(f) = \frac{P_{ri}(f)}{P_{to}(f)} \quad (4)$$

次に、 L_t 更新部においては、通話帯域内における $\alpha(f)$ と $\beta(f)$ の積の最大値 $\text{MAX}[\alpha(f) \cdot \beta(f)]$ とハウリングマージンの所望値（ $-HM$ dB）との比較により、 L_t を更新する。すなわち、 L_t は式（5）、式（6）で与えられる。なお、 HM はエコーキャンセラが未収束時などにおいてもハウリングを生じずに安定した通話を確保するためにあらかじめ設計したハウリングマージンを表す。

$$\begin{aligned} & \text{(if } \text{MAX}[\alpha(f) \cdot \beta(f)] > -HM \text{ dB)} \\ & L_t = \text{MAX}[\alpha(f) \cdot \beta(f)] + HM \text{ dB} \end{aligned} \quad (5)$$

$$\begin{aligned} & \text{(else)} \\ & L_t = 0 \text{ dB} \end{aligned} \quad (6)$$

図6にこれらの演算を模式的に示す。いま、FFTのポイント数を8点とすると、 $\alpha(f)$ 、 $\beta(f)$ はナイキスト周波数（サンプリング周波数の1/2）以下の帯域を4分割したときの各帯域におけるスペクトルを表す。図6においては、 $\alpha_1 \sim \alpha_4$ および $\beta_1 \sim \beta_4$ がこれに相当する。

次に、各帯域ごとに $\alpha_i \cdot \beta_i$ （ $i=1 \sim 4$ ）を算出し、その最大値を求める。図6に示す例では、最大値は $\alpha_2 \cdot \beta_2$ である。この値が $-HM$ dBを超えている場合には、その超過分が所要総損失量 L_t となる。このような演算により L_t を更新していくことにより、つねにハウリングマージンを設計値（ HM dB）に保つことができる。

2.4 エコー抑圧量の周波数特性の改善

ECにより得られる抑圧量はエコー経路の特性に依存し、複雑な周波数特性を有している。さらに通話機筐体にお

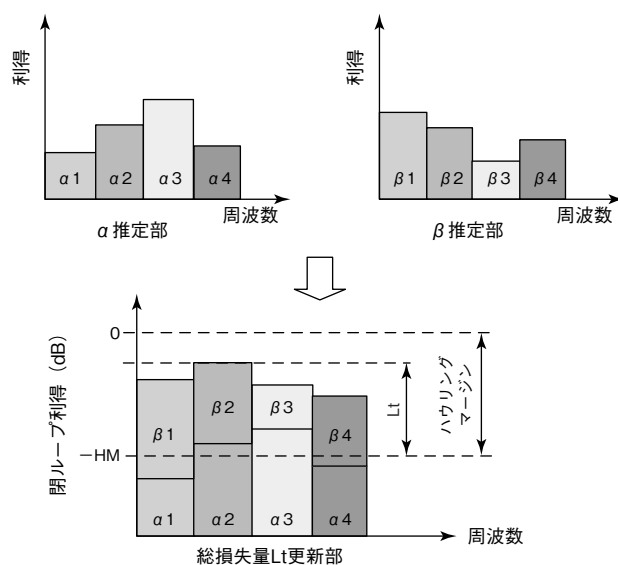


図6 所要総損失量 L_t の高精度推定

る音響結合特性によっては十分なエコー抑圧量を得られない周波数帯域が存在する可能性がある。一方、前節で述べたように、ハウリングマージンは $|\alpha(f) \cdot \beta(f)|$ が最大となる周波数における閉ループ一巡利得により定まる。したがって、通話帯域内においてエコー抑圧量が得られない帯域がある場合には、ハウリングマージンの増大が期待できない。この問題を改善するため、筆者らは、図7に示すようにマイクロホンからの出力信号 s に対して補正フィルタ CF （以下、 CF と記す）を挿入することにより、エコー経路の周波数特性を補正する方式を考案した。

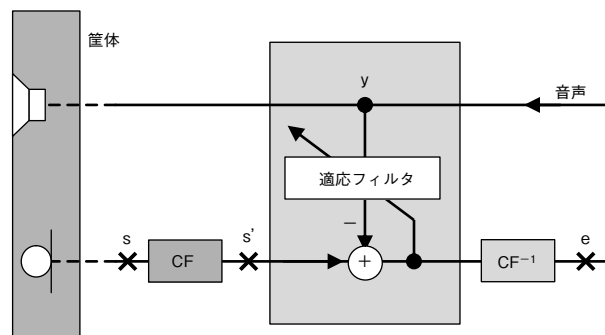
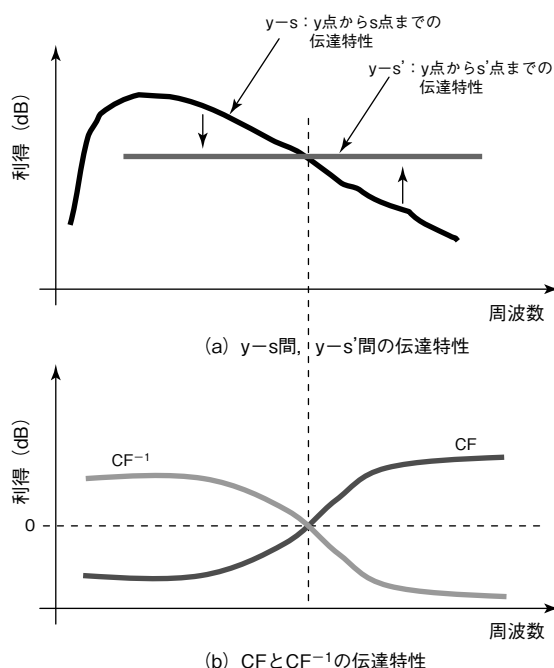


図7 エコー特性補正フィルタ CF の挿入

ここで、図8に示すように、 CF はエコー経路の周波数特性において利得が低い帯域を強調する特性をもつ。すなわち、エコー経路の周波数特性を平坦化することにより、全帯域で一様にエコー抑圧量を得ることをねらいとする。また、フィルタ CF^{-1} は CF の逆フィルタであり、 CF による強調処理を相殺するような周波数特性をもつ。このような構成により、マイクロホンにより集音される通話者の音声（すなわち、送話音声）に対しては周波数特性を変化さ



せずに相手側端末に伝送することができる。

図9は本方式の有効性を検証するために行ったエコー抑圧特性の実測結果を示す。測定手順は以下のとおりである。

- (1) ECを動作させずに、受話信号 y として白色雑音を入力したときに e 点に出力される信号のパワースペクトルを測定（学習前の e のスペクトル）。
- (2) y をテスト音声としてECを2秒間程度、適応動作。
- (3) ECの学習を止め（フィルタ係数は保持）、 y を白色雑音に戻し、 e 点に出力される信号のパワースペクトルを測定（学習後の e のスペクトル）。

なお、本測定においては、CFは2 kHz以上の帯域を強調する特性とした。従来方式（a）に対し、提案方式（b）ではCFにより強調処理を施した高域においても良好なエコー抑圧量（学習前と学習後の差）を得ており、本方式の有効性が確認できている。

2.5 周囲騒音下における通話性能向上

図3に示した通話系において、子器側の周囲騒音レベルが高い場合、以下の問題について考慮する必要がある。

- (1) VSを送話状態に移せにくい。
- (2) EC2の収束が進まず、 β の利得を低減できない。

ここで、(1)の解決は従来のVSを用いて構成される片方向通話型ハンズフリー通話機と共通の課題である。これは、VSの挿入利得制御部で挿入利得の配分を決定するための参照点(Tin, Rin)において、「送話信号パワー」≪「受話信号パワー」となるためである。さらに、図10を用

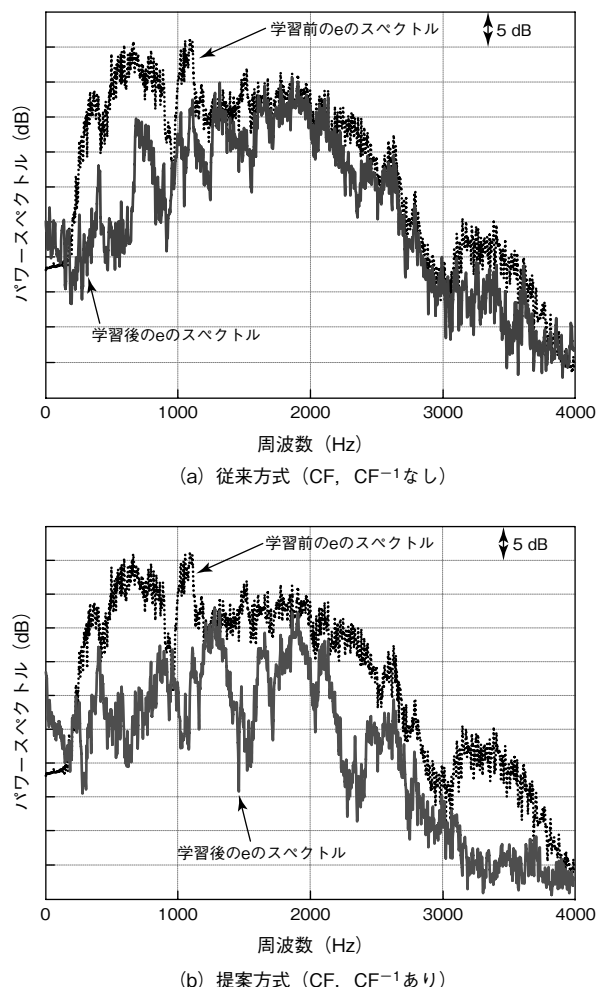


図9 エコー抑圧特性の実測結果

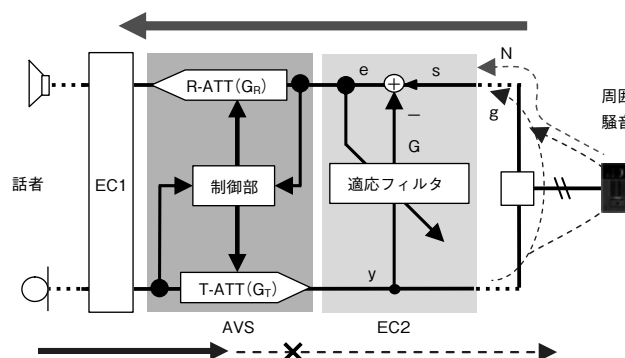
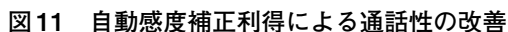


図10 周囲騒音により生じる課題

いて(2)に関して説明する。

- (a) G_T からの出力信号 y のレベルが小さくなる。
- (b) 回線側での回り込みによるエコー g が小さくなり、子器からの入力信号（周囲騒音） N との関係が $g \ll N$ となる。
- (c) 減算器への入力信号 s は N とほぼ等しくなり、 N は y とは相関がない外乱成分であるため収束しない。

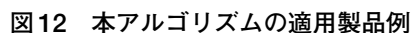
以上のように、(2)は(1)の結果として2次的に生じるものであるため、(1)を解決することが本質的な対策である。これに対して筆者らは、図11に示すように受話信号参照点(Rin)から挿入損失量制御部までの経路上に自動感度補正利得 σ_R を挿入する方式を考案した。



σ_R の利得は、初期値は0 dBに設定されている。また、通話状態においては、親機側および子器側の周囲騒音レベルをつねに推定しており、それらの推定値に基づき σ_R を更新する。すなわち、親機側に対して子器側の周囲騒音レベルがきわめて高い場合、AVSの制御部に入力される送受話信号は、「送話信号パワー」 $\ll$ 「受話信号パワー」となっており、通話状態の推定結果が受話側に片寄ってしまう。自動補正利得 σ_R はこの差を補正するものであり、このような状態において利得を低減することにより、「受話信号パワー」を実際よりも小さく見積もることとなり、推定される通話状態が受話側に片寄るのを防止することが可能となる。これにより、親機側の話者が発声すればAVSは送話状態となり、EC2に対して十分な大きさの参照信号 y を与えることができ、収束を早めることができる。

構築したアルゴリズムを高性能マイクロコンピュータまたはDSP（Digital Signal Processor）に実装し、図12に示す集合住宅系セキュリティインタホン「Smart MONION」および高機能玄関番「HOMITY」に搭載している。

いずれもハンズフリー通話機として十分な音量、音質を有しており、ハウリングや不快なエコーを生じることなく、80 dB以上の周囲騒音環境下においても言葉切れのない良好な通話性能を確認している。



通話機筐体ごとに異なる音響結合特性に依存せずにエコー抑圧量の周波数特性が平坦となるエコーキャンセラと、刻々と変化する音響結合利得およびエコー抑圧量に追従して通話路に挿入する損失量の総和（総損失量）を逐次更新することを特徴とする総損失量適応型ボイススイッチを用いて、高音圧の周囲騒音環境においても短時間で総挿入損失量を低減して双方向同時通話状態を実現するハンズフリー通話アルゴリズムを考案した。

本技術の適用により、業界初の双方向同時通話型ハンズフリーインタホンを実現した。

本アルゴリズムの特徴をまとめると、以下の3点が挙げられる。

- (1) FFTを用いた周波数領域での回り込み利得の算出による所要総損失量の高精度推定。
- (2) 筐体ごとに異なる音響経路の周波数特性の補正によるエコーキャンセラの抑圧性能の向上。
- (3) 周囲騒音下におけるボイススイッチの自動感度補正処理によるエコーキャンセラの収束の高速化。

また、本アルゴリズムを当社製インタホンに搭載し、ハンズフリー通話機として十分な音量、音質を有し、良好な通話性能が得られることを確認している。

*参考文献

- 1) 加藤 邦紘, 松田 亮一: 音声スイッチ回路の動作解析と設計法, 電子情報通信学会誌, Vol. 56-A, No. 5, p. 293 (1973)
- 2) 辻井 重男: エコーキャンセラ技術, 日刊工業技術センター (1986)
- 3) 古川 博基, 金森 丈郎, 茨木 悟, 直野 博之: ハウリング検出機能を有したエコーキャンセラ型拡声通話回路, 電子情報通信学会春季全国大会, 分冊 1, p. 338 (1991)

◆執筆者紹介



福島 実

情報機器 R & D センター
電気通信主任技術者



吉田 恵一

情報機器 R & D センター